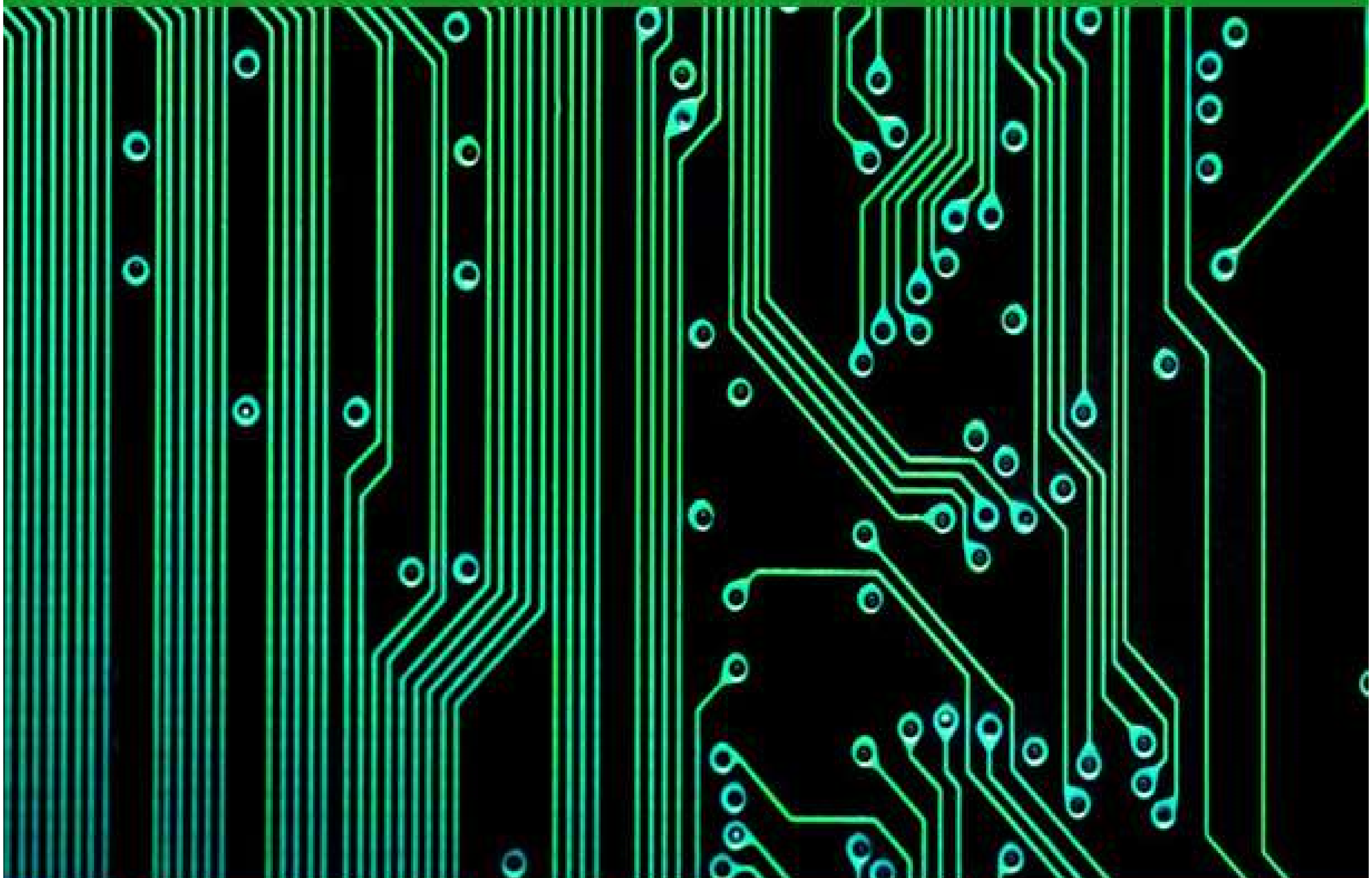
# Handbook of Electric Current and Electronics



Thomasena Brennan

Leo Goldberg

First Edition, 2012

ISBN 978-81-323-0901-7





*Published by:*
**Academic Studio**
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

# Table of Contents

## Chapter-1

# Electric Current

**Electric current** is a flow of electric charge through a medium. This flowing electric charge is typically carried by moving electrons in a conductor such as wire. It can also be carried by ions in an electrolyte, or, it can be carried by both ions and electrons in a plasma.

The SI unit for measuring the rate of flow of electric charge is the ampere, which is charge flowing through some surface at the rate of one coulomb per second. Electric current is measured using an ammeter.

### *Symbol*

The conventional symbol for current is *I*, which may seem puzzling. It originates from the French phrase *intensité de courant*, or in English *current intensity*. This phrase is frequently used when discussing the value of an electric current, especially in older texts; modern practice often shortens this to simply *current* but *current intensity* is still used in many recent textbooks. The *I* symbol was used by André-Marie Ampère himself, after whom the unit of electric current is named, in formulating the eponymous Ampère's force law which he discovered in 1820. The notation travelled from France to England where it became standard, although at least one journal did not change from using *C* to *I* until 1896.

### *Conduction mechanisms in various media*

In metallic solids, electricity flows by means of electrons, from lower to higher electrical potential. In other media, any stream of charged objects may constitute an electric current. To provide a definition of current that is independent of the type of charge carriers flowing, *conventional current* is defined to flow in the same direction as positive charges. So in metals where the charge carriers (electrons) are negative, conventional current flows in the opposite direction as the electrons. In conductors where the charge carriers are positive, conventional current flows in the same direction as the charge carriers.

In a vacuum, a beam of ions or electrons may be formed. In other conductive materials, the electric current is due to the flow of both positively and negatively charged particles at the same time. In still others, the current is entirely due to positive charge flow. For example, the electric currents in electrolytes are flows of electrically charged atoms (ions), which exist in both positive and negative varieties. In a common lead-acid electrochemical cell, electric currents are composed of positive hydrogen ions (protons) flowing in one direction, and negative sulfate ions flowing in the other. Electric currents in sparks or plasma are flows of electrons as well as positive and negative ions. In ice and in certain solid electrolytes, the electric current is entirely composed of flowing ions. In a semiconductor it is sometimes useful to think of the current as due to the flow of positive "holes" (the mobile positive charge carriers that are places where the semiconductor crystal is missing a valence electron). This is the case in a p-type semiconductor.

## Metals

A solid conductive metal contains mobile, or free electrons, originating in the conduction electrons. These electrons are bound to the metal lattice but no longer to any individual atom. Even with no external electric field applied, these electrons move about randomly due to thermal energy but, on average, there is zero net current within the metal. Given a surface through which a metal wire passes, the number of electrons moving from one side to the other in any period of time is on average equal to the number passing in the opposite direction. As George Gamow put in his science-popularizing book, *One, Two, Three...Infinity* (1947), "The metallic substances differ from all other materials by the fact that the outer shells of their atoms are bound rather loosely, and often let one of their electrons go free. Thus the interior of a metal is filled up with a large number of unattached electrons that travel aimlessly around like a crowd of displaced persons. When a metal wire is subjected to electric force applied on its opposite ends, these free electrons rush in the direction of the force, thus forming what we call an electric current."

When a metal wire is connected across the two terminals of a DC voltage source such as a battery, the source places an electric field across the conductor. The moment contact is made, the free electrons of the conductor are forced to drift toward the positive terminal under the influence of this field. The free electrons are therefore the charge carrier in a typical solid conductor. For an electric current of 1 ampere, 1 coulomb of electric charge (which consists of about $6.242 \times 10^{18}$ elementary charges) drifts every second through any plane through which the conductor passes.

For a steady flow of charge through a surface, the current $I$ in amperes can be calculated with the following equation:

$$I = \frac{Q}{t},$$

where $Q$ is the electric charge transferred through the surface over some time $t$. If $Q$ and $t$ are measured in coulombs and seconds respectively, $I$ is in amperes.

More generally, electric current can be represented as the rate at which charge flows through a given surface as:

$$I = \frac{\mathrm{d}Q}{\mathrm{d}t}\,.$$

## Electrolytes

Electric currents in electrolytes are flows of electrically charged particles (ions). For example, if an electric field is placed across a solution of $Na^+$ and $Cl^-$ (and conditions are right) the sodium ions move towards the negative electrode (cathode), while the chloride ions move towards the positive electrode (anode). Reactions take place at both electrode surfaces, absorbing each ion.

Water-ice and certain solid electrolytes called proton conductors contain positive hydrogen ions or "protons" which are mobile. In these materials, electric currents are composed of moving protons, as opposed to the moving electrons found in metals.

In certain electrolyte mixtures, brightly-colored ions form the moving electric charges. The slow migration of these ions means that the current is visible.

## Gases and plasmas

In air and other ordinary gases below the breakdown field, the dominant source of electrical conduction is via a relatively small number of mobile ions produced by radioactive gases, ultraviolet light, or cosmic rays. Since the electrical conductivity is low, gases are dielectrics or insulators. However, once the applied electric field approaches the breakdown value, free electrons become sufficiently accelerated by the electric field to create additional free electrons by colliding, and ionizing, neutral gas atoms or molecules in a process called avalanche breakdown. The breakdown process forms a plasma that contains a significant number of mobile electrons and positive ions, causing it to behave as an electrical conductor. In the process, it forms a light emitting conductive path, such as a spark, arc or lightning.

Plasma is the state of matter where some of the electrons in a gas are stripped or "ionized" from their molecules or atoms. A plasma can be formed by high temperature, or by application of a high electric or alternating magnetic field as noted above. Due to their lower mass, the electrons in a plasma accelerate more quickly in response to an electric field than the heavier positive ions, and hence carry the bulk of the current.

## Vacuum

Since a "perfect vacuum" contains no charged particles, it normally behaves as perfect insulator. However, metal electrode surfaces can cause a region of the vacuum to become conductive by injecting free electrons or ions through either field electron emission or thermionic emission. Thermionic emission occurs when the thermal energy exceeds the metal's work function, while field electron emission occurs when the electric field at the surface of the metal is high enough to cause tunneling, which results in the ejection of free electrons from the metal into the vacuum. Externally heated electrodes are often used to generate an electron cloud as in the filament or indirectly heated cathode of vacuum tubes. Cold electrodes can also spontaneously produce electron clouds via thermionic emission when small incandescent regions (called **cathode spots** or **anode spots**) are formed. These are incandescent regions of the electrode surface that are created by a localized high current flow. These regions may be initiated by field electron emission, but are then sustained by localized thermionic emission once a vacuum arc forms. These small electron-emitting regions can form quite rapidly, even explosively, on a metal surface subjected to a high electrical field. Vacuum tubes and sprytrons are some of the electronic switching and amplifying devices based on vacuum conductivity.

### *Current density and Ohm's law*

Current density is a measure of the density of an electric current. It is defined as a vector whose magnitude is the electric current per cross-sectional area. In SI units, the current density is measured in amperes per square meter.

$$I = \overrightarrow{J} \cdot \overrightarrow{A}$$

where *I* is current in the conductor, **J** is the current density, and **A** is the cross-sectional area. The dot product of the two vector quantities (**A** and **J**) is a scalar that represents the electric current.

Current density (current per unit area) *J* in a material is proportional to the conductivity *σ* and electric field *E* in the medium:

$$J = \sigma E$$

Instead of conductivity, reciprocal quantity called resistivity *ρ*, can be used:

$$J = \frac{E}{\rho}$$

Conduction in semiconductor devices may occur by a combination of electric field (drift) and diffusion, which is proportional to diffusion constant *D* and charge density $\alpha_q$. The current density is then:

$$J = \sigma E + Dq\nabla n,$$

with *q* being the elementary charge and *n* the electron density. The carriers move in the direction of decreasing concentration, so for electrons a positive current results for a positive density gradient. If the carriers are holes, replace electron density *n* by the negative of the hole density *p*.

In linear anisotropic materials, $\sigma$, $\rho$ and $D$ are tensors.

In linear materials such as metals, and under low frequencies, the current density across the conductor surface is uniform. In such conditions, Ohm's law states that the current is directly proportional to the potential difference between two ends (across) of that metal (ideal) resistor (or other ohmic device):

$$I = \frac{V}{R},$$

where $I$ is the current, measured in amperes; $V$ is the potential difference, measured in volts; and $R$ is the resistance, measured in ohms. The letter $I$ stands for the German word, "Intensität" meaning "Intensity". For alternating currents, especially at higher frequencies, skin effect causes the current to spread unevenly across the conductor cross-section, with higher density near the surface, thus increasing the apparent resistance.

### *Drift speed*

The mobile charged particles within a conductor move constantly in random directions, like the particles of a gas. In order for there to be a net flow of charge, the particles must also move together with an average drift rate. Electrons are the charge carriers in metals and they follow an erratic path, bouncing from atom to atom, but generally drifting in the opposite direction of the electric field. The speed at which they drift can be calculated from the equation:

$$I = nAvQ,$$

where

$I$ is the electric current
$n$ is number of charged particles per unit volume (or charge carrier density)
$A$ is the cross-sectional area of the conductor
$v$ is the drift velocity, and
$Q$ is the charge on each particle.

Electric currents in solids typically flow very slowly. For example, in a copper wire of cross-section 0.5 $mm^2$, carrying a current of 5 A, the *drift velocity* of the electrons is on the order of a millimetre per second. To take a different example, in the near-vacuum inside a cathode ray tube, the electrons travel in near-straight lines at about a tenth of the speed of light.

Any accelerating electric charge, and therefore any changing electric current, gives rise to an electromagnetic wave that propagates at very high speed outside the surface of the conductor. This speed is usually a significant fraction of the speed of light, as can be deduced from Maxwell's Equations, and is therefore many times faster than the drift velocity of the electrons. For example, in AC power lines, the waves of electromagnetic energy propagate through the space between the wires, moving from a source to a distant load, even though the electrons in the wires only move back and forth over a tiny distance.

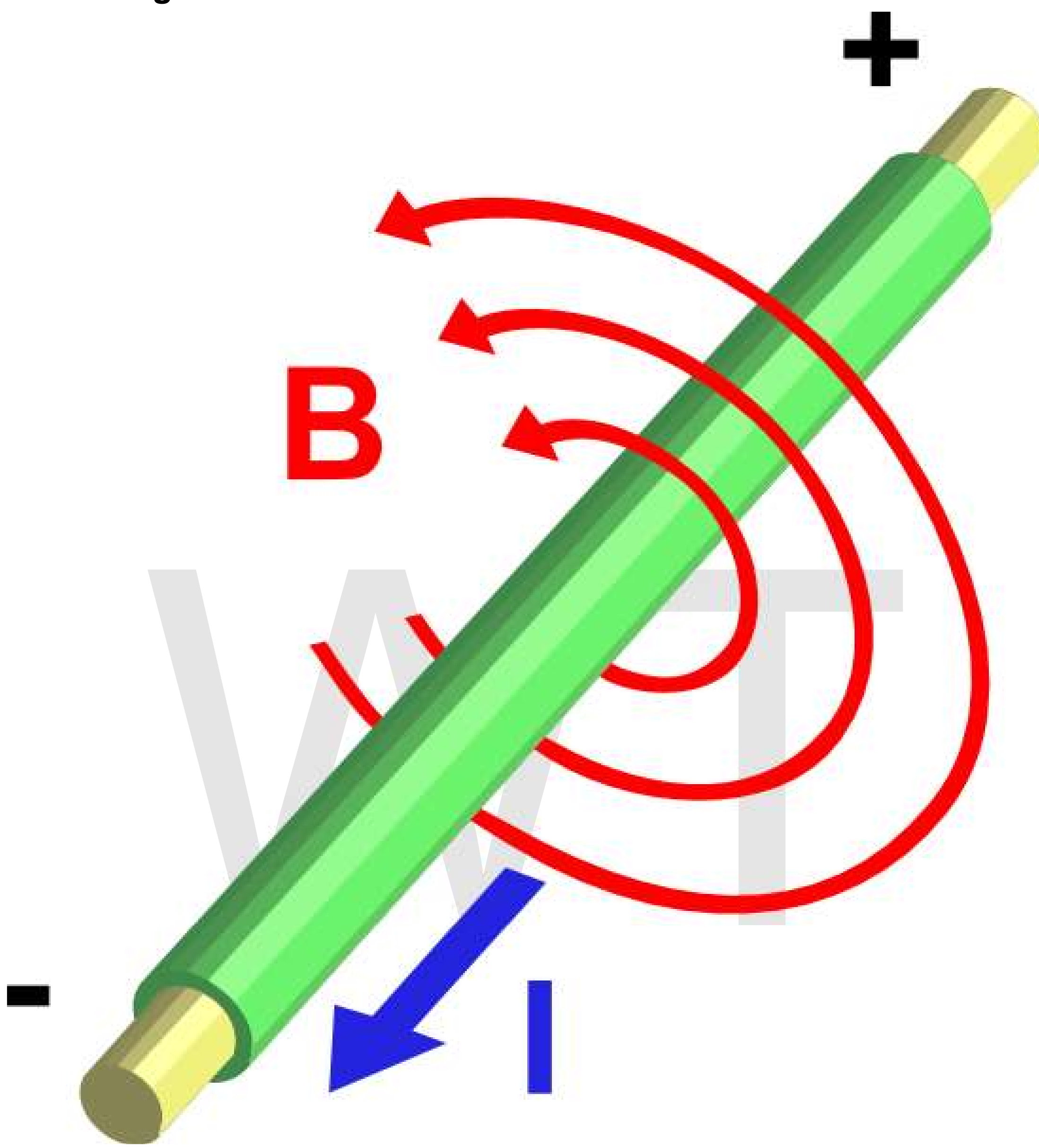
The ratio of the speed of the electromagnetic wave to the speed of light in free space is called the velocity factor, and depends on the electromagnetic properties of the conductor and the insulating materials surrounding it, and on their shape and size.

The magnitudes (but, not the natures) of these three velocities can be illustrated by an analogy with the three similar velocities associated with gases.

- The low drift velocity of charge carriers is analogous to air motion; in other words, winds.
- The high speed of electromagnetic waves is roughly analogous to the speed of sound in a gas (these waves move through the medium much faster than any individual particles do)
- The random motion of charges is analogous to heat - the thermal velocity of randomly vibrating gas particles.

This analogy is extremely simplistic and incomplete: The rapid propagation of a sound wave doesn't impart any change in the air molecules' drift velocity, whereas EM waves do carry the energy to propagate the actual current at a rate which is much, much higher than the electrons' drift velocity. To illustrate the difference: The sound and the change in the air's drift velocity (the force of the wind gust) cross distance at rates equaling the speeds of sound and of mechanical transmission of force (**not higher** than rate of drift velocity); while a change in an EM field and the **change** in current (electrons' drift velocity) both propagate across distance at rates **much higher** than the actual drift velocity. You can hear wind much earlier than the force of the gust reaches you, but you don't observe a change in an EM field earlier than you can observe the change of current.

### *Electromagnetism*



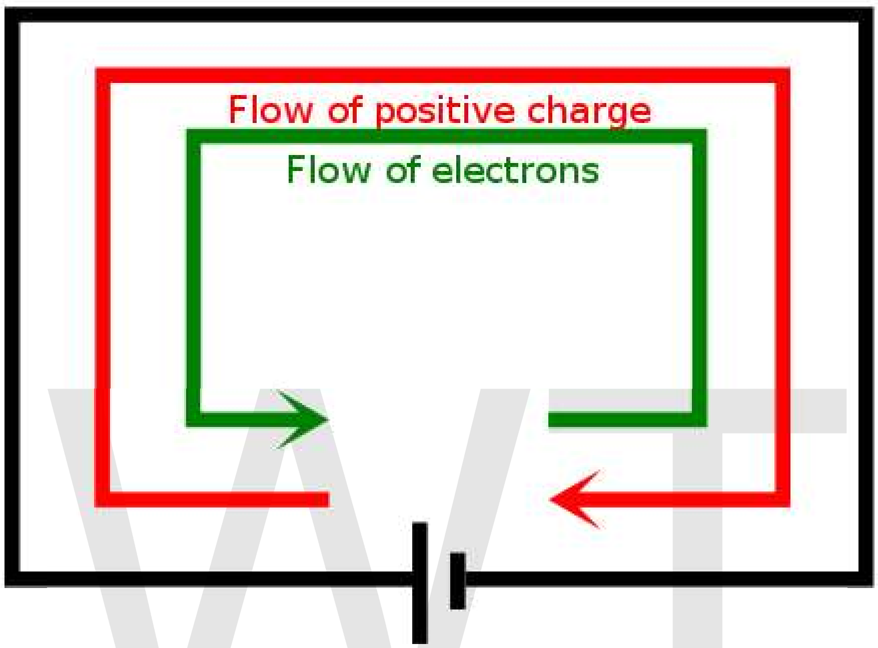

According to Ampère's law, an electric current produces a magnetic field.

Electric current produces a magnetic field. The magnetic field can be visualized as a pattern of circular field lines surrounding the wire.

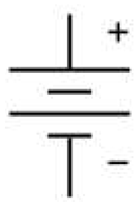
Electric current can be directly measured with a galvanometer, but this method involves breaking the electrical circuit, which is sometimes inconvenient. Current can also be measured without breaking the circuit by detecting the magnetic field associated with the current. Devices used for this include Hall effect sensors, current clamps, current transformers, and Rogowski coils.

The theory of Special Relativity allows one to transform the magnetic field into a static electric field for an observer moving at the same speed as the charge in the diagram. The amount of current is particular to a reference frame.

### Conventions




The electrons, the charge carriers in an electrical circuit, flow in the opposite direction of the *conventional* electric current.




The symbol for a battery in a circuit diagram.

A flow of positive charges gives the same *electric* current as a flow of negative charges in the opposite direction. Since current can be the flow of either positive or negative charges, or both, a convention for the direction of current which is independent of the type of charge carriers is needed. Therefore the direction of *conventional current* is defined to be the direction of the flow of positive charges.

In metals, which make up the wires and other conductors in most electrical circuits, the positive charges are immobile, and only the negatively charged electrons flow. Because the electron carries negative charge, the *electron* motion in a metal conductor is in the direction opposite to that of conventional (or *electric*) current.

### Reference direction

When analyzing electrical circuits, the actual direction of current through a specific circuit element is usually unknown. Consequently, each circuit element is assigned a current variable with an arbitrarily chosen *reference direction*. When the circuit is solved, the circuit element currents may have positive or negative values. A negative value means that the actual direction of current through that circuit element is opposite that of the chosen reference direction. In electronic circuits the reference current directions are usually chosen so that all currents flow toward ground. This often matches conventional current direction, because in many circuits the power supply voltage is positive with respect to ground.

## *Occurrences*

Natural examples include lightning and the solar wind, the source of the polar auroras (the aurora borealis and aurora australis). The artificial form of electric current is the flow of conduction electrons in metal wires, such as the overhead power lines that deliver electrical energy across long distances and the smaller wires within electrical and electronic equipment. In electronics, other forms of electric current include the flow of electrons through resistors or through the vacuum in a vacuum tube, the flow of ions inside a battery or a neuron, and the flow of holes within a semiconductor.

## *Current Measurement*

Current can be measured using an ammeter.

At the circuit level there are various techniques that can be used to measure current:

- Shunt resistor
- Hall effect current sensor transducers
- Transformer(however dc cannot be measured)
- Magnetoresistive Field Sensors

**Chapter-2**

# Passivity (Engineering) and Voltage Source

## Passivity (engineering)

**Passivity** is a property of engineering systems, used in a variety of engineering disciplines, but most commonly found in analog electronics and control systems. A **passive component**, depending on field, may be either a component that consumes (but does not produce) energy (thermodynamic passivity), or a component that is incapable of power gain (incremental passivity).

A component that is not passive is called an **active component**. An electronic circuit consisting entirely of passive components is called a passive circuit (and has the same properties as a passive component). Used without qualifier, the term **passive** is ambiguous. Typically, analog designers use this term to refer to **incrementally passive** components and systems, while control systems engineers will use this to refer to **thermodynamically passive** ones.

### *Thermodynamic passivity*

In control systems and circuit network theory, a passive component or circuit is one that consumes energy, but does not produce energy. Under this methodology, voltage and current sources are considered active, while resistors, transistors, tunnel diodes, glow tubes, capacitors, metamaterials and other dissipative and energy-neutral components are considered passive. Circuit designers will sometimes refer to this class of components as dissipative, or thermodynamically passive.

While many books give definitions for passivity, many of these contain subtle errors in how initial conditions are treated (and, occasionally, the definitions do not generalize to all types of nonlinear time-varying systems with memory). Below is a correct, formal definition, taken from Wyatt et al. (which also explains the problems with many other definitions). Given an $n$-port $R$ with a state representation $S$, and initial state $x$, define available energy $E_A$ as:

$$E_A(x) = \sup_{x \to T \geq 0} \int_0^T -\langle v(t), i(t) \rangle \, \mathrm{d}t$$

where the notation $\sup_{x \to T \geq 0}$ indicates that the supremum is taken over all $T \geq 0$ and all admissible pairs $\{v(\cdot), i(\cdot)\}$ with the fixed initial state $x$ (e.g., all voltage–current trajectories for a given initial condition of the system). A system is considered passive if $E_A$ is finite for all initial states $x$. Otherwise, the system is considered active. Roughly speaking, the inner product $\langle v(t), i(t) \rangle$ is the instantaneous power (e.g., the product of voltage and current), and $E_A$ is the upper bound on the integral of the instantaneous power (i.e., energy). This upper bound (taken over all $T \geq 0$) is the *available energy* in the system for the particular initial condition $x$. If, for all possible initial states of the system, the energy available is finite, then the system is called *passive.*

### Incremental passivity

In circuit design, informally, passive components refer to ones that are not capable of power gain. Under this definition, passive components include capacitors, inductors, resistors, diodes, transformers, voltage sources, and current sources. They exclude devices like transistors, vacuum tubes, relays, tunnel diodes, and glow tubes. Formally, for a memoryless two-terminal element, this means that the current–voltage characteristic is monotonically increasing. For this reason, control systems and circuit network theorists refer to these devices as locally passive, incrementally passive, increasing, monotone increasing, or monotonic. It is not clear how this definition would be formalized to multiport devices with memory – as a practical matter, circuit designers use this term informally, so it may not be necessary to formalize it.

Systems for which the small signal model is not passive are sometimes called locally active (e.g. transistors and tunnel diodes). Systems that can generate power about a time-variant unperturbed state are often called parametrically active (e.g. certain types of nonlinear capacitors) .

### Other definitions of passivity

In some very informal settings, passivity may refer to the simplicity of the device, although this definition is now almost universally considered incorrect. Here, devices like diodes would be considered active, and only very simple devices like capacitors, inductors, and resistors are considered passive. In some cases, the term "linear element" may be a more appropriate term than "passive device." In other cases, "solid state device" may be a more appropriate term than "active device."

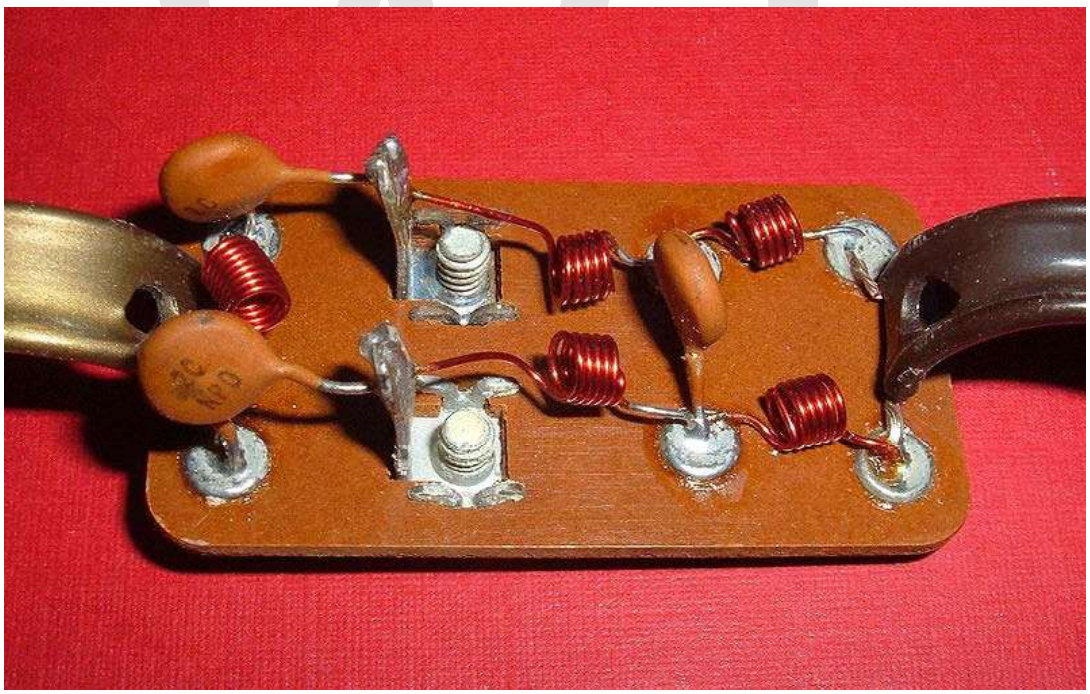
### Stability

Passivity, in most cases, can be used to demonstrate that passive circuits will be stable under specific criteria. Note that this only works if only one of the above definitions of passivity is used – if components from the two are mixed, the systems may be unstable under any criteria. In addition, passive circuits will not necessarily be stable under all stability criteria. For instance, a resonant series LC circuit will have unbounded voltage output for a bounded voltage input, but will be stable in the sense of Lyapunov, and given bounded energy input will have bounded energy output.

Passivity is frequently used in control systems to design stable control systems or to show stability in control systems. This is especially important in the design of large, complex control systems (e.g. stability of airplanes). Passivity is also used in some areas of circuit design, especially filter design.

## Passive filter

A passive filter is a kind of electronic filter that is made only from passive elements – in contrast to an active filter, it does not require an external power source (beyond the signal). Since most filters are linear, in most cases, passive filters are composed of just the four basic linear elements – resistors, capacitors, inductors, and transformers. More complex passive filters may involve nonlinear elements, or more complex linear elements, such as transmission lines.



Television signal splitter consisting of a passive high-pass filter (left) and a passive low-pass filter (right). The antenna is connected to the screw terminals to the left of center.

A passive filter has several advantages over an active filter:

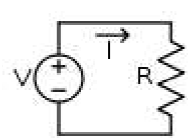
- Guaranteed stability
- Passive filters scale better to large signals (tens of amperes, hundreds of volts), where active devices are often impractical

- No power consumption.
- May be less expensive in discrete designs (unless large coils are required)
- For linear filters, may be, more linear than filters including active (and therefore non-linear) elements, depending on components required.

They are commonly used in speaker crossover design (due to the moderately large voltages and currents, and the lack of easy access to power), filters in power distribution networks (due to the large voltages and currents), power supply bypassing (due to low cost, and in some cases, power requirements), as well as a variety of discrete and home brew circuits (for low-cost and simplicity). Passive filters are uncommon in monolithic integrated circuit design, where active devices are inexpensive compared to resistors and capacitors, and inductors are prohibitively expensive. Passive filters are still found, however, in hybrid integrated circuits. Indeed, it may be the desire to incorporate a passive filter that leads the designer to use the hybrid format.

# Voltage source

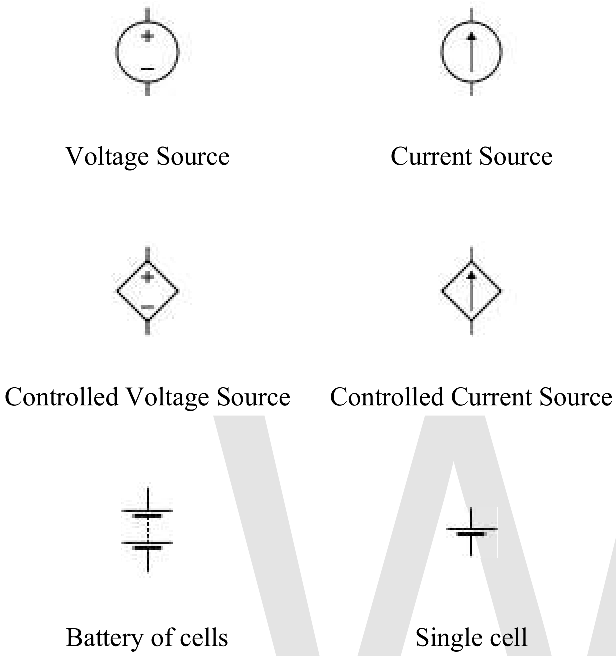
In electric circuit theory, an **ideal voltage source** is a circuit element where the voltage across it is independent of the current through it. A voltage source is the dual of a current source. In analysis, a voltage source supplies a constant DC or AC potential between its terminals for any current flow through it. Real-world sources of electrical energy, such as batteries, generators, or power systems, can be modeled for analysis purposes as a combination of an ideal voltage source and additional combinations of impedance elements.




A schematic diagram of an ideal voltage source, *V*, driving a resistor, *R*, and creating a current *I*

## Ideal voltage sources

An **ideal voltage source** is a mathematical abstraction that simplifies the analysis of electric circuits. If the voltage across an ideal voltage source can be specified independently of any other variable in a circuit, it is called an **independent** voltage source. Conversely, if the voltage across an ideal voltage source is determined by some other voltage or current in a circuit, it is called a **dependent** or **controlled voltage source**. A mathematical model of an amplifier will include dependent voltage sources whose magnitude is governed by some fixed relation to an input signal, for example. In the analysis of faults on electrical power systems, the whole network of interconnected sources and transmission lines can be usefully replaced by an ideal (AC) voltage source and a single equivalent impedance.




Symbols used for voltage sources

The internal resistance of an ideal voltage source is zero; it is able to supply or absorb any amount of current. The current through an ideal voltage source is completely determined by the external circuit. When connected to an open circuit, there is zero current and thus zero power. When connected to a load resistance, the current through the source approaches infinity as the load resistance approaches zero (a short circuit). Thus, an ideal voltage source can supply unlimited power.

No real voltage source is ideal; all have a non-zero effective internal resistance, and none can supply unlimited current. However, the internal resistance of a real voltage source is effectively modeled in linear circuit analysis by combining a non-zero resistance in series with an ideal voltage source (a Thévenin equivalent circuit).

### *Comparison between voltage and current sources*

Most sources of electrical energy (the mains, a battery) are modeled as voltage sources. An *ideal* voltage source provides no energy when it is loaded by an open circuit (i.e. an infinite impedance), but approaches infinite energy and current when the load resistance approaches zero (a short circuit). Such a theoretical device would have a zero ohm output impedance in series with the source. A real-world voltage source has a very low, but non-zero output impedance: often much less than 1 ohm.

Conversely, a current source provides a constant current, as long as the load connected to the source terminals has sufficiently low impedance. An ideal current source would provide no energy to a short circuit and approach infinite energy and voltage as the load resistance approaches infinity (an open circuit). An *ideal* current source has an infinite output impedance in parallel with the source. A *real-world* current source has a very high, but finite output impedance. In the case of transistor current sources, impedance of a few megohms (at low frequencies) is typical.

Since no ideal sources of either variety exist (all real-world examples have finite and non-zero source impedance), any current source can be considered as a voltage source with the *same* source impedance and vice versa. Voltage sources and current sources are sometimes said to be duals of each other and any non ideal source can be converted from one to the other by applying Norton's or Thevenin's theorems.

## Chapter-3

# Current Source

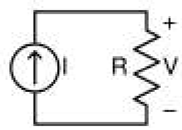



Figure 1: An ideal current source, *I*, driving a resistor, *R*, and creating a voltage *V*

A **current source** is an electrical or electronic device that delivers or absorbs electric current. A current source is the dual of a voltage source. The term constant-current **sink** is sometimes used for sources fed from a negative voltage supply. Figure 1 shows a schematic for an ideal current source driving a resistor load.



### *Ideal current sources*

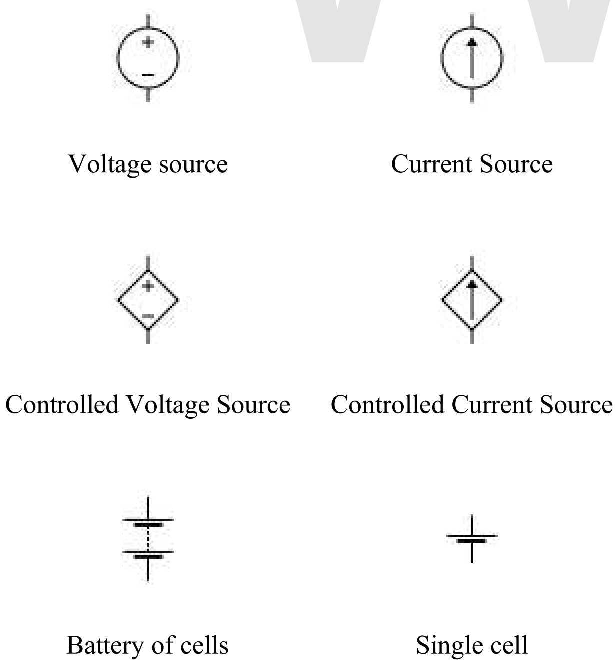



Figure 2: Source symbols

In circuit theory, an **ideal current source** is a circuit element where the current through it is independent of the voltage across it. It is a mathematical model, which real devices can only approach in performance. If the current through an ideal current source can be specified independently of any other variable in a circuit, it is called an *independent* current source. Conversely, if the current through an ideal current source is determined by some other voltage or current in a circuit, it is called a **dependent** or **controlled current source**. Symbols for these sources are shown in Figure 2.

An independent current source with zero current is identical to an ideal open circuit. For this reason, the internal resistance of an ideal current source is infinite. The voltage across an ideal current source is completely determined by the circuit it is connected to. When connected to a short circuit, there is zero voltage and thus zero power delivered. When connected to a load resistance, the voltage across the source approaches infinity as the load resistance approaches infinity (an open circuit). Thus, an ideal current source, if such a thing existed in reality, could supply unlimited power and so would represent an unlimited source of energy.

No real current source is ideal (no unlimited energy sources exist) and all have a finite internal resistance (none can supply unlimited voltage). However, the internal resistance of a physical current source is effectively modeled in circuit analysis by combining a non-zero resistance in parallel with an ideal current source (the Norton equivalent circuit). The connection of an ideal open circuit to an ideal non-zero current source does not represent any physically realizable system.

## *Physical current sources*

### Resistor current source

The simplest non-ideal current source consists of a voltage source in series with a resistor. The current available from such a source is given by the ratio of the voltage across the voltage source to the resistance of the resistor. This value of current will only be delivered to a load with zero voltage drop across its terminals (a short circuit, an uncharged capacitor, a charged inductor, a virtual ground circuit, etc.) The current delivered to a load with nonzero voltage (drop) across its terminals (a linear or nonlinear resistor with a finite resistance, a charged capacitor, an uncharged inductor, a voltage source, etc.) will always be different. It is given by the ratio of the voltage drop across the resistor (the difference between the exciting voltage and the voltage across the load) to its resistance. For a nearly ideal current source, the value of the resistor should be very large but this implies that, for a specified current, the voltage source must be very large (in the limit as the resistance and the voltage go to infinity, the current source will become ideal and the current will not depend at all on the voltage across the load). Thus, efficiency is low (due to power loss in the resistor) and it is usually impractical to construct a 'good' current source this way. Nonetheless, it is often the case that such a circuit will provide adequate performance when the specified current and load resistance are small. For example, a 5 V voltage source in series with a 4.7 kilohm resistor will provide an *approximately* constant current of 1 mA (±5%) to a load resistance in the range of 50 to 450 ohm.

A Van de Graaff generator is an example of such a high voltage current source. It behaves as an almost constant current source because of its very high output voltage coupled with its very high output resistance and so it supplies the same few microamperes at any output voltage up to hundreds of thousands of volts (or even tens of megavolts) for large laboratory versions.

## Active current sources without negative feedback

Active current sources have many important applications in electronic circuits. Current sources (current-stable resistors) are often used in place of ohmic resistors in analog integrated circuits to generate a current that depends slightly on the voltage across the load.

## Transistor current sources with constant input voltage

The common collector, common drain and common cathode configurations driven by a constant input voltage naturally behave as current sources (or sinks) because the output impedance of these devices is naturally high. The simple current mirror is an example of such a current source widely used in integrated circuits.

## FET current sources with zero input voltage

A JFET can be made to act as a current source by tying its gate to its source. The current then flowing is the $I_{DSS}$ of the FET. These can be purchased with this connection already made and in this case the devices are called current regulator diodes or constant current diodes or current limiting diodes (CLD). An enhancement mode N channel MOSFET can be used in the circuits listed below.

## Current sources with series negative feedback

## Simple transistor current source

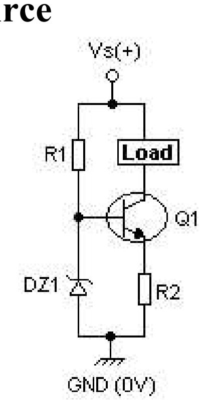



Figure 3: Typical constant current source (CCS)

Figure 3 shows a typical constant current source (CCS). DZ1 is a zener diode which, when reverse biased (as shown in the circuit) has a constant voltage drop across it irrespective of the current flowing through it. Thus, as long as the zener current ($I_Z$) is above a certain level (called holding current), the voltage across the zener diode ($V_Z$) will be constant. Resistor R1 supplies the zener current and the base current ($I_B$) of NPN transistor (Q1). The constant zener voltage is applied across the base of Q1 and emitter resistor R2. The operation of the circuit is as follows:

Voltage across R2 ($V_{R2}$) is given by $V_Z$ - $V_{BE}$, where $V_{BE}$ is the base-emitter drop of Q1. The emitter current of Q1 which is also the current through R2 is given by

$$I_{R2}(= I_E) = \frac{V_{R2}}{R2} = \frac{V_Z - V_{BE}}{R2}$$

Since $V_Z$ is constant and $V_{BE}$ is also (approximately) constant for a given temperature, it follows that $V_{R2}$ is constant and hence $I_E$ is also constant. Due to transistor action, emitter current $I_E$ is very nearly equal to the collector current $I_C$ of the transistor (which in turn, is the current through the load). Thus, the load current is constant (neglecting the output resistance of the transistor due to the Early effect) and the circuit operates as a constant current source. As long as the temperature remains constant (or doesn't vary much), the load current will be independent of the supply voltage, R1 and the transistor's gain. R2 allows the load current to be set at any desirable value and is calculated by

$$R2 = \frac{V_Z - V_{BE}}{I_{R2}}$$

or

$$R2 = \frac{V_Z - 0.65}{I_{R2}},$$

since $V_{BE}$ is typically 0.65 V for a silicon device.

($I_{R2}$ is also the emitter current and is assumed to be the same as the collector or required load current, provided $h_{FE}$ is sufficiently large). Resistance $R_1$ at resistor R1 is calculated as

$$R_1 = \frac{V_S - V_Z}{I_Z + K \cdot I_B}$$

where $K$ = 1.2 to 2 (so that $R_1$ is low enough to ensure adequate $I_B$),

$$I_B = \frac{I_C(= I_E = I_{R2})}{h_{FE(min)}}$$

and $h_{FE(min)}$ is the lowest acceptable current gain for the particular transistor type being used.

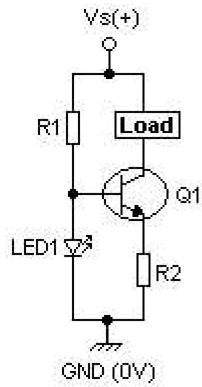



Figure 5: Typical constant current source (CCS) using LED instead of zener

The Zener diode can be replaced by any other diode, e.g. a light-emitting diode LED1 as shown in Figure 5. The LED voltage drop ($V_D$) is now used to derive the constant voltage and also has the additional advantage of tracking (compensating) $V_{BE}$ changes due to temperature. $R_2$ is calculated as

$$R_2 = \frac{V_D - V_{BE}}{I_{R2}}$$

and $R_1$ as

$$R_1 = \frac{V_S - V_D}{I_D + K \cdot I_B}$$, where $I_D$ is the LED current.

## Simple transistor current source with diode compensation

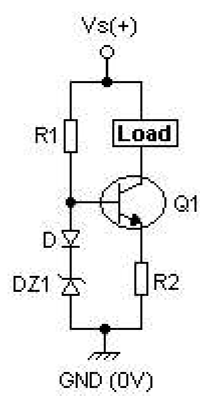



Figure 4: Typical constant current source (CCS) with diode compensation

Temperature changes will change the output current delivered by the circuit of Figure 3 because $V_{BE}$ is sensitive to temperature. Temperature dependence can be compensated using the circuit of Figure 4 that includes a standard diode D (of the same semiconductor material as the transistor) in series with the Zener diode as shown in the image on the left. The diode drop ($V_D$) tracks the $V_{BE}$ changes due to temperature and thus significantly counteracts temperature dependence of the CCS.

Resistance $R_2$ is now calculated as

$$R_2 = \frac{V_Z + V_D - V_{BE}}{I_{R2}}$$

Since $V_D = V_{BE} = 0.65$ V,

$$R_2 = \frac{V_Z}{I_{R2}}$$

(In practice $V_D$ is never exactly equal to $V_{BE}$ and hence it only suppresses the change in $V_{BE}$ rather than nulling it out.)

$R_1$ is calculated as

$$R_1 = \frac{V_S - V_Z - V_D}{I_Z + K \cdot I_B}$$ (the compensating diode's forward voltage drop $V_D$ appears in the equation and is typically 0.65 V for silicon devices.)

This method is most effective for Zener diodes rated at 5.6 V or more. For breakdown diodes of less than 5.6 V, the compensating diode is usually not required because the breakdown mechanism is not as temperature dependent as it is in breakdown diodes above this voltage.

## Current mirror with emitter degeneration

Series negative feedback is also used in the two-transistor current mirror with emitter degeneration. Negative feedback is a basic feature in some current mirrors using multiple transistors, such as the Widlar current source and the Wilson current source.

**Op-amp current sources...**

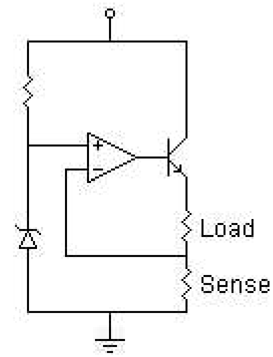



Figure 6: Typical op-amp current source. The transistor is not needed if the required current doesn't exceed the sourcing ability of the op-amp. The current will be the zener voltage divided by the sense resistor.

## ...with series negative feedback...

Another common method is to use a negative feedback to set the current and remove the dependence on the Vbe of the transistor. Figure 6 shows a very common approach using an op amp with the non-inverting input connected to a voltage source (such as the Zener in an above example) and the inverting input connected to the same node as the resistor and emitter of the transistor. The circuit is actually a non-inverting amplifier driven by a constant input voltage. It keeps up this constant voltage across the constant sense resistor; as a result, the current flowing through the load is constant as well.

## ...with parallel negative feedback

In the case of op-amp circuits (e.g., an op-amp voltage-to-current converter) sometimes it is desired to inject a precisely known current through a resistor into the inverting input (as an offset of signal input for instance). The combination of the input voltage source and the resistor will approximate an ideal current source with value *V*/*R*. The op-amp inverting input will be at virtual ground.

### Current source made by a voltage regulator

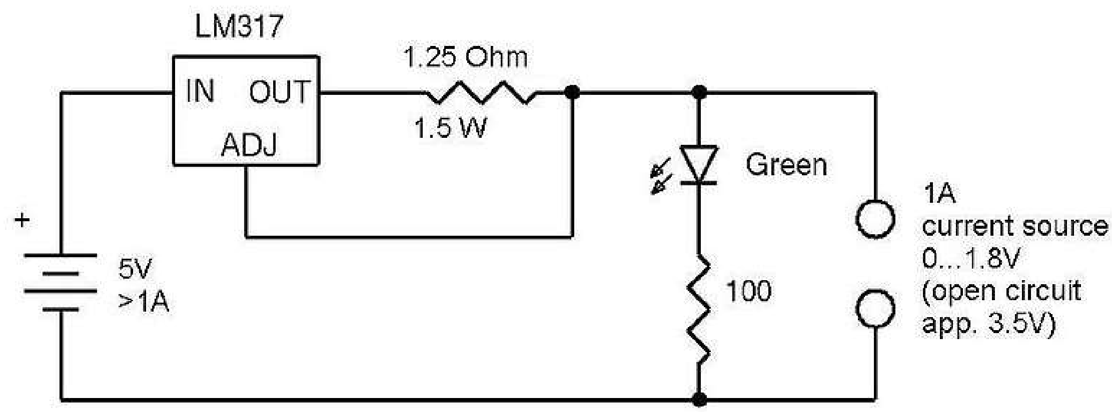



Figure 7: Constant current source using the LM317 voltage regulator

The circuit of Figure 7 using the LM317 voltage regulator is used to present a constant current source. The voltage regulator keeps up a constant voltage drop (1.25 V) across a constant resistor (1.25 Ω); so, a constant current (1 A) flows through the resistor and the load. The LED is on when the voltage across the load exceeds 1.8 V (the indicator circuit introduces some error).

## *Current and voltage source comparison*

Most sources of electrical energy (mains electricity, a battery, ...) are best modeled as voltage sources. Such sources provide constant voltage, which means that as long as the amount of current drawn from the source is within the source's capabilities, its output voltage stays constant. An ideal voltage source provides no energy when it is loaded by an open circuit (i.e. an infinite impedance), but approaches infinite power and current when the load resistance approaches zero (a short circuit). Such a theoretical device would have a zero ohm output impedance in series with the source. A real-world voltage source has a very low, but non-zero output impedance: often much less than 1 ohm.

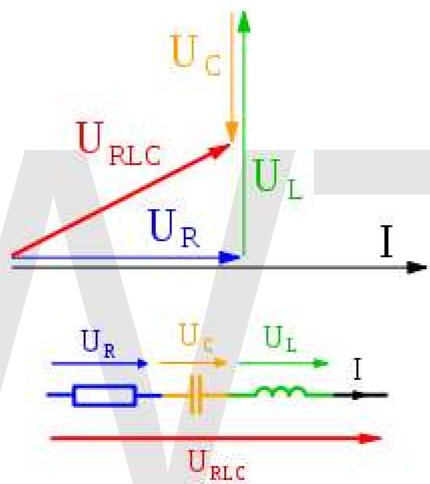
Conversely, a current source provides a constant current, as long as the load connected to the source terminals has sufficiently low impedance. An ideal current source would provide no energy to a short circuit and approach infinite energy and voltage as the load resistance approaches infinity (an open circuit). An *ideal* current source has an infinite output impedance in parallel with the source. A *real-world* current source has a very high, but finite output impedance. In the case of transistor current sources, impedances of a few megohms (at DC) are typical.

An *ideal* current source cannot be connected to an *ideal* open circuit because this would create the paradox of running a constant, non-zero current (from the current source) through an element with a defined zero current (the open circuit). Nor can an *ideal* voltage source be connected to an *ideal* short circuit (R=0), since this would result a similar paradox of finite non zero voltage across an element with defined zero voltage (the short circuit).

Because no ideal sources of either variety exist (all real-world examples have finite and non-zero source impedance), any current source can be considered as a voltage source with the *same* source impedance and vice versa. These concepts are dealt with by Norton's and Thévenin's theorems.

**Chapter-4**

# Phasor




An example of series RLC circuit and respective **phasor diagram**

In physics and engineering, a **phase vector**, or **phasor**, is a representation of a sine wave whose amplitude (**A**), phase (**θ**), and frequency (**ω**) are time-invariant. It is a subset of a more general concept called analytic representation. Phasors reduce the dependencies on these parameters to three independent factors, thereby simplifying certain kinds of calculations. In particular the frequency factor, which also includes the time-dependence of the sine wave, is often common to all the components of a linear combination of sine waves. Using phasors, it can be factored out, leaving just the static amplitude and phase information to be combined algebraically (rather than trigonometrically). Similarly, linear differential equations can be reduced to algebraic ones. The term *phasor* therefore often refers to just those two factors. In older texts, a **phasor** is also referred to as a **sinor**.

## *Definition*

Euler's formula indicates that sine waves can be represented mathematically as the sum of two complex-valued functions**:**

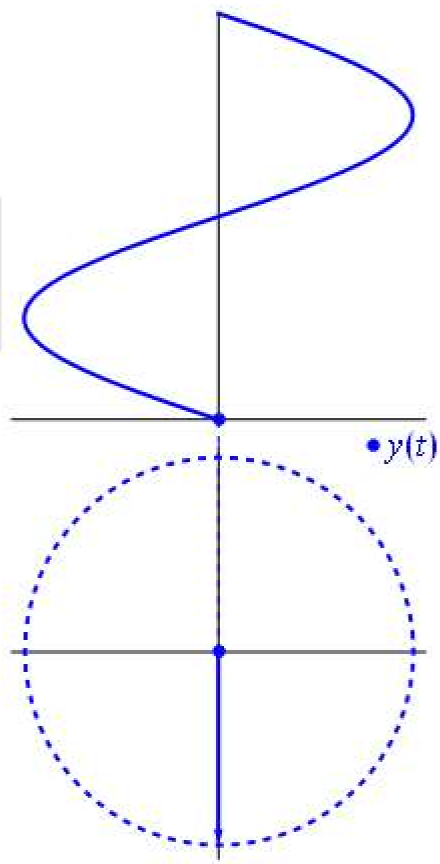
$$A \cdot \cos(\omega t + \theta) = A/2 \cdot e^{i(\omega t+\theta)} + A/2 \cdot e^{-i(\omega t+\theta)},$$

or as the real part of one of the functions:

$$\begin{aligned} A \cdot \cos(\omega t + \theta) &= \operatorname{Re}\left\{A \cdot e^{i(\omega t+\theta)}\right\} \\ &= \operatorname{Re}\left\{Ae^{i\theta} \cdot e^{i\omega t}\right\}. \end{aligned}$$

As indicated above, *phasor* can refer to either $Ae^{i\theta}e^{i\omega t}$ or just the complex constant, $Ae^{i\theta}$. In the latter case, it is understood to be a shorthand notation, encoding the amplitude and phase of an underlying sinusoid.

An even more compact shorthand is angle notation: $A\angle\theta$.




A phasor can be seen as a rotating vector. The graphical representation (on paper) is at $t = 0$.

The sine wave can be understood as the projection onto the real axis of a rotating vector on the complex plane. The modulus of this vector is the amplitude of the oscillations, while its argument is the total phase $\omega t + \theta$. The phase constant $\theta$ represents the angle that the complex vector forms with the real axis at $t = 0$.

## *Phasor arithmetic*

### Multiplication by a constant (scalar)

Multiplication of the phasor $Ae^{i\theta}e^{i\omega t}$ by a complex constant, $Be^{i\phi}$ produces another phasor. That means its only effect is to change the amplitude and phase of the underlying sinusoid:

$$\begin{aligned}\operatorname{Re}\{(Ae^{i\theta}\cdot Be^{i\phi})\cdot e^{i\omega t}\} &= \operatorname{Re}\{(ABe^{i(\theta+\phi)})\cdot e^{i\omega t}\}\\ &= AB\cos(\omega t+(\theta+\phi))\end{aligned}$$

In electronics, $Be^{i\phi}$ would represent an impedance, which is independent of time. In particular it is *not* the shorthand notation for another phasor. Multiplying a phasor current by an impedance produces a phasor voltage. But the product of two phasors (or squaring a phasor) would represent the product of two sine waves, which is a non-linear operation that produces new frequency components. Phasor notation can only represent systems with one frequency, such as a linear system stimulated by a sinusoid.

### Differentiation and integration

The time derivative or integral of a phasor produces another phasor. For example:

$$\begin{aligned}\operatorname{Re}\left\{\frac{d}{dt}(Ae^{i\theta}\cdot e^{i\omega t})\right\} &= \operatorname{Re}\{Ae^{i\theta}\cdot i\omega e^{i\omega t}\}\\ &= \operatorname{Re}\{Ae^{i\theta}\cdot e^{i\pi/2}\omega e^{i\omega t}\}\\ &= \operatorname{Re}\{\omega Ae^{i(\theta+\pi/2)}\cdot e^{i\omega t}\}\\ &= \omega A\cdot\cos(\omega t+\theta+\pi/2)\end{aligned}$$

Therefore, in phasor representation, the time derivative of a sinusoid becomes just multiplication by the constant, $i\omega = (e^{i\pi/2}\cdot\omega)$. Similarly, integrating a phasor corresponds to multiplication by $\frac{1}{i\omega} = \frac{e^{-i\pi/2}}{\omega}.$ The time-dependent factor, $e^{i\omega t}$, is unaffected. When we solve a linear differential equation with phasor arithmetic, we are merely factoring $e^{i\omega t}$ out of all terms of the equation, and reinserting it into the answer. For example, consider the following differential equation for the voltage across the capacitor in an RC circuit:

$$\frac{d\ v_C(t)}{dt} + \frac{1}{RC}v_C(t) = \frac{1}{RC}v_S(t)$$

When the voltage source in this circuit is sinusoidal:

$$v_S(t) = V_P \cdot \cos(\omega t + \theta),$$

we may substitute:

$$v_S(t) = \mathrm{Re}\{V_s \cdot e^{i\omega t}\}$$
$$v_C(t) = \mathrm{Re}\{V_c \cdot e^{i\omega t}\},$$

where phasor $V_s = V_P e^{i\theta}$, and phasor $V_c$is the unknown quantity to be determined.

In the phasor shorthand notation, the differential equation reduces to:

$$i\omega V_c + \frac{1}{RC} V_c = \frac{1}{RC} V_s$$

Solving for the phasor capacitor voltage gives:

$$V_c = \frac{1}{1 + i\omega RC} \cdot (V_s) = \frac{1 - i\omega RC}{1 + (\omega RC)^2} \cdot (V_P e^{i\theta})$$

As we have seen, the factor multiplying $V_s$ represents differences of the amplitude and phase of $v_C(t)$ relative to $V_P$ and $\theta$.

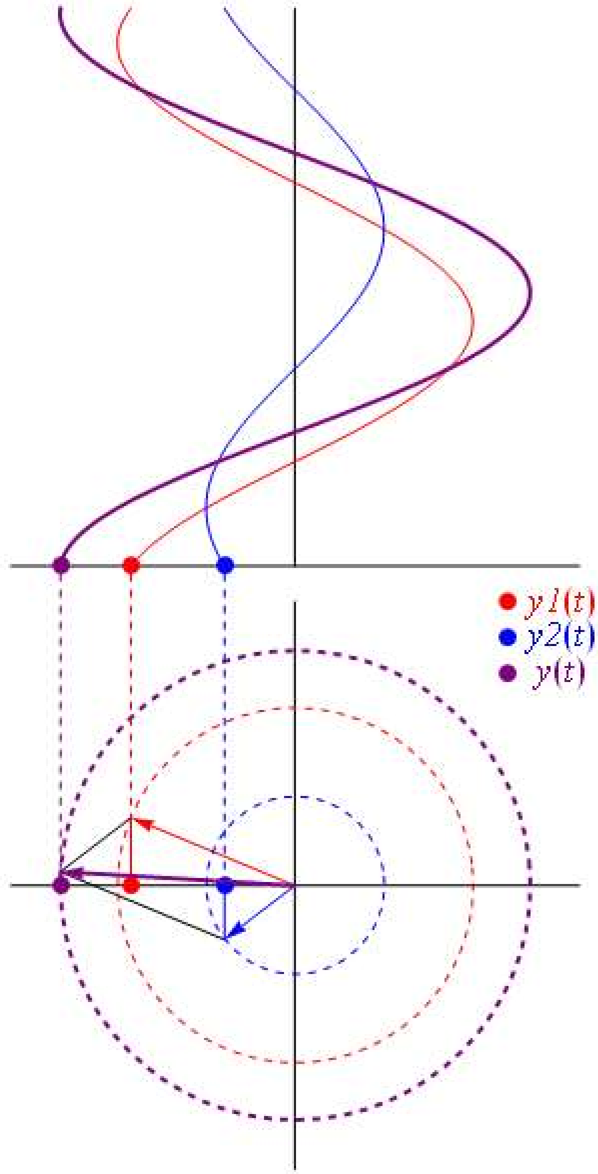
In polar coordinate form, it is:

$$\frac{1}{\sqrt{1 + (\omega RC)^2}} \cdot e^{-i\phi(\omega)}, \text{ where } \phi(\omega) = \arctan(\omega RC).$$

Therefore:

$$v_C(t) = \frac{1}{\sqrt{1 + (\omega RC)^2}} \cdot V_P \cos(\omega t + \theta - \phi(\omega))$$

## Addition




The sum of phasors as addition of rotating vectors

The sum of multiple phasors produces another phasor. That is because the sum of sine waves with the same frequency is also a sine wave with that frequency:

$$\begin{aligned} A_1\cos(\omega t+\theta_1)+A_2\cos(\omega t+\theta_2) &= \operatorname{Re}\{A_1e^{i\theta_1}e^{i\omega t}\}+\operatorname{Re}\{A_2e^{i\theta_2}e^{i\omega t}\} \\ &= \operatorname{Re}\{A_1e^{i\theta_1}e^{i\omega t}+A_2e^{i\theta_2}e^{i\omega t}\} \\ &= \operatorname{Re}\{(A_1e^{i\theta_1}+A_2e^{i\theta_2})e^{i\omega t}\} \\ &= \operatorname{Re}\{(A_3e^{i\theta_3})e^{i\omega t}\} \\ &= A_3\cos(\omega t+\theta_3), \end{aligned}$$

where:

$$A_3^2 = (A_1 \cos\theta_1 + A_2 \cos\theta_2)^2 + (A_1 \sin\theta_1 + A_2 \sin\theta_2)^2,$$
$$\theta_3 = \arctan\left(\frac{A_1 \sin\theta_1 + A_2 \sin\theta_2}{A_1 \cos\theta_1 + A_2 \cos\theta_2}\right)$$

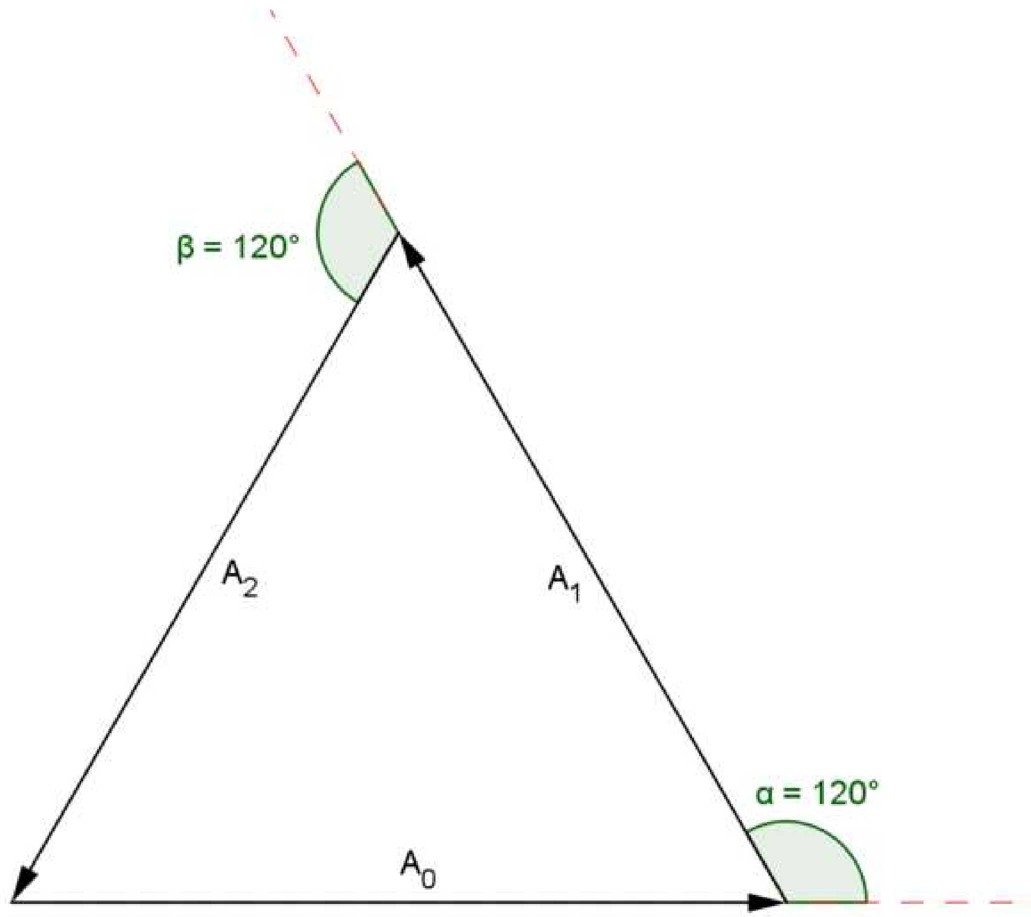
or, via the law of cosines on the complex plane (or the trigonometric identity for angle differences):

$$A_3^2 = A_1^2 + A_2^2 - 2A_1A_2\cos(180^\circ - \Delta\theta), = A_1^2 + A_2^2 + 2A_1A_2\cos(\Delta\theta),$$

where $\Delta\theta = \theta_1 - \theta_2$. A key point is that $A_3$ and $\theta_3$ do not depend on ω or t, which is what makes phasor notation possible. The time and frequency dependence can be suppressed and re-inserted into the outcome as long as the only operations used in between are ones that produce another phasor. In angle notation, the operation shown above is written:

$$A_1\angle\theta_1 + A_2\angle\theta_2 = A_3\angle\theta_3.$$

Another way to view addition is that two **vectors** with coordinates $[A_1 \cos(\omega t+\theta_1), A_1 \sin(\omega t+\theta_1)]$ and $[A_2 \cos(\omega t+\theta_2), A_2 \sin(\omega t+\theta_2)]$ are added vectorially to produce a resultant vector with coordinates $[A_3 \cos(\omega t+\theta_3), A_3 \sin(\omega t+\theta_3)]$.




Phasor diagram of three waves in perfect destructive interference

In physics, this sort of addition occurs when sine waves "interfere" with each other, constructively or destructively. The static vector concept provides useful insight into questions like this: "What phase difference would be required between three identical waves for perfect cancellation?" In this case, simply imagine taking three vectors of equal length and placing them head to tail such that the last head matches up with the first tail. Clearly, the shape which satisfies these conditions is an equilateral triangle, so the angle between each phasor to the next is 120° (2π/3 radians), or one third of a wavelength $^{\lambda}/_{3}$. So the phase difference between each wave must also be 120°, as is the case in three-phase power

In other words, what this shows is:

$$\cos(\omega t) + \cos(\omega t + 2\pi/3) + \cos(\omega t + 4\pi/3) = 0.$$

In the example of three waves, the phase difference between the first and the last wave was 240 degrees, while for two waves destructive interference happens at 180 degrees. In the limit of many waves, the phasors must form a circle for destructive interference, so that the first phasor is nearly parallel with the last. This means that for many sources, destructive interference happens when the first and last wave differ by 360 degrees, a full wavelength $\lambda$. This is why in single slit diffraction, the minima occurs when light from the far edge travels a full wavelength further than the light from the near edge.

### Phasor diagrams

Electrical engineers, electronics engineers, electronic engineering technicians and aircraft engineers all use phasor diagrams to visualize complex constants and variables (phasors). Like vectors, arrows drawn on graph paper or computer displays represent phasors. Cartesian and polar representations each have advantages.

## *Circuit laws*

With phasors, the techniques for solving DC circuits can be applied to solve AC circuits. A list of the basic laws is given below.

- **Ohm's law for resistors:** a resistor has no time delays and therefore doesn't change the phase of a signal therefore $V=IR$ remains valid.
- **Ohm's law for resistors, inductors, and capacitors:** $V = IZ$ where $Z$ is the complex impedance.
- In an AC circuit we have real power ($P$) which is a representation of the average power into the circuit and reactive power ($Q$) which indicates power flowing back and forward. We can also define the complex power $S = P + jQ$ and the apparent power which is the magnitude of $S$. The power law for an AC circuit expressed in phasors is then $S = VI^*$ (where $I^*$ is the complex conjugate of $I$).
- Kirchhoff's circuit laws work with phasors in complex form

Given this we can apply the techniques of analysis of resistive circuits with phasors to analyze single frequency AC circuits containing resistors, capacitors, and inductors. Multiple frequency linear AC circuits and AC circuits with different waveforms can be analyzed to find voltages and currents by transforming all waveforms to sine wave components with magnitude and phase then analyzing each frequency separately, as allowed by the superposition theorem.

## *Power engineering*

In analysis of three phase AC power systems, usually a set of phasors is defined as the three complex cube roots of unity, graphically represented as unit magnitudes at angles of 0, 120 and 240 degrees. By treating polyphase AC circuit quantities as phasors, balanced circuits can be simplified and unbalanced circuits can be treated as an algebraic combination of symmetrical circuits. This approach greatly simplifies the work required in electrical calculations of voltage drop, power flow, and short-circuit currents. In the context of power systems analysis, the phase angle is often given in degrees, and the magnitude in rms value rather than the peak amplitude of the sinusoid.

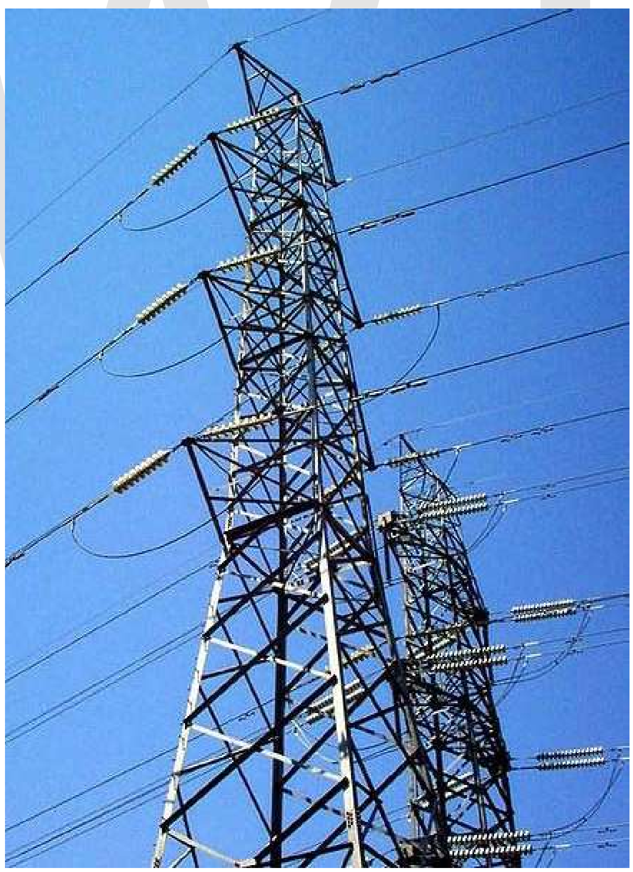
The technique of synchrophasors uses digital instruments to measure the phasors representing transmission system voltages at widespread points in a transmission network. Small changes in the phasors are sensitive indicators of power flow and system stability.

Chapter-5

# Electric Power and Thévenin's Theorem

## Electric power

**Electric power** is the rate at which electric energy is transferred by an electric circuit. The SI unit of power is the watt.



Electric energy is transmitted with **overhead lines** on pylons like these in Brisbane, Australia.

When electric current flows in a circuit, it can transfer energy to do mechanical or thermodynamic work. Devices convert electrical energy into many useful forms, such as heat (electric heaters), light (light bulbs), motion (electric motors), sound (loudspeaker), information technological processes (computers), or even chemical changes. Electricity can be produced mechanically by generation, or chemically, or by direct conversion from light in photovoltaic cells, also it can be stored chemically in batteries.

## *Mathematics of electric power*

### Circuits

Electric power, like mechanical power, is represented by the letter *P* in electrical equations. The term *wattage* is used colloquially to mean "electric power in watts."

### Direct current

In direct current resistive circuits, electrical power is calculated using Joule's law:

$$P = VI$$

where *P* is the electric power, *V* the potential difference, and *I* the electric current.

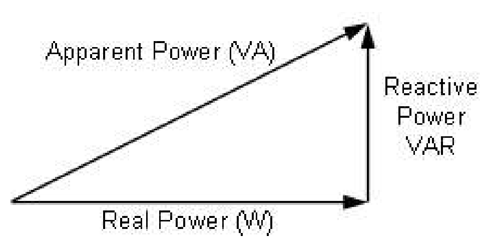
In the case of resistive (Ohmic, or linear) loads, Joule's law can be combined with Ohm's law ($I = V/R$) to produce alternative expressions for the dissipated power:

$$P = I^2 R = \frac{V^2}{R},$$

where *R* is the electrical resistance.

### Alternating current

In alternating current circuits, energy storage elements such as inductance and capacitance may result in periodic reversals of the direction of energy flow. The portion of power flow that, averaged over a complete cycle of the AC waveform, results in net transfer of energy in one direction is known as real power (also referred to as active power). That portion of power flow due to stored energy, that returns to the source in each cycle, is known as reactive power.




Power triangle The components of AC power

The relationship between real power, reactive power and apparent power can be expressed by representing the quantities as vectors. Real power is represented as a horizontal vector and reactive power is represented as a vertical vector. The apparent power vector is the hypotenuse of a right triangle formed by connecting the real and reactive power vectors. This representation is often called the *power triangle*. Using the Pythagorean Theorem, the relationship among real, reactive and apparent power is:

$$(\text{apparent power})^2 = (\text{real power})^2 + (\text{reactive power})^2$$

Real and reactive powers can also be calculated directly from the apparent power, when the current and voltage are both sinusoids with a known phase angle between them:

$$(\text{real power}) = (\text{apparent power})\cos(\theta)$$
$$(\text{reactive power}) = (\text{apparent power})\sin(\theta)$$

The ratio of real power to apparent power is called power factor and is a number always between 0 and 1. Where the currents and voltages have non-sinusoidal forms, power factor is generalized to include the effects of distortion.

### In space

Electrical power flows wherever electric and magnetic fields exist together and fluctuate in the same place. The simplest example of this is in electrical circuits, as the preceding section showed. In the general case, however, the simple equation $P = IV$ must be replaced by a more complex calculation, the integral of the cross-product of the electrical and magnetic field vectors over a specified area, thus:

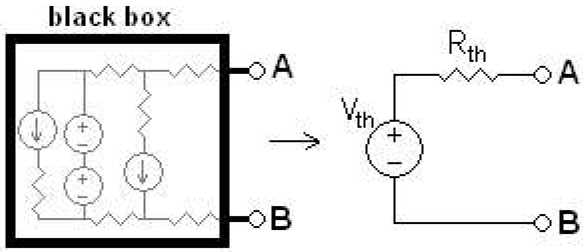
$$P = \int_S (\mathbf{E} \times \mathbf{H}) \cdot \mathbf{dA}.$$

The result is a scalar since it is the *surface integral* of the *Poynting vector*.

# Thévenin's theorem

In circuit theory, **Thévenin's theorem** for linear electrical networks states that any combination of voltage sources, current sources, and resistors with two terminals is electrically equivalent to a single voltage source *V* and a single series resistor *R*. For single frequency AC systems the theorem can also be applied to general impedances, not just resistors. The theorem was first discovered by German scientist Hermann von Helmholtz in 1853, but was then rediscovered in 1883 by French telegraph engineer Léon Charles Thévenin (1857–1926).

This theorem states that a circuit of voltage sources and resistors can be converted into a **Thévenin equivalent**, which is a simplification technique used in circuit analysis. The Thévenin equivalent can be used as a good model for a power supply or battery (with the resistor representing the internal impedance and the source representing the electromotive force). The circuit consists of an ideal voltage source in series with an ideal resistor.




Any black box containing only voltage sources, current sources, and other resistors can be converted to a Thévenin equivalent circuit, comprising exactly one voltage source and one resistor.

### *Calculating the Thévenin equivalent*

To calculate the equivalent circuit, the resistance and voltage are needed, so two equations are required. These two equations are usually obtained by using the following steps, but any conditions placed on the terminals of the circuit should also work:

1. Calculate the output voltage, $V_{AB}$, when in open circuit condition (no load resistor—meaning infinite resistance). This is $V_{Th}$.
2. Calculate the output current, $I_{AB}$, when the output terminals are short circuited (load resistance is 0). $R_{Th}$ equals $V_{Th}$ divided by this $I_{AB}$.

The equivalent circuit is a voltage source with voltage $V_{Th}$ in series with a resistance $R_{Th}$.

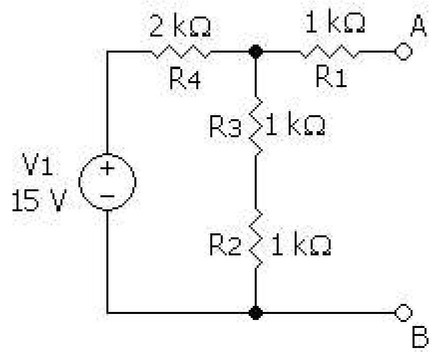
Step 2 could also be thought of as:

2a. Replace voltage sources with short circuits, and current sources with open circuits.
2b. Calculate the resistance between terminals A and B. This is $R_{Th}$.

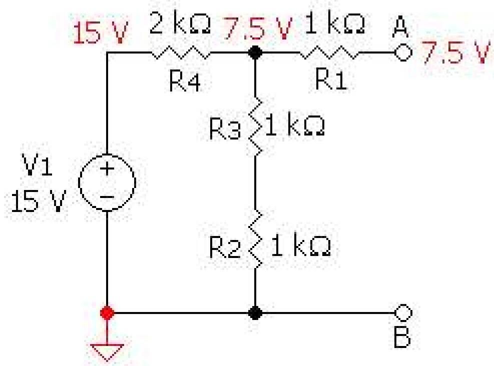
The Thévenin-equivalent voltage is the voltage at the output terminals of the original circuit. When calculating a Thévenin-equivalent voltage, the voltage divider principle is often useful, by declaring one terminal to be $V_{out}$ and the other terminal to be at the ground point.

The Thévenin-equivalent resistance is the resistance measured across points A and B "looking back" into the circuit. It is important to first replace all voltage- and current-sources with their internal resistances. For an ideal voltage source, this means replace the voltage source with a short circuit. For an ideal current source, this means replace the current source with an open circuit. Resistance can then be calculated across the terminals using the formulae for series and parallel circuits. This method is valid only for circuits with independent sources. If there are dependent sources in the circuit, another method must be used such as connecting a test source across A and B and calculating the voltage across or current through the test source.

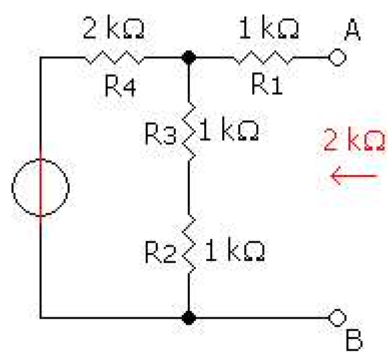
## Example



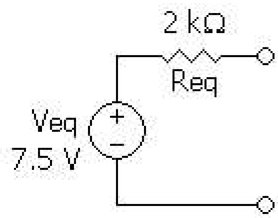

Step 0: The original circuit




Step 1: Calculating the equivalent output voltage




Step 2: Calculating the equivalent resistance




Step 3: The equivalent circuit

In the example, calculating the equivalent voltage:

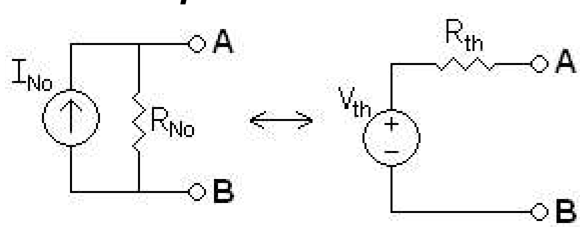
$$V_{\mathrm{Th}} = \frac{R_2 + R_3}{(R_2 + R_3) + R_4} \cdot V_1$$
$$= \frac{1\,\mathrm{k}\Omega + 1\,\mathrm{k}\Omega}{(1\,\mathrm{k}\Omega + 1\,\mathrm{k}\Omega) + 2\,\mathrm{k}\Omega} \cdot 15\,\mathrm{V}$$
$$= \frac{1}{2} \cdot 15\,\mathrm{V} = 7.5\,\mathrm{V}$$

(notice that $R_1$ is not taken into consideration, as above calculations are done in an open circuit condition between A and B, therefore no current flows through this part, which means there is no current through $R_1$ and therefore no voltage drop along this part)

Calculating equivalent resistance:

$$R_{\mathrm{Th}} = R_1 + [(R_2 + R_3) \| R_4)]$$
$$= 1\,\mathrm{k}\Omega + [(1\,\mathrm{k}\Omega + 1\,\mathrm{k}\Omega) \| 2\,\mathrm{k}\Omega)]$$
$$= 1\,\mathrm{k}\Omega + \left(\frac{1}{(1\,\mathrm{k}\Omega + 1\,\mathrm{k}\Omega)} + \frac{1}{(2\,\mathrm{k}\Omega)}\right)^{-1} = 2\,\mathrm{k}\Omega$$

## *Conversion to a Norton equivalent*




A Norton equivalent circuit is related to the Thévenin equivalent by the following:

$$R_{\mathrm{Th}} = R_{\mathrm{No}}$$
$$V_{\mathrm{Th}} = I_{\mathrm{No}} R_{\mathrm{No}}$$
$$I_{\mathrm{No}} = V_{\mathrm{Th}} / R_{\mathrm{Th}}$$

### *Practical limitations*

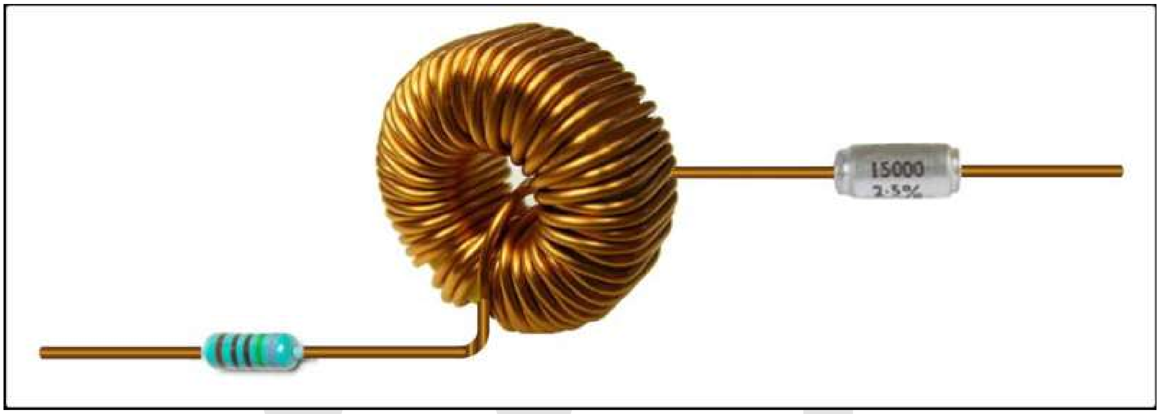
- Many, if not most circuits are only linear over a certain range of values, thus the Thévenin equivalent is valid only within this linear range and may not be valid outside the range.

- The Thévenin equivalent has an equivalent I-V characteristic only from the point of view of the load.

- The power dissipation of the Thévenin equivalent is not necessarily identical to the power dissipation of the real system. However, the power dissipated by an external resistor between the two output terminals is the same however the internal circuit is represented.

## Chapter-6

# RLC Circuit




A series RLC circuit: a resistor, inductor, and a capacitor

An **RLC circuit** (or **LCR circuit**) is an electrical circuit consisting of a resistor, an inductor, and a capacitor, connected in series or in parallel. The RLC part of the name is due to those letters being the usual electrical symbols for resistance, inductance and capacitance respectively. The circuit forms a harmonic oscillator for current and will resonate in just the same way as an LC circuit will. The difference that the presence of the resistor makes is that any oscillation induced in the circuit will die away over time if it not kept going by a source. This effect of the resistor is called damping. Some resistance is unavoidable in real circuits, even if a resistor is not specifically included as a component. A pure LC circuit is an ideal which really only exists in theory.

There are many applications for this circuit. They are used in many different types of oscillator circuit. Another important application is for tuning, such as in radio receivers or television sets, where they are used to select a narrow range of frequencies from the ambient radio waves. In this role the circuit is often referred to as a tuned circuit. An RLC circuit can be used as a band-pass filter or a band-stop filter. The tuning application, for instance, is an example of band-pass filtering. The RLC filter is described as a *second-order* circuit, meaning that any voltage or current in the circuit can be described by a second-order differential equation in circuit analysis.

The three circuit elements can be combined in a number of different topologies. All three elements in series or all three elements in parallel are the simplest in concept and the most straightforward to analyse. There are, however, other arrangements, some with practical importance in real circuits. One issue often encountered is the need to take into account inductor resistance. Inductors are typically constructed from coils of wire, the resistance of which is not usually desirable, but it often has a significant effect on the circuit.

## *Basic concepts*

### Resonance

An important property of this circuit is its ability to resonate at a specific frequency, the resonance frequency, $f_0$. Frequencies are measured in units of hertz. Here, however, angular frequency, $\omega_0$, is used which is more mathematically convenient. This is measured in radians per second. They are related to each other by a simple proportion,

$$\omega_0 = 2\pi f_0$$

Resonance occurs because energy is stored in two different ways: in an electric field as the capacitor is charged and in a magnetic field as current flows through the inductor. Energy can be transferred from one to the other within the circuit and this can be oscillatory. A mechanical analogy is a weight suspended on a spring which will oscillate up and down when released. This is no passing metaphor, a weight on a spring is described by exactly the same second order differential equation as an RLC circuit and for all the properties of the one system there will be found an analogous property of the other. The mechanical property answering to the resistor in the circuit is friction in the spring/weight system. Friction will slowly bring any oscillation to a halt if there is no external force driving it. Likewise, the resistance in an RLC circuit will "damp" the oscillation, diminishing it with time if there is no driving AC power source in the circuit.

The resonance frequency is defined as the frequency at which the impedance of the circuit is at a minimum. Equivalently, it can be defined as the frequency at which the impedance is purely real (that is, purely resistive). This occurs because the impedance (reactance) of the inductor and capacitor at resonance are equal but of opposite sign and cancel out. Circuits where L and C are in parallel rather than series actually have a maximum impedance rather than a minimum impedance. For this reason they are often described as antiresonators, it is still usual, however, to name the frequency at which this occurs as the resonance frequency.

## Natural frequency

The resonance frequency is defined in terms of the impedance presented to a driving source. It is still possible for the circuit to carry on oscillating (for a time) after the driving source has been removed or it is subjected to a step in voltage (including a step down to zero). This is similar to the way that a tuning fork will carry on ringing after it has been struck, and the effect is often called ringing. This effect is the undriven natural resonance frequency of the circuit and in general is not exactly the same as the driven resonance frequency, although the two will usually be quite close to each other. Various terms are used by different authors to distinguish the two, but resonance frequency unqualified usually means the driven resonance frequency. The driven frequency may be called the undamped resonance frequency or undamped natural frequency and the undriven frequency may be called the damped resonance frequency or the damped natural frequency. The reason for this terminology is that the driven resonance frequency in a series or parallel resonant circuit has the value

$$\omega_0 = \frac{1}{\sqrt{LC}}$$

This is exactly the same as the resonance frequency of an LC circuit, that is, one with no resistor present, that is, it is the same as a circuit in which there is no damping, hence undamped resonance frequency. The undriven resonance frequency, on the other hand, depends on the value of the resistor and hence is described as the damped resonance frequency. A highly damped circuit will fail to resonate at all when undriven. A circuit with a value of resistor that causes it to be just on the edge of ringing is called critically damped. Either side of critically damped are described as underdamped (ringing happens) and overdamped (ringing is suppressed).

Circuits with topologies more complex than straightforward series or parallel (some examples described later) have a driven resonance frequency that deviates from $\omega_0 = \frac{1}{\sqrt{LC}}$and for those the undamped resonance frequency, damped resonance frequency and driven resonance frequency can all be different.

## Damping

Damping is caused by the resistance in the circuit. It determines whether or not the circuit will resonate naturally (that is, without a driving source). Circuits which will resonate in this way are described as underdamped and those that will not are overdamped. Damping attenuation (symbol $\alpha$) is measured in nepers per second. However, the unitless damping factor (symbol $\zeta$) is often a more useful measure, which is related to $\alpha$ by

$$\zeta = \frac{\alpha}{\omega_0}$$

The special case of $\zeta = 1$ is called critical damping and represents the case of a circuit that is just on the border of oscillation. It is the minimum damping that can be applied without causing oscillation.

## Bandwidth

The resonance effect can be used for filtering, the rapid change in impedance near resonance can be used to pass or block signals close to the resonance frequency. Both band-pass and band-stop filters can be constructed and some filter circuits are shown later. A key parameter in filter design is bandwidth. The bandwidth is measured between the 3dB-points, that is, the frequencies at which the power passed through the circuit has fallen to half the value passed at resonance. There are two of these half-power frequencies, one above, and one below the resonance frequency

$$\Delta\omega = \omega_2 - \omega_1$$

where $\Delta\omega$is the bandwidth, $\omega_1$is the lower half-power frequency and $\omega_2$is the upper half-power frequency. The bandwidth is related to attenuation by,

$$\Delta\omega = 2\alpha$$

when the units are radians per second and nepers per second respectively. Other units may require a conversion factor. A more general measure of bandwidth is the fractional bandwidth, which expresses the bandwidth as a fraction of the resonance frequency and is given by

$$F_\mathrm{b} = \frac{\Delta\omega}{\omega_0}$$

The fractional bandwidth is also often stated as a percentage. The damping of filter circuits is adjusted to result in the required bandwidth. A narrow band filter, such as a notch filter, requires low damping. A wide band filter requires high damping.

## Q factor

The Q factor is a widespread measure used to characterise resonators. It is defined as the peak energy stored in the circuit divided by the average energy dissipated in it per cycle at resonance. Low *Q* circuits are therefore damped and lossy and high *Q* circuits are underdamped. *Q* is related to bandwidth; low *Q* circuits are wide band and high *Q* circuits are narrow band. In fact, it happens that *Q* is the inverse of fractional bandwidth

$$Q = \frac{1}{F_\mathrm{b}} = \frac{\omega_0}{\Delta\omega}$$

Q factor is directly proportional to selectivity, as Q factor depends inversely on bandwidth.

### Scaled parameters

The parameters $\zeta$, $F_b$, and $Q$ are all scaled to $\omega_0$. This means that circuits which have similar parameters share similar characteristics regardless of whether or not they are operating in the same frequency band.

## *Series RLC circuit*

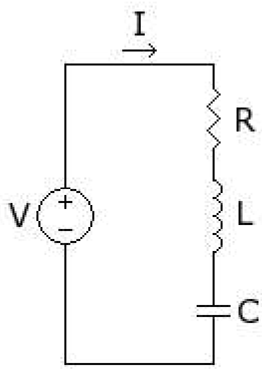



**Figure 1.** RLC series circuit

**V** - the voltage of the power source
**I** - the current in the circuit
**R** - the resistance of the resistor
**L** - the inductance of the inductor
**C** - the capacitance of the capacitor

In this circuit, the three components are all in series with the voltage source. The governing differential equation can be found by substituting into Kirchhoff's voltage law (KVL) the constitutive equation for each of the three elements. From KVL,

$$v_R + v_L + v_C = v(t)$$

where $v_R, v_L, v_C$ are the voltages across R, L and C respectively and $v(t)$ is the time varying voltage from the source. Substituting in the constitutive equations,

$$Ri(t) + L\frac{di}{dt} + \frac{1}{C}\int_{-\infty}^{\tau=t} i(\tau)\, d\tau = v(t)$$

For the case where the source is an unchanging voltage, differentiating and dividing by L leads to the second order differential equation:

$$\frac{d^2i(t)}{dt^2} + \frac{R}{L}\frac{di(t)}{dt} + \frac{1}{LC}i(t) = 0$$

This can usefully be expressed in a more generally applicable form:

$$\frac{d^2i(t)}{dt^2} + 2\alpha\frac{di}{dt} + {\omega_0}^2 i(t) = 0$$

$\alpha$ and $\omega_0$ are both in units of angular frequency. $\alpha$ is called the *neper frequency*, or *attenuation*, and is a measure of how fast the transient response of the circuit will die away after the stimulus has been removed. Neper occurs in the name because the units can also be considered to be nepers per second, neper being a unit of attenuation. $\omega_0$ is the angular resonance frequency and is discussed later.

For the case of the series RLC circuit these two parameters are given by:

$$\alpha = \frac{R}{2L} \text{ and } \omega_0 = \frac{1}{\sqrt{LC}}$$

A useful parameter is the *damping factor*, ζ which is defined as the ratio of these two,

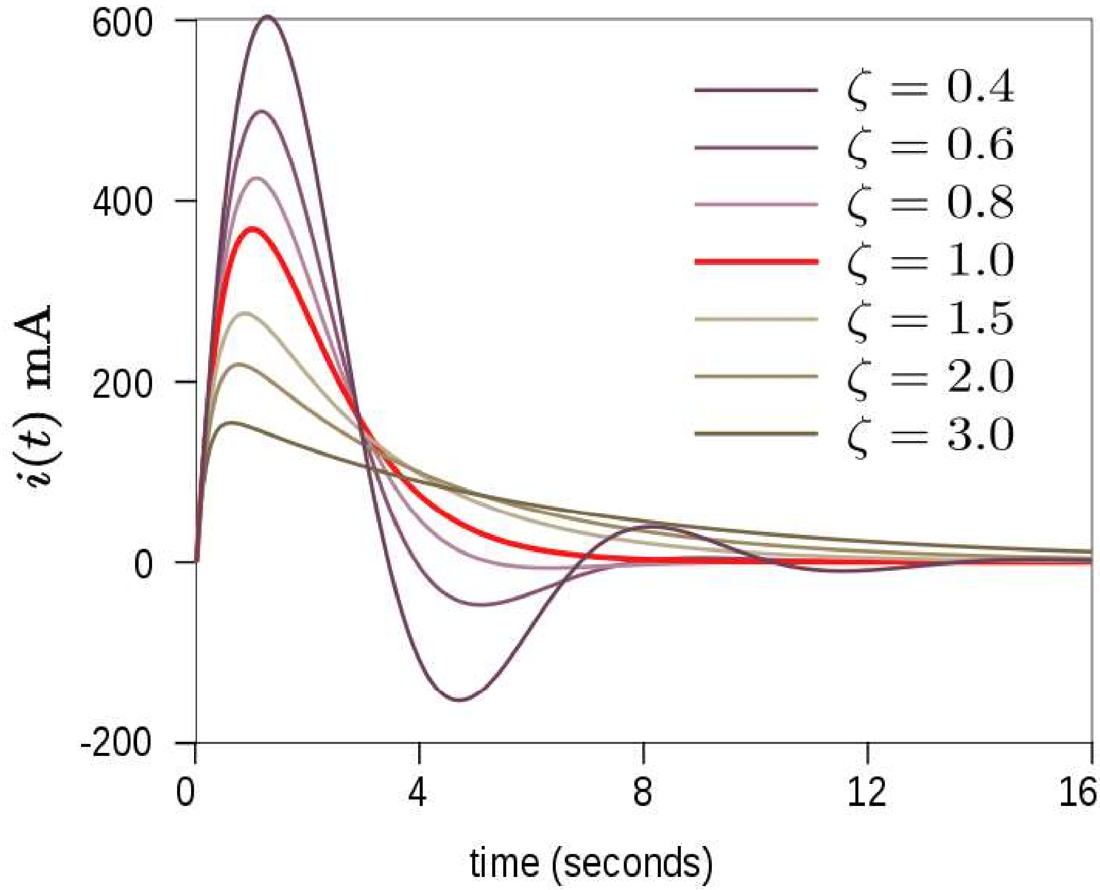
$$\zeta = \frac{\alpha}{\omega_0}$$

In the case of the series RLC circuit, the damping factor is given by,

$$\zeta = \frac{R}{2}\sqrt{\frac{C}{L}}$$

The value of the damping factor determines the type of transient that the circuit will exhibit. Some authors do not use $\zeta$ and call $\alpha$ the damping factor.

**Transient response**




Plot showing underdamped and overdamped responses of a series RLC circuit. The critical damping plot is the bold red curve. The plots are normalised for L=1, C=1 and $\omega_0=1$

The differential equation for the circuit solves in three different ways depending on the value of $\zeta$. These are underdamped ($\zeta<1$), overdamped ($\zeta>1$) and critically damped ($\zeta=1$). The differential equation has the characteristic equation,

$$s^2 + 2\alpha s + {\omega_0}^2 = 0$$

The roots of the equation in *s* are,

$$s_1 = -\alpha + \sqrt{\alpha^2 - {\omega_0}^2}$$
$$s_2 = -\alpha - \sqrt{\alpha^2 - {\omega_0}^2}$$

The general solution of the differential equation is an exponential in either root or a linear superposition of both,

$$i(t) = A_1 e^{s_1 t} + A_2 e^{s_2 t}$$

The coefficients $A_1$ and $A_2$ are determined by the boundary conditions of the specific problem being analysed. That is, they are set by the values of the currents and voltages in the circuit at the onset of the transient and the presumed value they will settle to after infinite time.

The overdamped response ($\zeta > 1$) is,

$$i(t) = A_1 e^{-\omega_0(\zeta + \sqrt{\zeta^2 - 1})t} + A_2 e^{-\omega_0(\zeta - \sqrt{\zeta^2 - 1})t}$$

The overdamped response is a decay of the transient current without oscillation.

The underdamped response ($\zeta < 1$) is,

$$i(t) = B_1 e^{-\alpha t} \cos(\omega_d t) + B_2 e^{-\alpha t} \sin(\omega_d t)$$

By applying standard trigonometric identities the two trigonometric functions may be expressed as a single sinusoid with phase shift,

$$i(t) = B_3 e^{-\alpha t} \sin(\omega_d t + \varphi)$$

The underdamped response is a decaying oscillation at frequency $\omega_d$. The oscillation decays at a rate determined by the attenuation $\alpha$. The exponential in $\alpha$describes the envelope of the oscillation. $B_1$ and $B_2$ (or $B_3$ and the phase shift $\varphi$in the second form) are arbitrary constants determined by boundary conditions. The frequency $\omega_d$is given by,

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} = \omega_0 \sqrt{1 - \zeta^2}$$

This is called the damped resonance frequency or the damped natural frequency. It is the frequency the circuit will naturally oscillate at if not driven by an external source. The resonance frequency, $\omega_0$, which is the frequency at which the circuit will resonate when driven by an external oscillation, may often be referred to as the undamped resonance frequency to distinguish it.

The critically damped response ($\zeta = 1$) is,

$$i(t) = D_1 t e^{-\alpha t} + D_2 e^{-\alpha t}$$

The critically damped response represents the circuit response that decays in the fastest possible time without going into oscillation. This consideration is important in control systems where it is required to reach the desired state as quickly as possible without overshooting. $D_1$ and $D_2$ are arbitrary constants determined by boundary conditions.

## Laplace domain

The series RLC can be analyzed for both transient and steady AC state behavior using the Laplace transform. If the voltage source above produces a waveform with Laplace-transformed *V*(*s*) (where *s* is the complex frequency $s = \sigma + i\omega$), KVL can be applied in the Laplace domain:

$$V(s) = I(s)\left(R + Ls + \frac{1}{Cs}\right)$$

where *I*(*s*) is the Laplace-transformed current through all components. Solving for *I*(*s*):

$$I(s) = \frac{1}{R + Ls + \frac{1}{Cs}} V(s)$$

And rearranging, we have that

$$I(s) = \frac{s}{L\left(s^2 + \frac{R}{L}s + \frac{1}{LC}\right)} V(s)$$

**Laplace admittance**

Solving for the Laplace admittance *Y*(*s*):

$$Y(s) = \frac{I(s)}{V(s)} = \frac{s}{L\left(s^2 + \frac{R}{L}s + \frac{1}{LC}\right)}$$

Simplifying using parameters α and $\omega_o$ defined in the previous section, we have

$$Y(s) = \frac{I(s)}{V(s)} = \frac{s}{L\left(s^2 + 2\alpha s + {\omega_0}^2\right)}$$

**Poles and zeros**

The zeros of *Y*(*s*) are those values of *s* such that *Y*(*s*) = 0:

$s = 0$ and $|s| \rightarrow \infty$

The poles of *Y*(*s*) are those values of *s* such that $Y(s) \rightarrow \infty$. By the quadratic formula, we find

$$s = -\alpha \pm \sqrt{\alpha^2 - {\omega_0}^2}$$

The poles of $Y(s)$ are identical to the roots $s_1$ and $s_2$ of the characteristic polynomial of the differential equation in the section above.

## General solution

For an arbitrary E(t), the solution obtained by inverse transform of I(s) is:

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}\left(\cos\omega_d\tau - \frac{\alpha}{\omega_d}\sin\omega_d\tau\right)d\tau$$

in the underdamped case ($\omega_0 > \alpha$)

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}(1-\alpha\tau)d\tau$$

in the critically damped case ($\omega_0 = \alpha$)

$$I(t) = \frac{1}{L}\int_0^t E(t-\tau)e^{-\alpha\tau}\left(\cosh\omega_r\tau - \frac{\alpha}{\omega_r}\sinh\omega_r\tau\right)d\tau$$

in the overdamped case ($\omega_0 < \alpha$)

where $\omega_r = \sqrt{\alpha^2 - {\omega_0}^2}$, and cosh and sinh are the usual hyperbolic functions.

**Sinusoidal steady state**

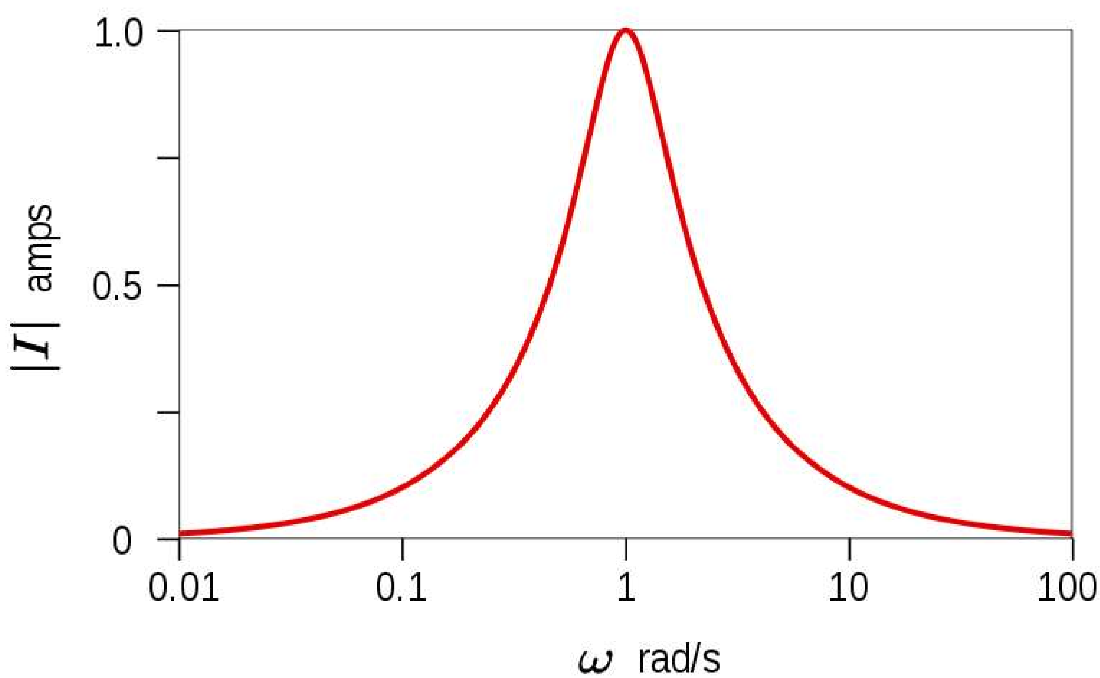



**Figure 4.** *Sinusoidal steady-state analysis*

normalised to $R$ = 1 ohm, $C$ = 1 farad, $L$ = 1 henry, and $V$ = 1.0 volt

Sinusoidal steady state is represented by letting $s = i\omega$

Taking the magnitude of the above equation with this substitution:

$$|Y(s = i\omega)| = \frac{1}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}.$$

and the current as a function of $\omega$ can be found from

$$|I(i\omega)| = |Y(i\omega)||V(i\omega)|.$$

Note that there is a peak at $i_{mag}(\omega) = 1$. This is known as the resonance frequency. Solving for this value:

$$\omega_0 = \frac{1}{\sqrt{LC}}.$$

### *Parallel RLC circuit*

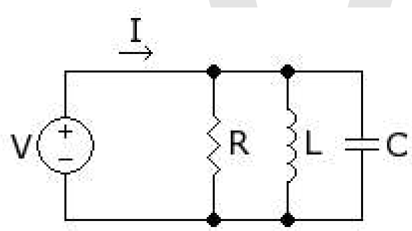



**Figure 5.** RLC parallel circuit

**V** - the voltage of the power source
**I** - the current in the circuit
**R** - the resistance of the resistor
**L** - the inductance of the inductor
**C** - the capacitance of the capacitor

The properties of the parallel RLC circuit can be obtained from the duality relationship of electrical circuits and considering that the parallel RLC is the dual impedance of a series RLC. From this consideration is immediately obtained the result that the differential equations describing this circuit will be identical to the general form of those describing a series RLC.

For the parallel circuit, the attenuation $\alpha$ is given by

$$\alpha = \frac{1}{2RC}$$

and the damping factor is consequently

$$\zeta = \frac{1}{2R}\sqrt{\frac{L}{C}}$$

This is the inverse of the expression for $\zeta$ in the series circuit. Likewise, the other scaled parameters, fractional bandwidth and $Q$ are also the inverse of each other. This means that a wide band, low $Q$ circuit in one topology will become a narrow band, high $Q$ circuit in the other topology when constructed from components with identical values. The $Q$ and fractional bandwidth of the parallel circuit are given by

$$Q = R\sqrt{\frac{C}{L}} \text{ and } F_{\mathrm{b}} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

## Frequency domain

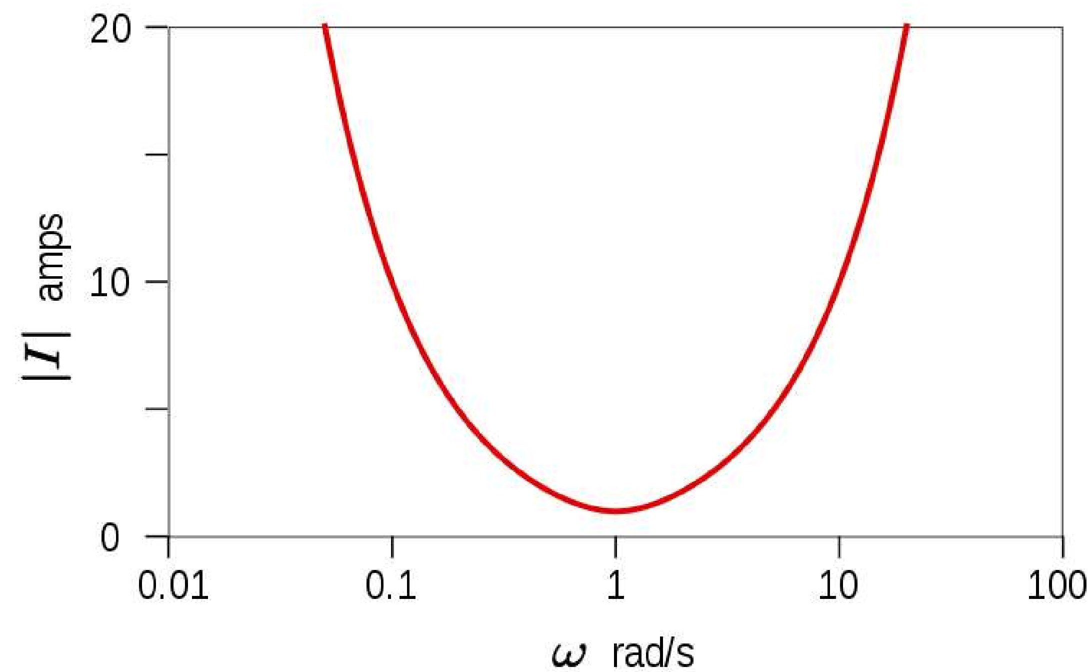



**Figure 6.** *Sinusoidal steady-state analysis*s

normalised to $R$ = 1 ohm, $C$ = 1 farad, $L$ = 1 henry, and $V$ = 1.0 volt

The complex admittance of this circuit is given by adding up the admittances of the components:

$$\frac{1}{Z} = \frac{1}{Z_L} + \frac{1}{Z_C} + \frac{1}{Z_R} = \frac{1}{j\omega L} + j\omega C + \frac{1}{R}$$

The change from a series arrangement to a parallel arrangement results in the circuit having a peak in impedance at resonance rather than a minimum, so the circuit is an antiresonator.

The graph opposite shows that there is a minimum in the frequency response of the current at the resonance frequency $\omega_0 = \frac{1}{\sqrt{LC}}$ when the circuit is driven by a constant voltage. On the other hand, if driven by a constant current, there would be a maximum in the voltage which would follow the same curve as the current in the series circuit.

### *Other configurations*

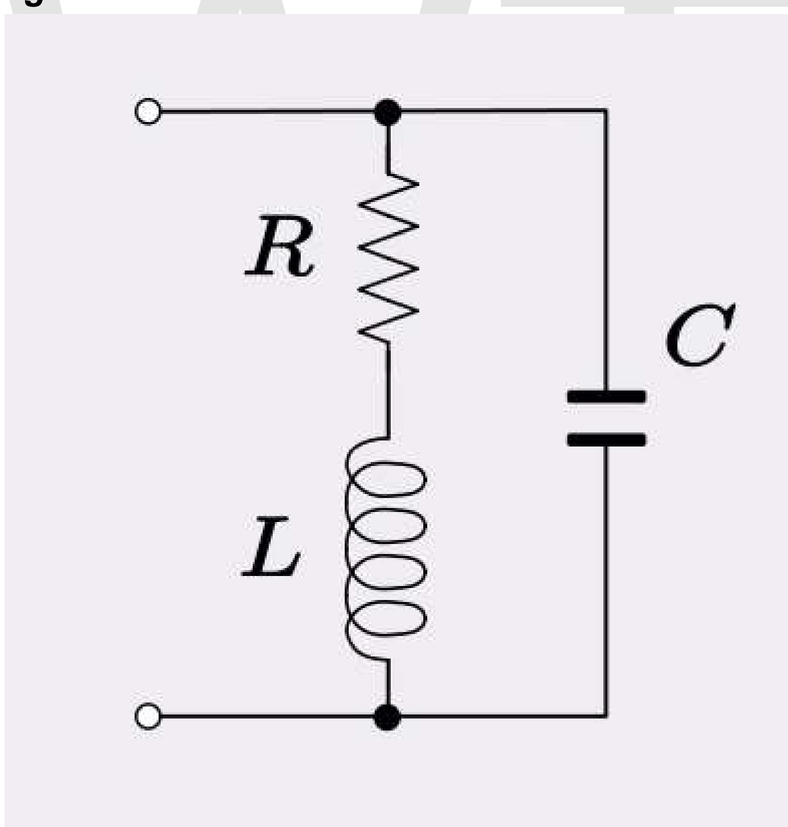



**Fig. 7.** RLC parallel circuit with resistance in series with the inductor

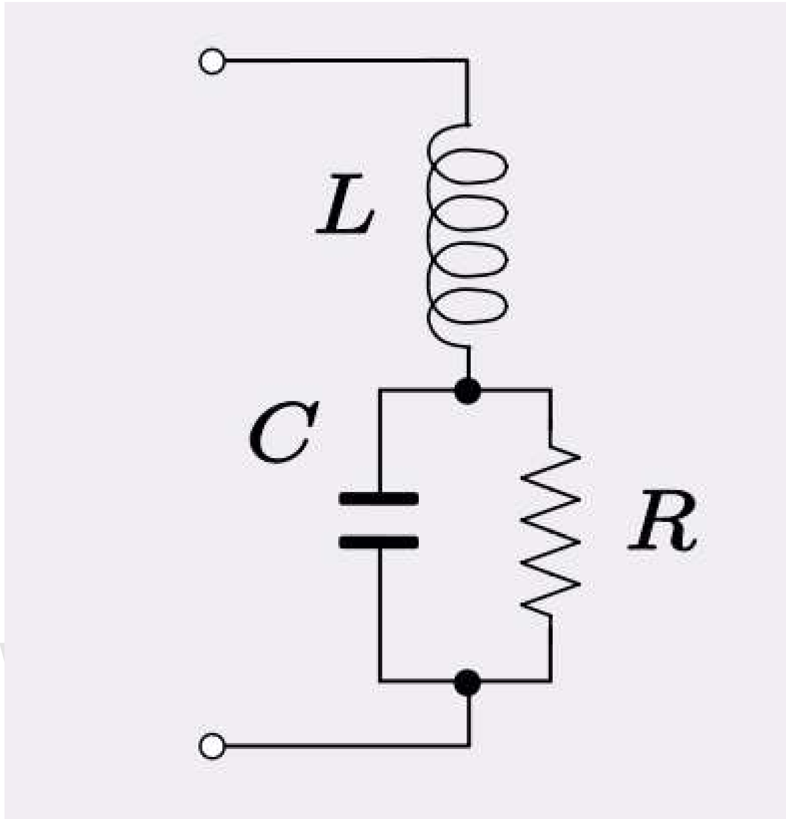



**Fig. 8.** RLC series circuit with resistance in parallel with the capacitor

A series resistor with the inductor in a parallel LC circuit as shown in figure 7 is a topology commonly encountered where there is a need to take into account the resistance of the coil winding. Parallel LC circuits are frequently used for bandpass filtering and the $Q$ is largely governed by this resistance. The resonant frequency of this circuit is,

$$\omega_0 = \sqrt{\frac{1}{LC} - \left(\frac{R}{L}\right)^2}$$

This is the resonant frequency of the circuit defined as the frequency at which the admittance has zero imaginary part. The frequency that appears in the generalised form of the characteristic equation (which is the same for this circuit as previously)

$$s^2 + 2\alpha s + {\omega'_0}^2 = 0$$

is not the same frequency. In this case it is the undamped resonant frequency

$$\omega'_0 = \sqrt{\frac{1}{LC}}$$

In the same vein, a resistor in parallel with the capacitor in a series LC circuit can be used to represent a capacitor with a lossy dielectric. This configuration is shown in figure 8. The resonant frequency in this case is given by

$$\omega_0 = \sqrt{\frac{1}{LC} - \frac{1}{(RC)^2}}$$

## *Applications*

### Variable tuned circuits

A very frequent use of these circuits is in the tuning circuits of analogue radios. Adjustable tuning is commonly achieved with a parallel plate variable capacitor which allows the value of C to be changed and tune to stations on different frequencies. For the IF stage in the radio where the tuning is preset in the factory the more usual solution is an adjustable core in the inductor to adjust L. In this design the core (made of a high permeability material that has the effect of increasing inductance) is threaded so that it can be screwed further in, or screwed further out of the inductor winding as required.

### Filters

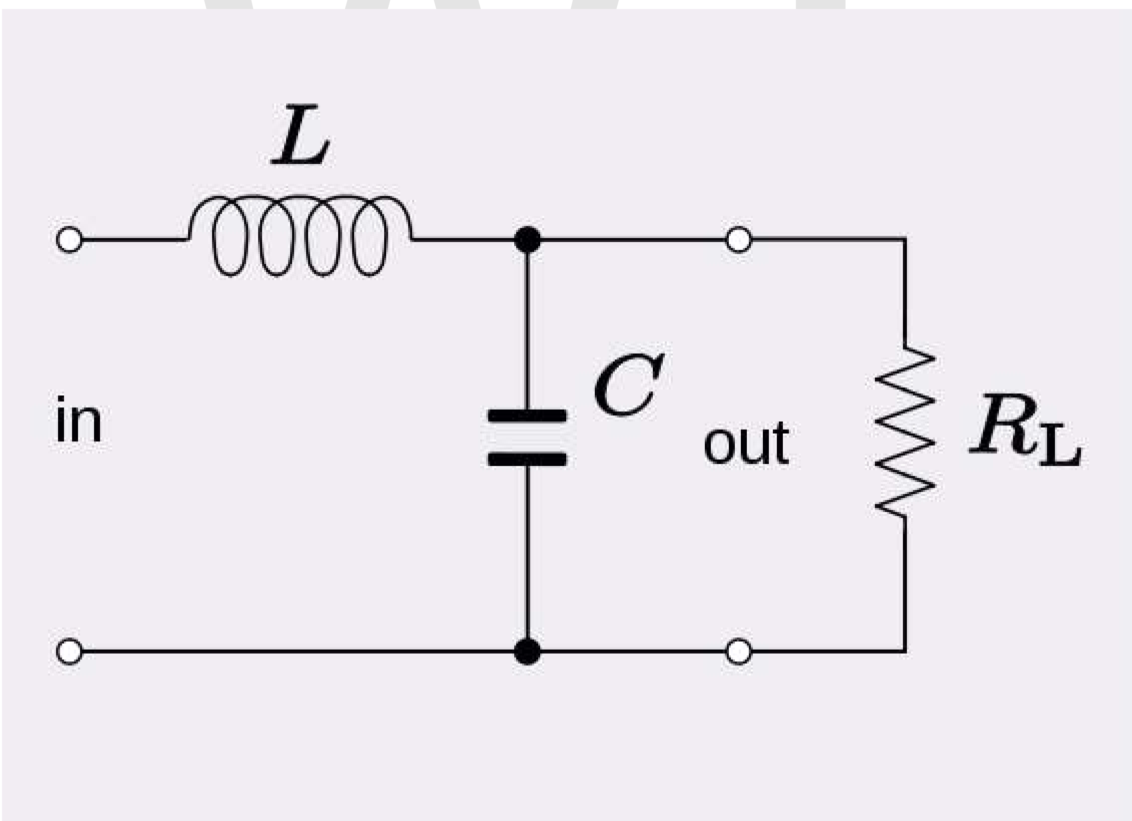



**Fig. 9.** RLC circuit as a low-pass filter

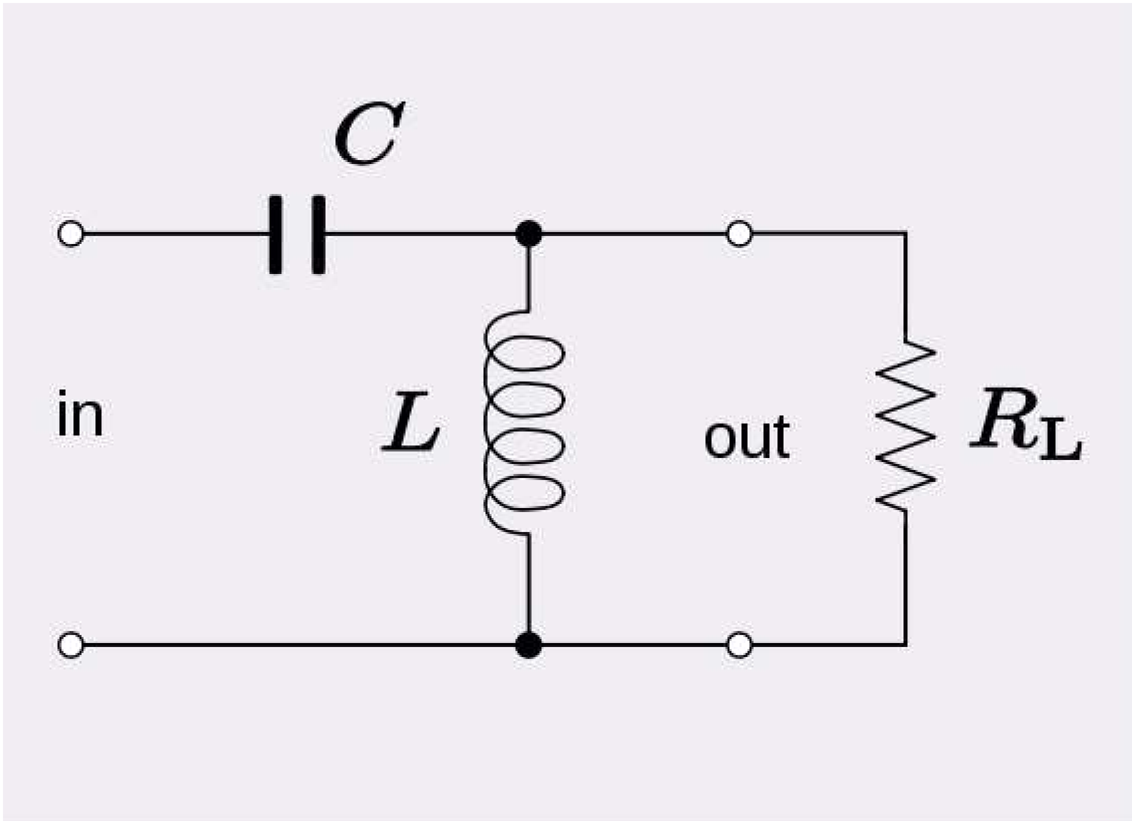



**Fig. 10.** RLC circuit as a high-pass filter

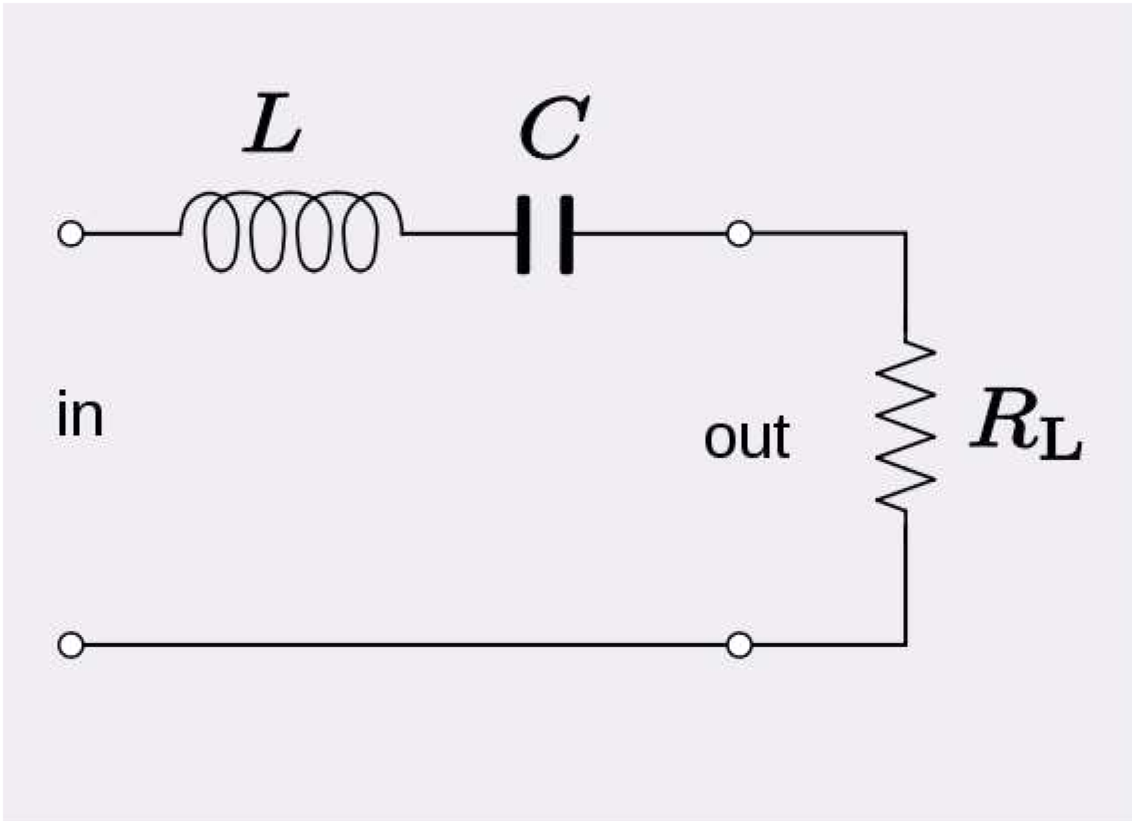



**Fig. 11.** RLC circuit as a series band-pass filter in series with the line

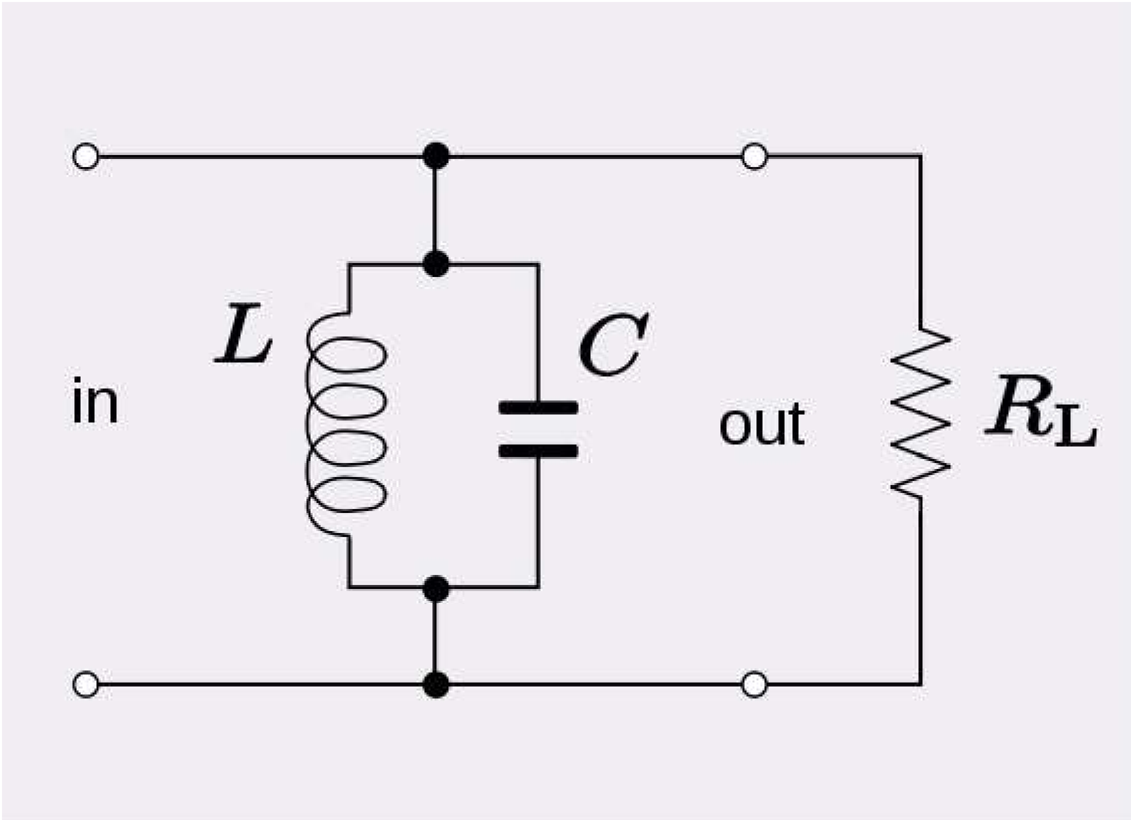



**Fig. 12.** RLC circuit as a parallel band-pass filter in shunt across the line

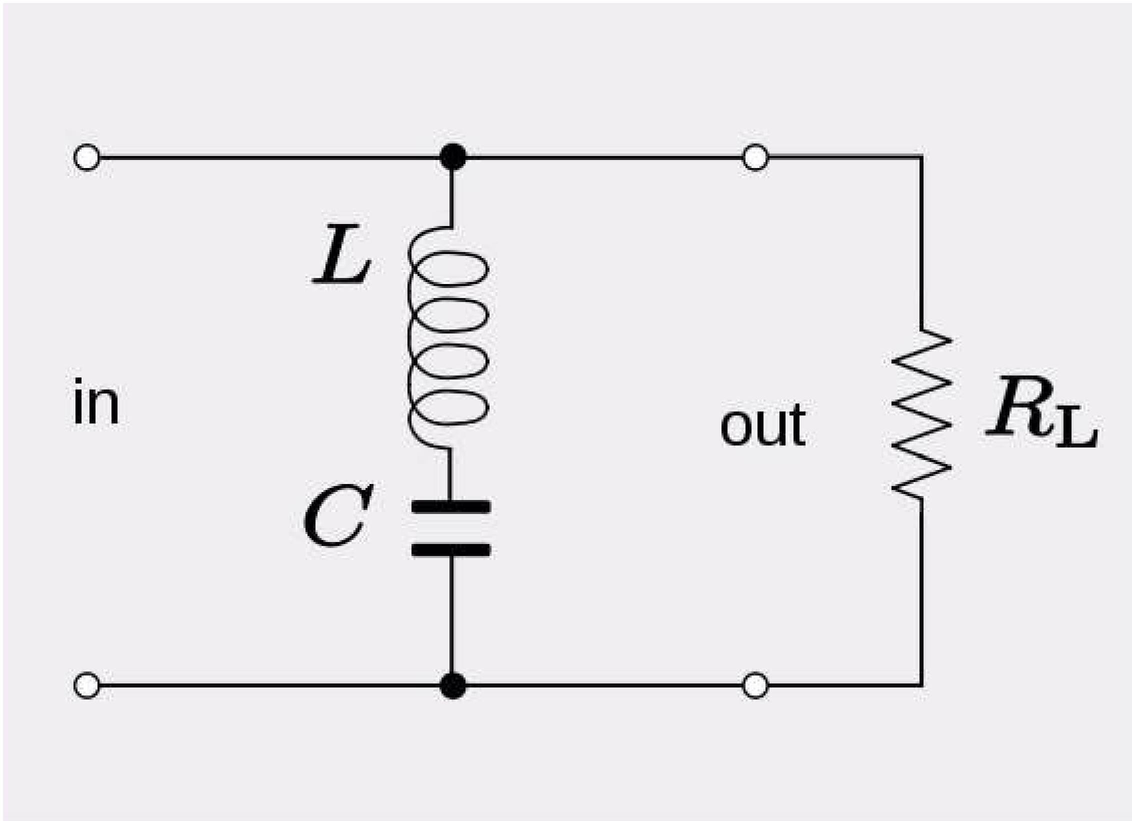



**Fig. 13.** RLC circuit as a series band-stop filter in shunt across the line

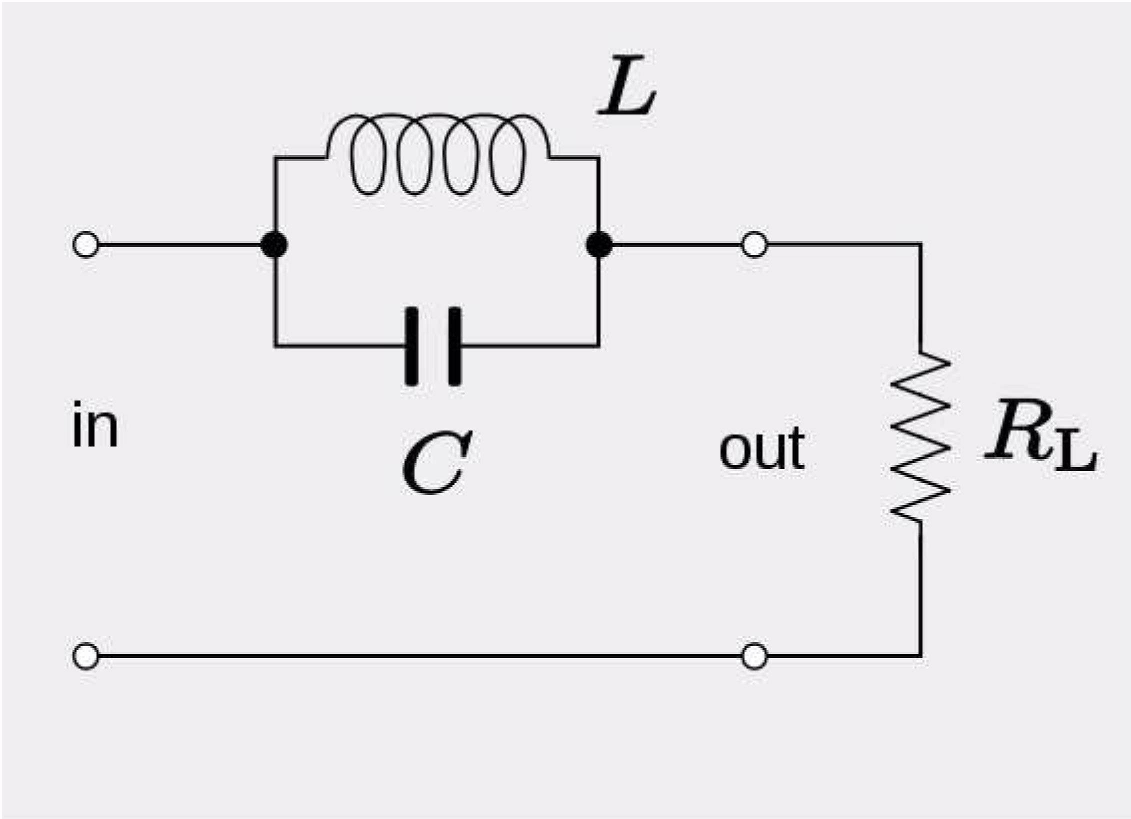



**Fig. 14.** RLC circuit as a parallel band-stop filter in series with the line

In the filtering application, the resistor *R* becomes the load that the filter is working into. The value of the damping factor is chosen based on the desired bandwidth of the filter. For a wider bandwidth, a larger value of the damping factor is required (and vice versa). The three components give the designer three degrees of freedom. Two of these are required to set the bandwidth and resonant frequency. The designer is still left with one which can be used to scale *R*, *L* and *C* to convenient practical values. Alternatively, *R* may be predetermined by the external circuitry which will use the last degree of freedom.

Low-pass filter

An RLC circuit can be used as a low-pass filter. The circuit configuration is shown in figure 9. The corner frequency, that is, the frequency of the 3dB point, is given by

$$\omega_c = \frac{1}{\sqrt{LC}}$$

This is also the bandwidth of the filter. The damping factor is given by

$$\zeta = \frac{1}{2R_L}\sqrt{\frac{L}{C}}$$

High-pass filter

A high-pass filter is shown in figure 10. The corner frequency is the same as the low-pass filter

$$\omega_c = \frac{1}{\sqrt{LC}}$$

The filter has a stop-band of this width.

Band-pass filter

A band-pass filter can be formed with an RLC circuit by either placing a series LC circuit in series with the load resistor or else by placing a parallel LC circuit in parallel with the load resistor. These arrangements are shown in figures 11 and 12 respectively. The centre frequency is given by

$$\omega_c = \frac{1}{\sqrt{LC}}$$

and the bandwidth for the series circuit is

$$\Delta\omega = \frac{R_L}{L}$$

The shunt version of the circuit is intended to be driven by a high impedance source, that is, a constant current source. Under those conditions the bandwidth is

$$\Delta\omega = \frac{1}{CR_L}$$

Band-stop filter

Figure 13 shows a band-stop filter formed by a series LC circuit in shunt across the load. Figure 14 is a band-stop filter formed by a parallel LC circuit in series with the load. The first case requires a high impedance source so that the current is diverted into the resonator when it becomes low impedance at resonance. The second case requires a low impedance source so that the voltage is dropped across the antiresonator when it becomes high impedance at resonance.

## Oscillators

For applications in oscillator circuits, it is generally desirable to make the attenuation (or equivalently, the damping factor) as small as possible. In practice, this objective requires making the circuit's resistance *R* as small as physically possible for a series circuit, or alternatively increasing *R* to as much as possible for a parallel circuit. In either case, the *RLC circuit* becomes a good approximation to an ideal LC circuit. However, for very low attenuation circuits (high Q-factor) circuits, issues such as dielectric losses of coils and capacitors can become important.

In an oscillator circuit

$$\alpha \ll \omega_0.$$

or equivalently

$$\zeta \ll 1.$$

As a result

$$\omega_d \approx \omega_0.$$

### Voltage multiplier

In a series RLC circuit at resonance, the current is limited only by the resistance of the circuit

$$I = \frac{V}{R}$$

If R is small, consisting only of the inductor winding resistance say, then this current will be large. It will drop a voltage across the inductor of

$$V_\mathrm{L} = \frac{V}{R}\omega_0 L$$

An equal magnitude voltage will also be seen across the capacitor but in antiphase to the inductor. If R can be made sufficiently small, these voltages can be several times the input voltage. The voltage ratio is, in fact, the *Q* of the circuit,

$$\frac{V_\mathrm{L}}{V} = Q$$

A similar effect is observed with currents in the parallel circuit. Even though the circuit appears as high impedance to the external source, there is a large current circulating in the internal loop of the parallel inductor and capacitor.

## Pulse discharge circuit

An overdamped series RLC circuit can be used as a pulse discharge circuit. Often it is useful to know the values of components that could be used to produce a waveform this is described by the form:

$$I(t) = I_0(e^{-\alpha t} - e^{-\beta t})$$

Such a circuit could consist of an energy storage capacitor, a load in the form of a resistance, some circuit inductance and a switch - all in series. The initial conditions are that the capacitor is at voltage $V_0$ and there is no current flowing in the inductor. If the inductance $L$ is known, then the remaining parameters are given by the following - Capacitance:

$$C = \frac{1}{L\alpha\beta}$$

Resistance (total of circuit and load):

$$R = L(\alpha + \beta)$$

Initial terminal voltage of capacitor:

$$V_0 = -I_0 L\alpha\beta(\frac{1}{\beta} - \frac{1}{\alpha})$$

Rearranging for the case where R is known - Capacitance:

$$C = \frac{(\alpha + \beta)}{R\alpha\beta}$$

Inductance (total of circuit and load):

$$L = \frac{R}{(\alpha + \beta)}$$

Initial terminal voltage of capacitor:

$$V_0 = \frac{-I_0 R\alpha\beta}{(\alpha + \beta)}(\frac{1}{\beta} - \frac{1}{\alpha})$$

**Chapter-7**

# p-n Junction


A silicon p–n junction with no applied voltage.

A **p–n junction** is formed by joining P-type and N-type semiconductors together in very close contact. The term *junction* refers to the boundary interface where the two regions of the semiconductor meet. If they were constructed of two separate pieces this would introduce a grain boundary, so p–n junctions are created in a single crystal of semiconductor by doping, for example by ion implantation, diffusion of dopants, or by epitaxy (growing a layer of crystal doped with one type of dopant on top of a layer of crystal doped with another type of dopant).

P-N junctions are elementary "building blocks" of almost all semiconductor electronic devices such as diodes, transistors, solar cells, LEDs, and integrated circuits; they are the active sites where the electronic action of the device takes place. For example, a common type of transistor, the bipolar junction transistor, consists of two p–n junctions in series, in the form n–p–n or p–n–p.

The discovery of the p–n junction is usually attributed to American physicist Russell Ohl of Bell Laboratories.

Schottky junction is a special case of a p-n junction, where metal serves the role of the n-type semiconductor.

### *Manufacture*

Normally, p–n junctions are manufactured from a single crystal with different dopant concentrations diffused across it. Creating a semiconductor from two separate pieces of material would introduce a grain boundary between the semiconductors which severely inhibits its utility by scattering the electrons and holes..

Note that, in the case of solar cells, polycrystalline silicon is often used to reduce expense, despite the lower efficiency caused by the grain boundaries. These boundaries are not related to the p-n junctions in the cell. If they would be the same spacially, the disturbing effects would make the solar cell useless.

## *Properties of a p-n junction*

The p–n junction possesses some interesting properties which have useful applications in modern electronics. A p-doped semiconductor is relatively conductive. The same is true of an n-doped semiconductor, but the junction between them can become depleted of charge carriers, and hence nonconductive, depending on the relative voltages of the two semiconductor regions. By manipulating this non-conductive layer, p–n junctions are commonly used as diodes: circuit elements that allow a flow of electricity in one direction but not in the other (opposite) direction. This property is explained in terms of *forward bias* and *reverse bias*, where the term *bias* refers to an application of electric voltage to the p–n junction.

## *Equilibrium (zero bias)*

In a p–n junction, without an external applied voltage, an equilibrium condition is reached in which a potential difference is formed across the junction. This potential difference is called built-in potential $V_{bi}$.

After joining p-type and n-type semiconductors, electrons near the p–n interface tend to diffuse into the p region. As electrons diffuse, they leave positively charged ions (donors) in the n region. Similarly, holes near the p–n interface begin to diffuse into the n-type region leaving fixed ions (acceptors) with negative charge. The regions nearby the p–n interfaces lose their neutrality and become charged, forming the space charge region or depletion layer (see figure A).

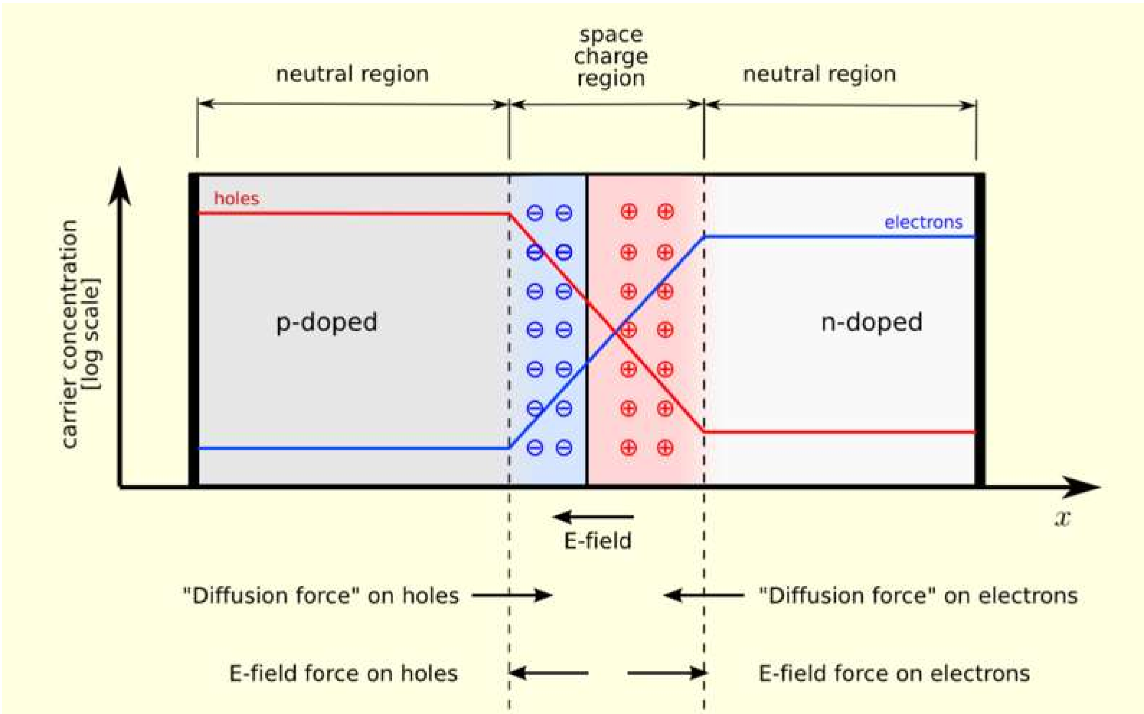



**Figure A.** A p–n junction in thermal equilibrium with zero bias voltage applied. Electrons and holes concentration are reported respectively with blue and red lines. Gray regions are charge neutral. Light red zone is positively charged. Light blue zone is negatively charged. The electric field is shown on the bottom, the electrostatic force on electrons and holes and the direction in which the diffusion tends to move electrons and holes.

The electric field created by the space charge region opposes the diffusion process for both electrons and holes. There are two concurrent phenomena: the diffusion process that tends to generate more space charge, and the electric field generated by the space charge that tends to counteract the diffusion. The carrier concentration profile at equilibrium is shown in figure A with blue and red lines. Also shown are the two counterbalancing phenomena that establish equilibrium.

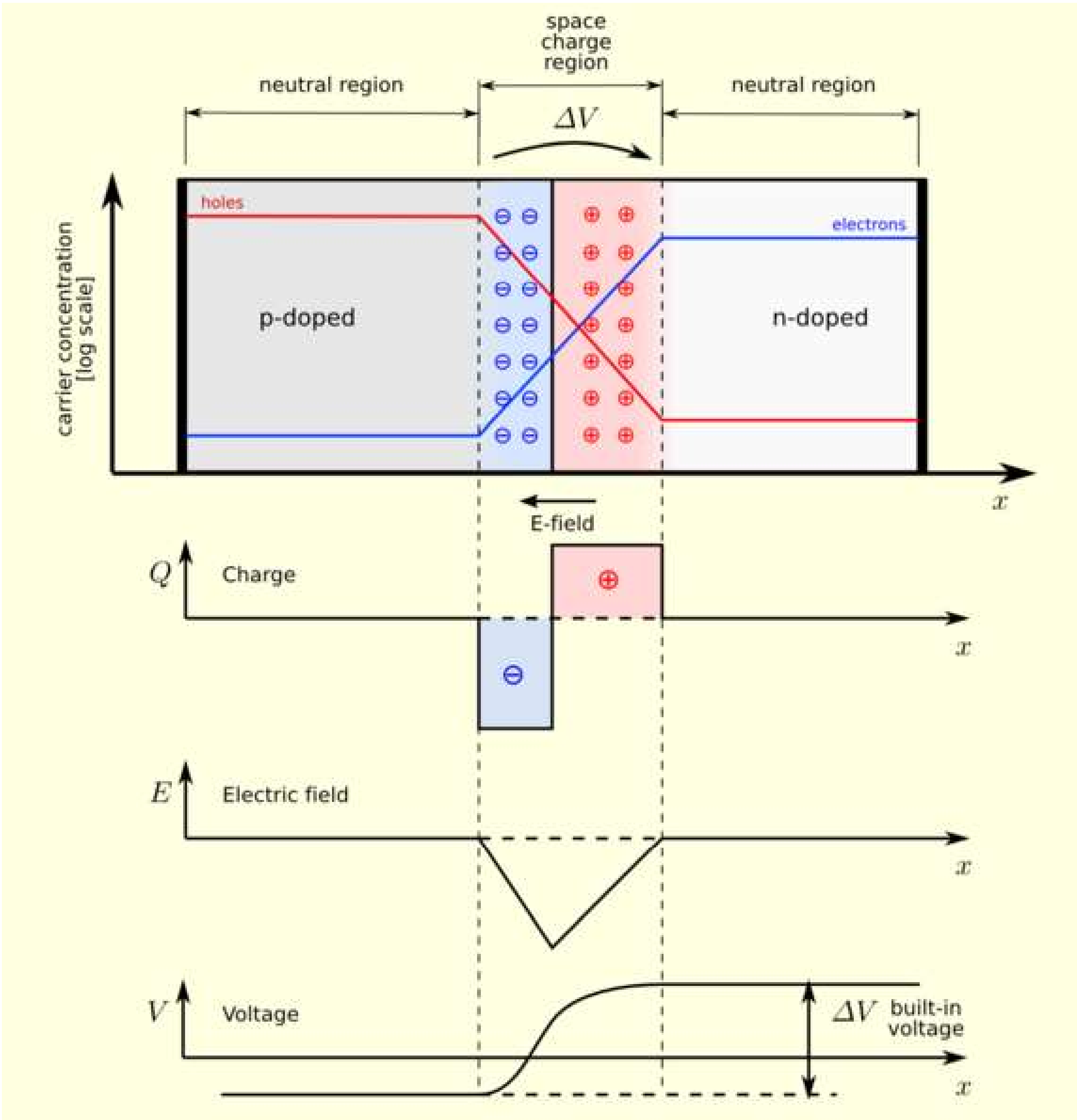
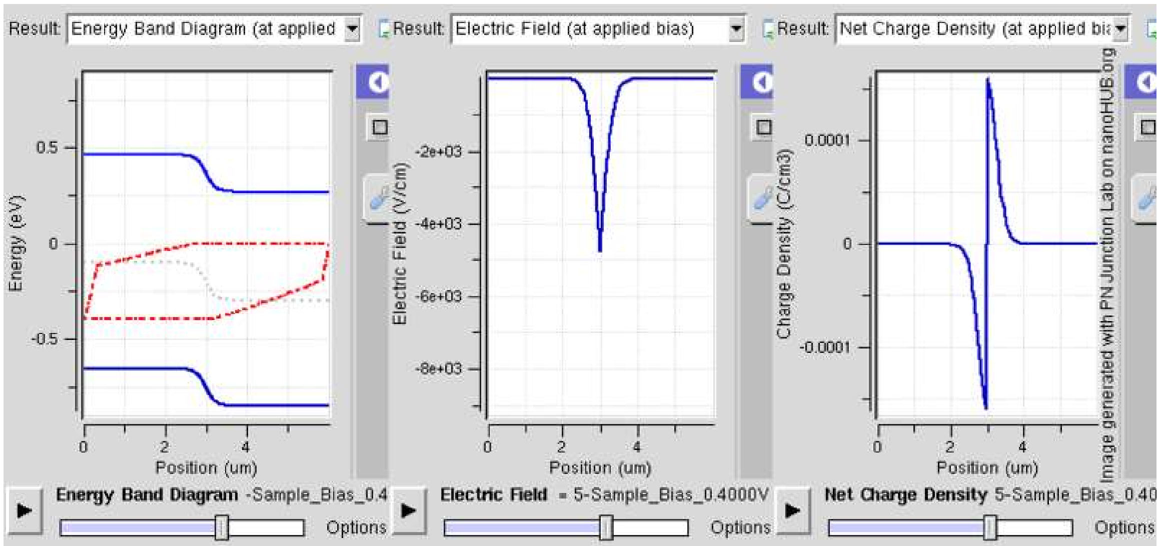



**Figure B.** A p–n junction in thermal equilibrium with zero bias voltage applied. Under the junction, plots for the charge density, the electric field and the voltage are reported.

The space charge region is a zone with a net charge provided by the fixed ions (donors or acceptors) that have been left *uncovered* by majority carrier diffusion. When equilibrium is reached, the charge density is approximated by the displayed step function. In fact, the region is completely depleted of majority carriers (leaving a charge density equal to the net doping level), and the edge between the space charge region and the neutral region is quite sharp (see figure B, Q(x) graph). The space charge region has the same magnitude of charge on both sides of the p–n interfaces, thus it extends farther on the less doped side (the n side in figures A and B).

### *Forward bias*

In forward bias, the p-type is connected with the positive terminal and the n-type is connected with the negative terminal.




PN junction operation in forward bias mode showing reducing depletion width. Both p and n junctions are doped at a 1e15/cm3 doping level, leading to built-in potential of ~0.59V. Reducing depletion width can be inferred from the shrinking charge profile, as fewer dopants are exposed with increasing forward bias.

With a battery connected this way, the holes in the P-type region and the electrons in the N-type region are pushed towards the junction. This reduces the width of the depletion zone. The positive charge applied to the P-type material repels the holes, while the negative charge applied to the N-type material repels the electrons. As electrons and holes are pushed towards the junction, the distance between them decreases. This lowers the barrier in potential. With increasing forward-bias voltage, the depletion zone eventually becomes thin enough that the zone's electric field can't counteract charge carrier motion across the p–n junction, consequently reducing electrical resistance. The electrons which cross the p–n junction into the P-type material (or holes which cross into the N-type material) will diffuse in the near-neutral region. Therefore, the amount of minority diffusion in the near-neutral zones determines the amount of current that may flow through the diode.

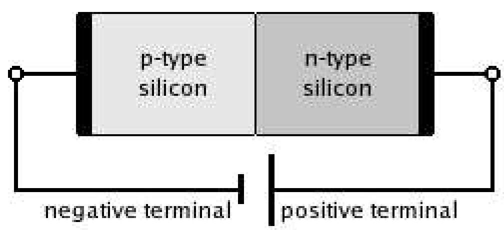
Only majority carriers (electrons in N-type material or holes in P-type) can flow through a semiconductor for a macroscopic length. With this in mind, consider the flow of electrons across the junction. The forward bias causes a force on the electrons pushing them from the N side toward the P side. With forward bias, the depletion region is narrow enough that electrons can cross the junction and *inject* into the P-type material. However, they do not continue to flow through the P-type material indefinitely, because it is energetically favorable for them to recombine with holes. The average length an electron travels through the P-type material before recombining is called the *diffusion length*, and it is typically on the order of microns.

Although the electrons penetrate only a short distance into the P-type material, the electric current continues uninterrupted, because holes (the majority carriers) begin to flow in the opposite direction. The total current (the sum of the electron and hole currents) is constant in space, because any variation would cause charge buildup over time (this is Kirchhoff's current law). The flow of holes from the P-type region into the N-type region is exactly analogous to the flow of electrons from N to P (electrons and holes swap roles and the signs of all currents and voltages are reversed).

Therefore, the macroscopic picture of the current flow through the diode involves electrons flowing through the N-type region toward the junction, holes flowing through the P-type region in the opposite direction toward the junction, and the two species of carriers constantly recombining in the vicinity of the junction. The electrons and holes travel in opposite directions, but they also have opposite charges, so the overall current is in the same direction on both sides of the diode, as required.

The Shockley diode equation models the forward-bias operational characteristics of a p–n junction outside the avalanche (reverse-biased conducting) region.

### *Reverse bias*




A silicon p–n junction in reverse bias.

Reverse biased usually refers to how a diode is used in a circuit. If a diode is reverse biased, the voltage at the cathode is higher than that at the anode. Therefore, no current will flow until the diode breaks down. Connecting the *P-type* region to the *negative* terminal of the battery and the *N-type* region to the *positive* terminal, corresponds to reverse bias. The connections are illustrated in the following diagram:

Because the p-type material is now connected to the negative terminal of the power supply, the 'holes' in the P-type material are pulled away from the junction, causing the width of the depletion zone to increase. Similarly, because the N-type region is connected to the positive terminal, the electrons will also be pulled away from the junction. Therefore the depletion region widens, and does so increasingly with increasing reverse-bias voltage. This increases the voltage barrier causing a high resistance to the flow of charge carriers thus allowing minimal electric current to cross the p–n junction. The increase in resistance of the p-n junction results in the junction behaving as an insulator.

The strength of the depletion zone electric field increases as the reverse-bias voltage increases. Once the electric field intensity increases beyond a critical level, the p–n junction depletion zone breaks-down and current begins to flow, usually by either the Zener or avalanche breakdown processes. Both of these breakdown processes are non-destructive and are reversible, so long as the amount of current flowing does not reach levels that cause the semiconductor material to overheat and cause thermal damage.

This effect is used to one's advantage in zener diode regulator circuits. Zener diodes have a certain - low - breakdown voltage. A standard value for breakdown voltage is for instance 5.6V. This means that the voltage at the cathode can never be more than 5.6V higher than the voltage at the anode, because the diode will break down - and therefore conduct - if the voltage gets any higher. This effectively regulates the voltage over the diode.

Another application where reverse biased diodes are used is in Varicap diodes. The width of the depletion zone of any diode changes with voltage applied. This varies the capacitance of the diode. For more information, refer to the Varicap article.

## *Electrostatics*

For a p-n junction Poisson's equation becomes

$$\Delta\varphi = -\frac{\rho}{\varepsilon} = \frac{q}{\varepsilon}\left(\underbrace{n_0 - p_0}_{\substack{\text{equilibrium concentration}\\ \text{difference of free charges } (\approx 0)}} + \underbrace{N_A - N_D}_{\substack{\text{concentration difference}\\ \text{of acceptor and donor atoms}}}\right)$$

where $\varphi$ is the electric potential, ρ is the charge density, $\varepsilon$ is permittivity and *q* is the magnitude of the electron charge.

Since the total charge on either side of the depletion region must cancel out it is

$$\underbrace{d_p}_{\substack{\text{width of}\\ \text{electric field}\\ \text{within p-side}}} N_A = \underbrace{d_n}_{\substack{\text{width of}\\ \text{electric field}\\ \text{within n-side}}} N_D$$

From the above equations and by deploying basic calculus it can be shown that the total width of the depletion region is

$$d = d_p + d_n = \sqrt{\frac{2\varepsilon}{q}\frac{N_A + N_D}{N_A N_D}\left(\underbrace{V_{bi}}_{\text{built-in voltage}} - \underbrace{V}_{\text{external applied voltage}}\right)}$$

Furthermore, by implementing the Einstein relation and assuming the semiconductor is nondegenerate (i.e. the product $p_0 n_0$ is independent of the Fermi energy) it follows that

$$V_{bi} = \frac{kT}{q}\ln\left(\frac{N_A N_D}{p_0 n_0}\right)$$

where $T$ is the temperature of the semiconductor and $k$ is Boltzmann constant.

### *Summary*

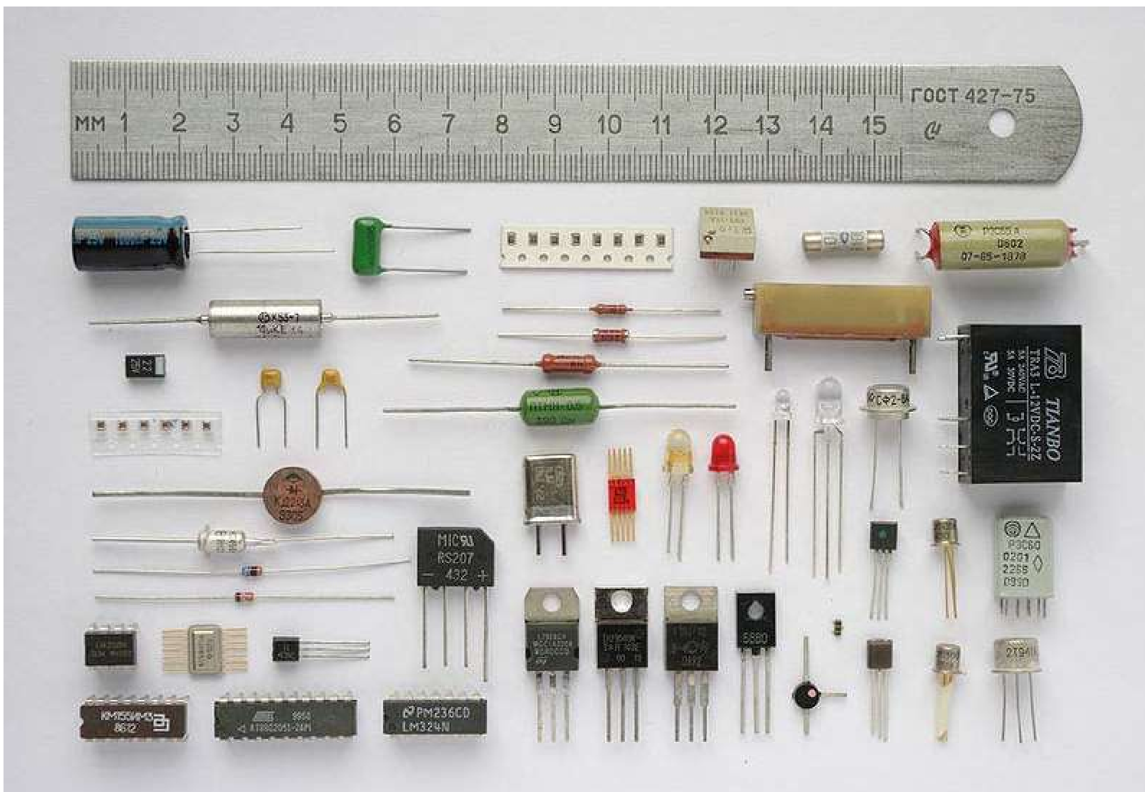
The forward-bias and the reverse-bias properties of the p–n junction imply that it can be used as a diode. A p–n junction diode allows electric charges to flow in one direction, but not in the opposite direction; negative charges (electrons) can easily flow through the junction from n to p but not from p to n and the reverse is true for holes. When the p–n junction is forward biased, electric charge flows freely due to reduced resistance of the p–n junction. When the p–n junction is reverse biased, however, the junction barrier (and therefore resistance) becomes greater and charge flow is minimal.

### *Non-rectifying junctions*

In the above diagrams, contact between the metal wires and the semiconductor material also creates metal-semiconductor junctions called Schottky diodes. In a simplified ideal situation a semiconductor diode would never function, since it would be composed of several diodes connected back-to-front in series. But in practice, surface impurities within the part of the semiconductor which touches the metal terminals will greatly reduce the width of those depletion layers to such an extent that the metal-semiconductor junctions do not act as diodes. These "nonrectifying junctions" behave as ohmic contacts regardless of applied voltage polarity.

Chapter-8

# Electronic Engineering



Electronic components

**Electronics engineering**, also referred to as **electronic engineering**, is an engineering discipline where non-linear and active electrical components such as electron tubes, and semiconductor devices, especially transistors, diodes and integrated circuits, are utilized to design electronic circuits, devices and systems, typically also including passive electrical components and based on printed circuit boards. The term denotes a broad engineering field that covers important subfields such as analog electronics, digital electronics, consumer electronics, embedded systems and power electronics. Electronics engineering deals with implementation of applications, principles and algorithms developed within many related fields, for example solid-state physics, radio engineering, telecommunications, control systems, signal processing, systems engineering, computer engineering, instrumentation engineering, electric power control, robotics, and many others.

The Institute of Electrical and Electronics Engineers (IEEE) is one of the most important and influential organizations for electronics engineers.

## *Relationship to electrical engineering*

Electronics is a subfield within the wider electrical engineering academic subject. An academic degree with a major in electronics engineering can be acquired from some universities, while other universitites use electrical engineering as the subject. The term electrical engineer is still used in the academic world to include electronic engineers. However, some people think the term 'electrical engineer' should be reserved for those having specialized in power and heavy current or high voltage engineering, while others believe that power is just one subset of electrical engineering (and indeed the term 'power engineering' is used in that industry) as well as 'electrical distribution engineering'. Again, in recent years there has been a growth of new separate-entry degree courses such as 'information engineering', 'systems engineering' and 'communication systems engineering', often followed by academic departments of similar name, which are typically not considered as subfields of electronics engineering but of electrical engineering.

Beginning in the 1980s, the term computer engineer was often used to refer to a subifeld of electronic or information engineers. However, Computer Engineering is now considered a subset of Electronics Engineering and computer science and the term is now becoming archaic.

## *History of electronic engineering*

Electronic engineering as a profession sprang from technological improvements in the telegraph industry in the late 19th century and the radio and the telephone industries in the early 20th century. People were attracted to radio by the technical fascination it inspired, first in receiving and then in transmitting. Many who went into broadcasting in the 1920s were only 'amateurs' in the period before World War I.

The modern discipline of electronic engineering was to a large extent born out of telephone, radio, and television equipment development and the large amount of electronic systems development during World War II of radar, sonar, communication systems, and advanced munitions and weapon systems. In the interwar years, the subject was known as radio engineering and it was only in the late 1950s that the term **electronic engineering** started to emerge.

The electronic laboratories (Bell Labs in the United States for instance) created and subsidized by large corporations in the industries of radio, television, and telephone equipment began churning out a series of electronic advances. In 1948, came the transistor and in 1960, the integrated circuit to revolutionize the electronic industry. In the UK, the subject of electronic engineering became distinct from electrical engineering as a university degree subject around 1960. Before this time, students of electronics and related subjects like radio and telecommunications had to enroll in the electrical engineering department of the university as no university had departments of electronics. Electrical engineering was the nearest subject with which electronic engineering could be aligned, although the similarities in subjects covered (except mathematics and electromagnetism) lasted only for the first year of the three-year course.

## Early electronics

In 1893, Nikola Tesla made the first public demonstration of radio communication. Addressing the Franklin Institute in Philadelphia and the National Electric Light Association, he described and demonstrated in detail the principles of radio communication. In 1896, Guglielmo Marconi went on to develop a practical and widely used radio system. In 1904, John Ambrose Fleming, the first professor of electrical Engineering at University College London, invented the first radio tube, the diode. One year later, in 1906, Robert von Lieben and Lee De Forest independently developed the amplifier tube, called the triode.

Electronics is often considered to have begun when Lee De Forest invented the vacuum tube in 1907. Within 10 years, his device was used in radio transmitters and receivers as well as systems for long distance telephone calls. In 1912, Edwin H. Armstrong invented the regenerative feedback amplifier and oscillator; he also invented the superheterodyne radio receiver and could be considered the father of modern radio. Vacuum tubes remained the preferred amplifying device for 40 years, until researchers working for William Shockley at Bell Labs invented the transistor in 1947. In the following years, transistors made small portable radios, or transistor radios, possible as well as allowing more powerful mainframe computers to be built. Transistors were smaller and required lower voltages than vacuum tubes to work.

Before the invention of the integrated circuit in 1959, electronic circuits were constructed from discrete components that could be manipulated by hand. These non-integrated circuits consumed much space and power, were prone to failure and were limited in speed although they are still common in simple applications. By contrast, integrated circuits packed a large number — often millions — of tiny electrical components, mainly transistors, into a small chip around the size of a coin.

## Tubes or valves

## The vacuum tube detector

The invention of the triode amplifier, generator, and detector made audio communication by radio practical. (Reginald Fessenden's 1906 transmissions used an electro-mechanical alternator.) The first known radio news program was broadcast 31 August 1920 by station 8MK, the unlicensed predecessor of WWJ (AM) in Detroit, Michigan. Regular wireless broadcasts for entertainment commenced in 1922 from the Marconi Research Centre at Writtle near Chelmsford, England.

While some early radios used some type of amplification through electric current or battery, through the mid 1920s the most common type of receiver was the crystal set. In the 1920s, amplifying vacuum tubes revolutionized both radio receivers and transmitters.

## Television

In 1928 Philo Farnsworth made the first public demonstration of a purely electronic television. During the 1930s several countries began broadcasting, and after World War II it spread to millions of receivers, eventually worldwide. Ever since then, electronics have been fully present in television devices.

Modern televisions and video displays have evolved from bulky electron tube technology to use more compact devices, such as plasma and LCD displays. The trend is for even lower power devices such as the organic light-emitting diode displays, and it is most likely to replace the LCD and plasma technologies.

## Radar and radio location

During World War II many efforts were expended in the electronic location of enemy targets and aircraft. These included radio beam guidance of bombers, electronic counter measures, early radar systems etc. During this time very little if any effort was expended on consumer electronics developments.

## Computers

A computer is a programmable machine that receives input, stores and manipulates data, and provides output in a useful format.

Although mechanical examples of computers have existed through much of recorded human history, the first electronic computers were developed in the mid-20th century (1940–1945). These were the size of a large room, consuming as much power as several hundred modern personal computers (PCs). Modern computers based on integrated circuits are millions to billions of times more capable than the early machines, and occupy a fraction of the space. Simple computers are small enough to fit into small pocket devices, and can be powered by a small battery. Personal computers in their various forms are icons of the Information Age and are what most people think of as "computers". However, the embedded computers found in many devices from MP3 players to fighter aircraft and from toys to industrial robots are the most numerous.

The ability to store and execute lists of instructions called programs makes computers extremely versatile, distinguishing them from calculators. The Church–Turing thesis is a mathematical statement of this versatility: any computer with a certain minimum capability is, in principle, capable of performing the same tasks that any other computer can perform. Therefore computers ranging from a netbook to a supercomputer are all able to perform the same computational tasks, given enough time and storage capacity.

### Microprocessors

In 1969, Ted Hoff conceived the commercial microprocessor at Intel and thus ignited the development of the personal computer. Hoff's invention was part of an order by a Japanese company for a desktop programmable electronic calculator, which Hoff wanted to build as cheaply as possible. The first realization of the microprocessor was the Intel 4004, a 4-bit processor, in 1969, but only in 1973 did the Intel 8080, an 8-bit processor, make the building of the first personal computer, the MITS Altair 8800, possible. The first PC was announced to the general public on the cover of the January 1975 issue of Popular Electronics.

Many electronics engineers today specialize in the development of programs for microprocessor based electronic systems, known as embedded systems. Due to the detailed knowledge of the hardware that is required for doing this, it is normally done by electronics engineers and not software engineers. Software engineers typically know and use microprocessors only at a conceptual level. Electronics engineers who exclusively carry out the role of programming embedded systems or microprocessors are referred to as "embedded systems engineers", or "firmware engineers".

## *Electronics*

In the field of electronic engineering, engineers design and test circuits that use the electromagnetic properties of electrical components such as resistors, capacitors, inductors, diodes and transistors to achieve a particular functionality. The tuner circuit, which allows the user of a radio to filter out all but a single station, is just one example of such a circuit.

In designing an integrated circuit, electronics engineers first construct circuit schematics that specify the electrical components and describe the interconnections between them. When completed, VLSI engineers convert the schematics into actual layouts, which map the layers of various conductor and semiconductor materials needed to construct the circuit. The conversion from schematics to layouts can be done by software but very often requires human fine-tuning to decrease space and power consumption. Once the layout is complete, it can be sent to a fabrication plant for manufacturing.

Integrated circuits and other electrical components can then be assembled on printed circuit boards to form more complicated circuits. Today, printed circuit boards are found in most electronic devices including televisions, computers and audio players.

## *Typical electronic engineering undergraduate syllabus*

Apart from electromagnetics and network theory, other items in the syllabus are particular to *electronics* engineering course. *Electrical* engineering courses have other specialisms such as machines, power generation and distribution. Note that the following list does not include the extensive engineering mathematics curriculum that is a prerequisite to a degree.

### Electromagnetics

Elements of vector calculus: divergence and curl; Gauss' and Stokes' theorems, Maxwell's equations: differential and integral forms. Wave equation, Poynting vector. Plane waves: propagation through various media; reflection and refraction; phase and group velocity; skin depth. Transmission lines: characteristic impedance; impedance transformation; Smith chart; impedance matching; pulse excitation. Waveguides: modes in rectangular waveguides; boundary conditions; cut-off frequencies; dispersion relations. Antennas: Dipole antennas; antenna arrays; radiation pattern; reciprocity theorem, antenna gain.

### Network analysis

Network graphs: matrices associated with graphs; incidence, fundamental cut set and fundamental circuit matrices. Solution methods: nodal and mesh analysis. Network theorems: superposition, Thevenin and Norton's maximum power transfer, Wye-Delta transformation. Steady state sinusoidal analysis using phasors. Linear constant coefficient differential equations; time domain analysis of simple RLC circuits, Solution of network equations using Laplace transform: frequency domain analysis of RLC circuits. 2-port network parameters: driving point and transfer functions. State equations for networks.

### Electronic devices and circuits

**Electronic devices**: Energy bands in silicon, intrinsic and extrinsic silicon. Carrier transport in silicon: diffusion current, drift current, mobility, resistivity. Generation and recombination of carriers. p-n junction diode, Zener diode, tunnel diode, BJT, JFET, MOS capacitor, MOSFET, LED, p-i-n and avalanche photo diode, LASERs. Device technology: integrated circuit fabrication process, oxidation, diffusion, ion implantation, photolithography, n-tub, p-tub and twin-tub CMOS process.

**Analog circuits**: Equivalent circuits (large and small-signal) of diodes, BJTs, JFETs, and MOSFETs. Simple diode circuits, clipping, clamping, rectifier. Biasing and bias stability of transistor and FET amplifiers. Amplifiers: single-and multi-stage, differential, operational, feedback and power. Analysis of amplifiers; frequency response of amplifiers. Simple op-amp circuits. Filters. Sinusoidal oscillators; criterion for oscillation; single-transistor and op-amp configurations. Function generators and wave-shaping circuits, Power supplies.

**Digital circuits**: of Boolean functions; logic gates digital IC families (DTL, TTL, ECL, MOS, CMOS). Combinational circuits: arithmetic circuits, code converters, multiplexers and decoders. Sequential circuits: latches and flip-flops, counters and shift-registers. Sample and hold circuits, ADCs, DACs. Semiconductor memories. Microprocessor 8086: architecture, programming, memory and I/O interfacing.

## Signals and systems

Definitions and properties of Laplace transform, continuous-time and discrete-time Fourier series, continuous-time and discrete-time Fourier Transform, z-transform. Sampling theorems. Linear Time-Invariant (LTI) Systems: definitions and properties; causality, stability, impulse response, convolution, poles and zeros frequency response, group delay, phase delay. Signal transmission through LTI systems. Random signals and noise: probability, random variables, probability density function, autocorrelation, power spectral density, function analogy between vectors & functions.

## Control systems

Basic control system components; block diagrammatic description, reduction of block diagrams — Mason's rule. Open loop and closed loop (negative unity feedback) systems and stability analysis of these systems. Signal flow graphs and their use in determining transfer functions of systems; transient and steady state analysis of LTI control systems and frequency response. Analysis of steady-state disturbance rejection and noise sensitivity.

Tools and techniques for LTI control system analysis and design: root loci, Routh-Hurwitz stability criterion, Bode and Nyquist plots. Control system compensators: elements of lead and lag compensation, elements of Proportional-Integral-Derivative controller (PID). Discretization of continuous time systems using Zero-order hold (ZOH) and ADCs for digital controller implementation. Limitations of digital controllers: aliasing. State variable representation and solution of state equation of LTI control systems. Linearization of Nonlinear dynamical systems with state-space realizations in both frequency and time domains. Fundamental concepts of controllability and observability for MIMO LTI systems. State space realizations: observable and controllable canonical form. Ackermann's formula for state-feedback pole placement. Design of full order and reduced order estimators.

## Communications

**Analog communication systems:** amplitude and angle modulation and demodulation systems, spectral analysis of these operations, superheterodyne noise conditions.

**Digital communication systems:** pulse code modulation (PCM), Differential Pulse Code Modulation (DPCM), Delta modulation (DM), digital modulation schemes-amplitude, phase and frequency shift keying schemes (ASK, PSK, FSK), matched filter receivers, bandwidth consideration and probability of error calculations for these schemes, GSM, TDMA.

## *Education and training*

Electronics engineers typically possess an academic degree with a major in electronic engineering. The length of study for such a degree is usually three or four years and the completed degree may be designated as a Bachelor of Engineering, Bachelor of Science, Bachelor of Applied Science, or Bachelor of Technology depending upon the university. Many UK universities also offer Master of Engineering (MEng) degrees at undergraduate level.

The degree generally includes units covering physics, chemistry, mathematics, project management and specific topics in electrical engineering. Initially such topics cover most, if not all, of the subfields of electronic engineering. Students then choose to specialize in one or more subfields towards the end of the degree.

Some electronics engineers also choose to pursue a postgraduate degree such as a Master of Science (MSc), Doctor of Philosophy in Engineering (PhD), or an Engineering Doctorate (EngD). The Master degree is being introduced in some European and American Universities as a first degree and the differentiation of an engineer with graduate and postgraduate studies is often difficult. In these cases, experience is taken into account. The Master's degree may consist of either research, coursework or a mixture of the two. The Doctor of Philosophy consists of a significant research component and is often viewed as the entry point to academia.

In most countries, a Bachelor's degree in engineering represents the first step towards certification and the degree program itself is certified by a professional body. After completing a certified degree program the engineer must satisfy a range of requirements (including work experience requirements) before being certified. Once certified the engineer is designated the title of Professional Engineer (in the United States, Canada and South Africa), Chartered Engineer or Incorporated Engineer (in the United Kingdom, Ireland, India and Zimbabwe), Chartered Professional Engineer (in Australia) or European Engineer (in much of the European Union).

Fundamental to the discipline are the sciences of physics and mathematics as these help to obtain both a qualitative and quantitative description of how such systems will work. Today most engineering work involves the use of computers and it is commonplace to use computer-aided design and simulation software programs when designing electronic systems. Although most electronic engineers will understand basic circuit theory, the theories employed by engineers generally depend upon the work they do. For example, quantum mechanics and solid state physics might be relevant to an engineer working on VLSI but are largely irrelevant to engineers working with macroscopic electrical systems.

## *Professional bodies*

Professional bodies of note for electrical engineers include the Institute of Electrical and Electronics Engineers (IEEE) and the Institution of Electrical Engineers (IEE) (now renamed the Institution of Engineering and Technology or IET). The IEEE claims to produce 30 percent of the world's literature in electrical/electronic engineering, has over 370,000 members, and holds more than 450 IEEE sponsored or cosponsored conferences worldwide each year.

## *Subfields*

Electronic engineering has many subfields. Here we, describes some of the most popular subfields in electronic engineering; although there are engineers who focus exclusively on one subfield, there are also many who focus on a combination of subfields.

### Overview of electronic engineering

**Electronic engineering** involves the design and testing of electronic circuits that use the electronic properties of components such as resistors, capacitors, inductors, diodes and transistors to achieve a particular functionality.

**Signal processing** deals with the analysis and manipulation of signals. Signals can be either analog, in which case the signal varies continuously according to the information, or digital, in which case the signal varies according to a series of discrete values representing the information.

For analog signals, signal processing may involve the amplification and filtering of audio signals for audio equipment or the modulation and demodulation of signals for telecommunications. For digital signals, signal processing may involve the compression, error checking and error detection of digital signals.

**Telecommunications engineering** deals with the transmission of information across a channel such as a co-axial cable, optical fiber or free space.

Transmissions across free space require information to be encoded in a carrier wave in order to shift the information to a carrier frequency suitable for transmission, this is known as modulation. Popular analog modulation techniques include amplitude modulation and frequency modulation. The choice of modulation affects the cost and performance of a system and these two factors must be balanced carefully by the engineer.

Once the transmission characteristics of a system are determined, telecommunication engineers design the transmitters and receivers needed for such systems. These two are sometimes combined to form a two-way communication device known as a transceiver. A key consideration in the design of transmitters is their power consumption as this is closely related to their signal strength. If the signal strength of a transmitter is insufficient the signal's information will be corrupted by noise.

**Control engineering** has a wide range of applications from the flight and propulsion systems of commercial airplanes to the cruise control present in many modern cars. It also plays an important role in industrial automation.

Control engineers often utilize feedback when designing control systems. For example, in a car with cruise control the vehicle's speed is continuously monitored and fed back to the system which adjusts the engine's power output accordingly. Where there is regular feedback, control theory can be used to determine how the system responds to such feedback.

**Instrumentation engineering** deals with the design of devices to measure physical quantities such as pressure, flow and temperature. These devices are known as instrumentation.

The design of such instrumentation requires a good understanding of physics that often extends beyond electromagnetic theory. For example, radar guns use the Doppler effect to measure the speed of oncoming vehicles. Similarly, thermocouples use the Peltier-Seebeck effect to measure the temperature difference between two points.

Often instrumentation is not used by itself, but instead as the sensors of larger electrical systems. For example, a thermocouple might be used to help ensure a furnace's temperature remains constant. For this reason, instrumentation engineering is often viewed as the counterpart of control engineering.

**Computer engineering** deals with the design of computers and computer systems. This may involve the design of new hardware, the design of PDAs or the use of computers to control an industrial plant. Computer engineers may also work on a system's software. However, the design of complex software systems is often the domain of software engineering, which is usually considered a separate discipline.

Desktop computers represent a tiny fraction of the devices a computer engineer might work on, as computer-like architectures are now found in a range of devices including video game consoles and DVD players.

## Project engineering

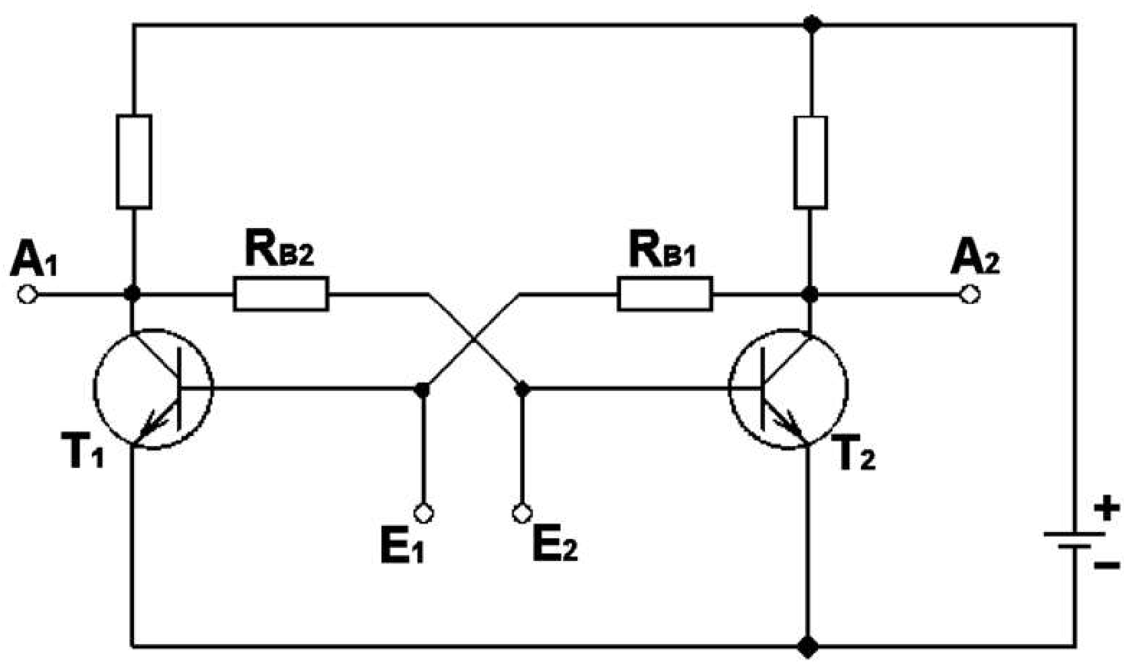
For most engineers not involved at the cutting edge of system design and development, technical work accounts for only a fraction of the work they do. A lot of time is also spent on tasks such as discussing proposals with clients, preparing budgets and determining project schedules. Many senior engineers manage a team of technicians or other engineers and for this reason project management skills are important. Most engineering projects involve some form of documentation and strong written communication skills are therefore very important.

The workplaces of electronics engineers are just as varied as the types of work they do. Electronics engineers may be found in the pristine laboratory environment of a fabrication plant, the offices of a consulting firm or in a research laboratory. During their working life, electronics engineers may find themselves supervising a wide range of individuals including scientists, electricians, computer programmers and other engineers.

Obsolescence of technical skills is a serious concern for electronics engineers. Membership and participation in technical societies, regular reviews of periodicals in the field and a habit of continued learning are therefore essential to maintaining proficiency. And these are mostly used in the field of consumer electronics products.

## Chapter-9

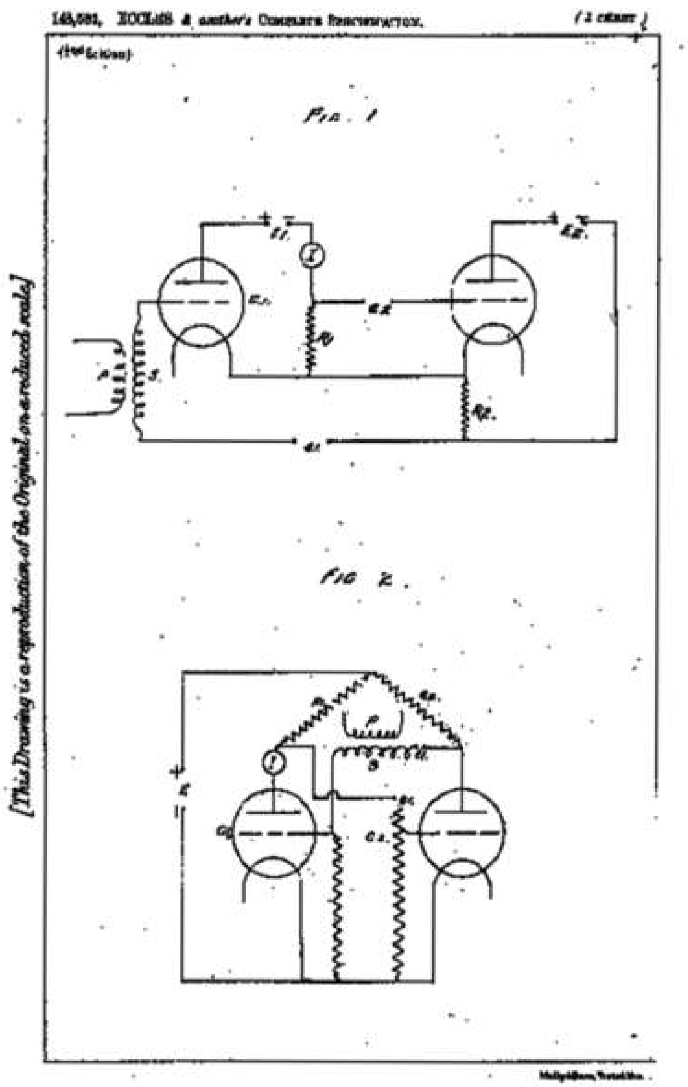
# Flip-flop




A traditional flip-flop circuit based on bipolar junction transistors

In electronics, a **flip-flop** is a circuit that has two stable states and can be used to store state information. The circuit can be made to change state by signals applied to one or more control inputs and will have one or two outputs. A circuit incorporating flip-flops has the attribute of *state*; its output depends not only on its current input, but also on its previous inputs. Such a circuit is described as sequential logic. Where a single input is provided, the circuit changes state every time a pulse appears on the input signal. Since the flip-flop retains the state after the signal pulses are removed, one type of flip-flop circuit is also called a "latch". Other types of flip-flops may have inputs that set a particular state, set the opposite state, or change states, depending on which input is pulsed.

Flip-flops are used as data storage elements, for counting of pulses, and for synchronizing randomly-timed input signals to some reference timing signal. Flip-flops are a fundamental building block of digital electronics systems used in computers, communications, and many other types of systems.



Flip-flop schematics from the Eccles and Jordan patent filed 1918, one drawn as a cascade of amplifiers with a positive feedback path, and the other as a symmetric cross-coupled pair

## History

The first electronic flip-flop was invented in 1918 by William Eccles and F. W. Jordan. It was initially called the *Eccles–Jordan trigger circuit* and consisted of two active elements (vacuum tubes). Such circuits and their transistorized versions were common in computers even after the introduction of integrated circuits, though flip-flops made from logic gates are also common now.

Early flip-flops were known variously as trigger circuits or multivibrators. A multivibrator is a two-state circuit; they come in several varieties, based on whether each state is stable or not: an *astable multivibrator* is not stable in either state, so it acts as a relaxation oscillator; a *monostable multivibrator* makes a pulse while in the unstable state, then returns to the stable state, and is known as a *one-shot*; a *bistable multivibrator* has two stable states, and this is the one usually known as a flip-flop. However, this terminology has been somewhat variable, historically. For example:

- 1942 – multivibrator implies astable: "The multivibrator circuit (Fig. 7-6) is somewhat similar to the flip-flop circuit, but the coupling from the anode of one valve to the grid of the other is by a condenser only, so that the coupling is not maintained in the steady state."
- 1942 – multivibrator as a particular flip-flop circuit: "Such circuits were known as 'trigger' or 'flip-flop' circuits and were of very great importance. The earliest and best known of these circuits was the multivibrator."
- 1943 – flip-flop as one-shot pulse generator: "It should be noted that an essential difference between the two-valve flip-flop and the multivibrator is that the flip-flop has one of the valves biased to cutoff."
- 1949 – monostable as flip-flop: "Monostable multivibrators have also been called 'flip-flops'."
- 1949 – monostable as flip-flop: "... a flip-flop is a monostable multivibrator and the ordinary multivibrator is an astable multivibrator."

According to P. L. Lindley, a JPL engineer, the flip-flop types discussed below (RS, D, T, JK) were first discussed in a 1954 UCLA course on computer design by Montgomery Phister, and then appeared in his book *Logical Design of Digital Computers.* Lindley was at the time working at Hughes Aircraft under Dr. Eldred Nelson, who had coined the term JK for a flip-flop which changed states when both inputs were on. The other names were coined by Phister. They differ slightly from some of the definitions given below. Lindley explains that he heard the story of the JK flip-flop from Dr. Eldred Nelson, who is responsible for coining the term while working at Hughes Aircraft. Flip-flops in use at Hughes at the time were all of the type that came to be known as J-K. In designing a logical system, Dr. Nelson assigned letters to flip-flop inputs as follows: #1: A & B, #2: C & D, #3: E & F, #4: G & H, #5: J & K.

## Implementation

Flip-flops can be either simple (transparent) or clocked; the transparent ones are commonly called latches.

The word *latch* is mainly used for storage elements, while clocked devices are described as **flip-flops**.

Simple flip-flops can be built around a pair of cross-coupled inverting elements: vacuum tubes, bipolar transistors, field effect transistors, inverters, and inverting logic gates have all been used in practical circuits. Clocked devices are specially designed for synchronous systems; such devices ignore their inputs except at the transition of a dedicated clock signal (known as clocking, pulsing, or strobing). Clocking causes the flip-flop to either change or retain its output signal based upon the values of the input signals at the transition. Some flip-flops change output on the rising edge of the clock, others on the falling edge.

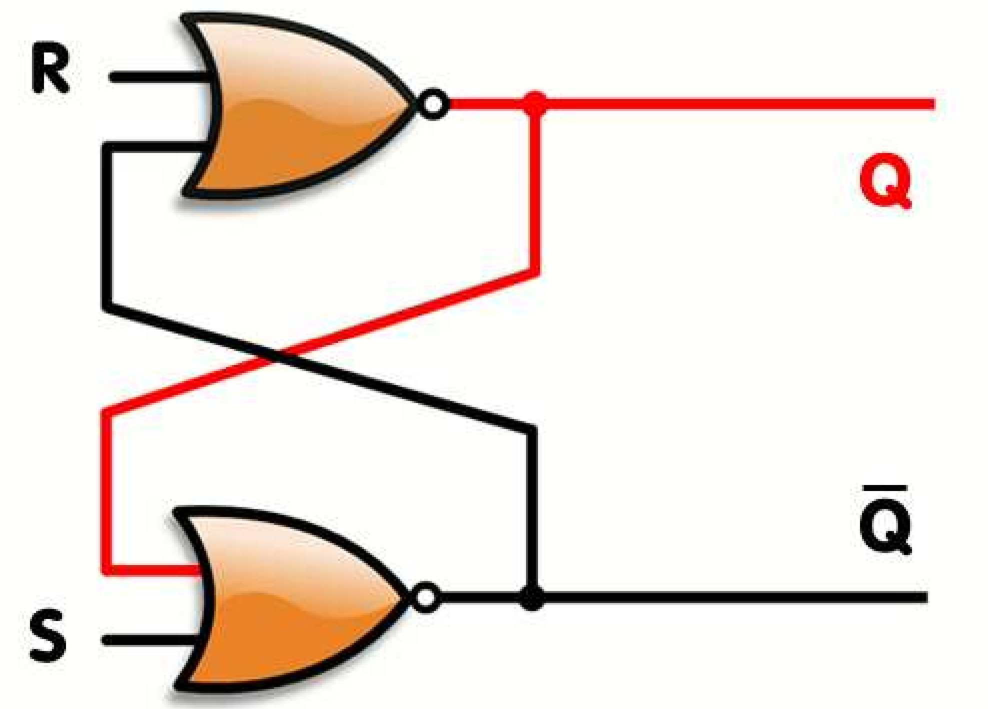
Since the elementary amplifying stages are inverting, two stages can be connected in succession (as a cascade) to form the needed non-inverting amplifier. In this configuration, each amplifier may be considered as an active inverting feedback network for the other inverting amplifier. Thus the two stages are connected in a non-inverting loop although the circuit diagram is usually drawn as a symmetric cross-coupled pair (both the drawings are initially introduced in the Eccles–Jordan patent).

## Flip-flop types

Flip-flops can be divided into common types: the **RS** ("set-reset"), **D** ("data" or "delay"), **T** ("toggle"), and **JK** types are the common ones. The behavior of a particular type can be described by what is termed the characteristic equation, which derives the "next" (i.e., after the next clock pulse) output, $Q_{next}$, in terms of the input signal(s) and/or the current output, $Q$.

## Simple set-reset latches

### SR NOR latch



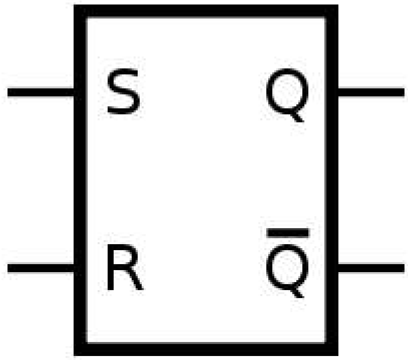

An RS latch, constructed from a pair of cross-coupled NOR gates. Red and black mean logical '1' and '0', respectively.

When using static gates as building blocks, the most fundamental latch is the simple *SR latch*, where S and R stand for *set* and *reset*. It can be constructed from a pair of cross-coupled NOR logic gates. The stored bit is present on the output marked Q.

While the S and R inputs are both low, feedback maintains the Q and Q outputs in a constant state, with Q the complement of Q. If S (*Set*) is pulsed high while R (*Reset*) is held low, then the Q output is forced high, and stays high when S returns to low; similarly, if R is pulsed high while S is held low, then the Q output is forced low, and stays low when R returns to low.

**SR latch operation**

| S R | Action |
|---|---|
| 0 0 | No Change |
| 0 1 | Q = 0 |
| 1 0 | Q = 1 |
| 1 1 | Restricted combination |




The symbol for an SR NOR latch

The R = S = 1 combination is called a **restricted combination** or a **forbidden state** because, as both NOR gates then output zeros, it breaks the logical equation Q = **not** Q. The combination is also inappropriate in circuits where *both* inputs may go low *simultaneously* (i.e. a transition from *restricted* to *keep*). The output would lock at either 1 or 0 depending on the propagation time relations between the gates (a race condition). In certain implementations, it could also lead to longer ringings (damped oscillations) before the output settles, and thereby result in undetermined values (errors) in high-frequency digital circuits. Although this condition is usually avoided, it can be useful in some applications.

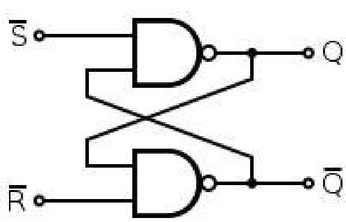
To overcome the restricted combination, one can add gates to the inputs that would convert `(S,R) = (1,1)` to one of the non-restricted combinations. That can be:

- Q = 1 (1,0) – referred to as an *S-latch*
- Q = 0 (0,1) – referred to as an *R-latch*
- Keep state (0,0) – referred to as an *E-latch*

Alternatively, the restricted combination can be made to *toggle* the output. The result is the JK latch.

Characteristic: Q+ = R'Q + R'S or Q+ = R'Q + S.

## SR NAND latch

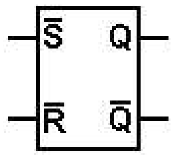



SR latch

This is an alternate model of the simple SR latch built with NAND (not AND) logic gates. *Set* and *reset* now become active low signals, denoted S and R respectively. Otherwise, operation is identical to that of the SR latch. Historically, SR-latches have been predominant despite the notational inconvenience of active-low inputs. This is because NAND gates are cheaper to produce than NOR gates in the diode-transistor logic (DTL), transistor-transistor logic (TTL) families, and complementary metal–oxide semiconductor (CMOS) logic families.

**SR latch operation**

| S | R | Action |
|---|---|---|
| 0 | 0 | Restricted combination |
| 0 | 1 | Q = 1 |
| 1 | 0 | Q = 0 |
| 1 | 1 | No Change |




Symbol for an SR NAND latch

## JK latch

The JK latch is much less used than the JK flip-flop. The JK latch follows the following state table:

**JK latch truth table**

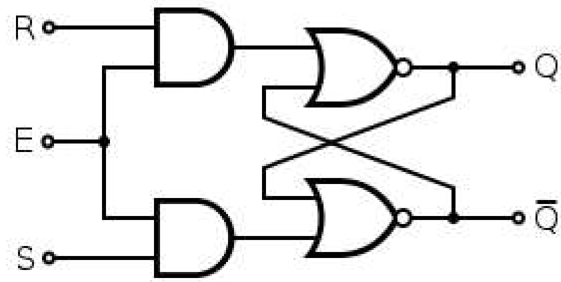
| J | K | $Q_{next}$ | Comment |
|---|---|---|---|
| 0 | 0 | Q | No change |
| 0 | 1 | 0 | Reset |
| 1 | 0 | 1 | Set |
| 1 | 1 | Q | Toggle |

Hence, the JK latch is an SR latch that is made to *toggle* its output when passed the restricted combination of 11. Unlike the JK Flip-Flop, in the JK latch, this is not a useful state because the speed of the toggling is not directed by a clock.

## Gated latches and conditional transparency

Latches are designed to be *transparent.* That is, input signal changes cause immediate changes in output; when several *transparent* latches follow each other, using the same clock signal, signals can propagate through all of them at once. Alternatively, additional logic can be added to a simple transparent latch to make it *non-transparent* or *opaque* when another input (an "enable" input) is not asserted. By following a *transparent-high* latch with a *transparent-low* (or *opaque-high*) latch, a master–slave flip-flop is implemented.

## Gated SR latch




A gated SR latch circuit diagram constructed from NOR gates.

A *synchronous SR latch* (sometimes *clocked SR flip-flop*) can be made by adding a second level of NAND gates to the inverted SR latch (or a second level of AND gates to the direct SR latch). The extra gates further invert the inputs so the simple SR latch becomes a gated SR latch (and a simple SR latch would transform into a gated SR latch with inverted enable).

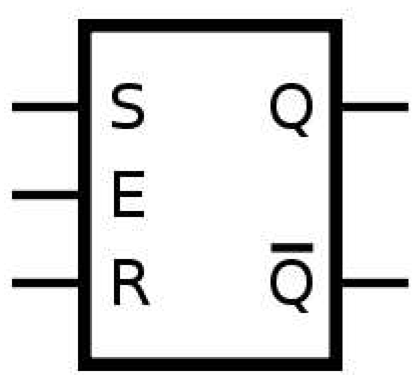
With E high (*enable* true), the signals can pass through the input gates to the encapsulated latch; all signal combinations except for (0,0) = *hold* then immediately reproduce on the (Q,Q) output, i.e. the latch is *transparent*.

With E low (*enable* false) the latch is *closed (opaque)* and remains in the state it was left the last time E was high.

The *enable* input is sometimes a clock signal, but more often a read or write strobe.

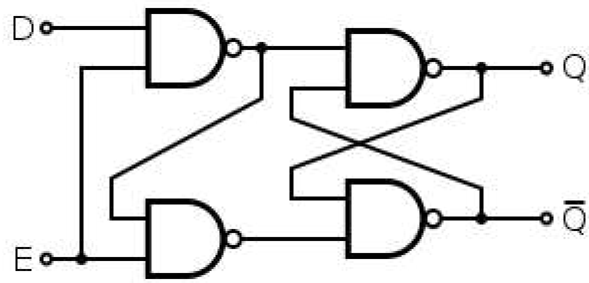
**Gated SR latch operation**

| E/C | Action |
|---|---|
| 0 | No action (keep state) |
| 1 | The same as non-clocked SR latch |



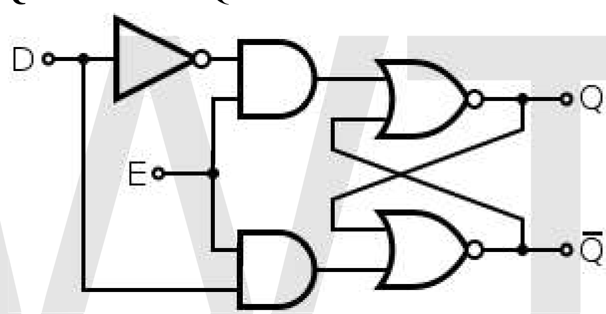

Symbol for a gated SR latch

## Gated D latch




A D-type transparent latch based on SR NAND latch

**D** Input
**E** Enable/clock
**Q** Output
**Q** Inverse of Q




A gated D latch based on SR NOR latch

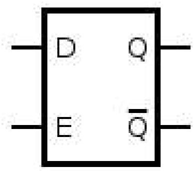
This latch exploits the fact that in the two active input combinations (01 and 10) of a gated SR latch R is the complement of S. The input NAND stage converts the two D input states (0 and 1) to these two input combinations for the next SR latch by inverting the data input signal. The low state of the *enable* signal produces the inactive "11" combination. Thus a gated D-latch may be considered as a *one-input synchronous SR latch*. This configuration prevents from applying the restricted combination to the inputs. It is also known as *transparent latch*, *data latch*, or simply *gated latch*. It has a *data* input and an *enable* signal (sometimes named *clock*, or *control*). The word *transparent* comes from the fact that, when the enable input is on, the signal propagates directly through the circuit, from the input D to the output Q.

Transparent latches are typically used as I/O ports or in asynchronous systems, or in synchronous two-phase systems (synchronous systems that use a two-phase clock), where two latches operating on different clock phases prevent data transparency as in a master–slave flip-flop.

Latches are available as integrated circuits, usually with multiple latches per chip. For example, 74HC75 is a quadruple transparent latch in the 7400 series.

Gated D latch truth table

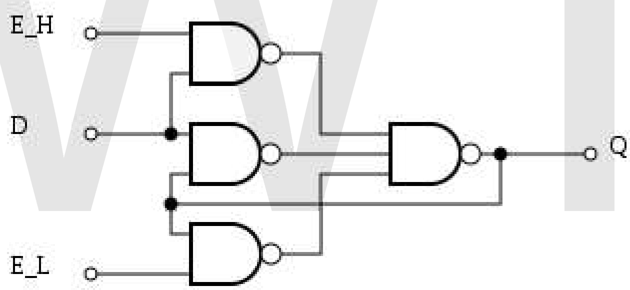
| **E/C** | **D** | **Q** | **Q** | **Comment** |
|---|---|---|---|---|
| 0 | X | $Q_{prev}$ | $Q_{prev}$ | No change |
| 1 | 0 | 0 | 1 | Reset |
| 1 | 1 | 1 | 0 | Set |




Symbol for a gated D latch

The truth table shows that when the *e*nable/*c*lock input is 0, the D input has no effect on the output. When E/C is high, the output equals D.

## Earle latch



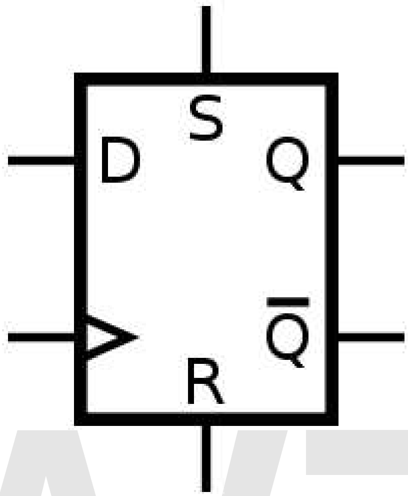

Earle latch uses complementary Enable inputs: Enable active Low (E_L) and Enable active H (E_H)

The classic gated latch designs have some undesirable characteristics. They require double-rail logic or an inverter. The input-to-output propagation may take up to three gate delays. The input-to-output propagation is not constant – some outputs take two gate delays while others take three.

Designers looked for alternatives. A successful alternative is the Earle latch. It requires only a single data input, and its output takes a constant two gate delays. In addition, the two gate levels of the Earle latch can be merged with the last two gate levels of the circuits driving the latch. Merging the latch function can implement the latch with no additional gate delays.

The Earle latch is hazard free. If the middle NAND gate is omitted, then one gets the **polarity hold latch**, which is commonly used because it demands less logic. Intentionally skewing the clock signal can avoid the hazard.

### D flip-flop




D flip-flop symbol

The D flip-flop is the most common flip-flop in use today. It is better known as *data* or *delay* flip-flop (as its output Q looks like a delay of input D).

The Q output takes on the state of the D input at the moment of a positive edge at the clock pin (or negative edge if the clock input is active low). It is called the D flip-flop for this reason, since the output takes the value of the D input or *data* input, and *delays* it by one clock cycle. The D flip-flop can be interpreted as a primitive memory cell, zero-order hold, or delay line. Whenever the clock pulses, the value of $Q_{next}$ is D and $Q_{prev}$ otherwise.

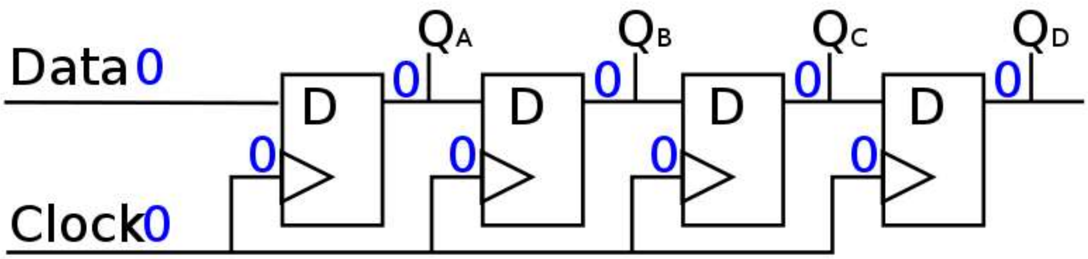
Truth table:

| **Clock** | **D** | **Q** | **$Q_{prev}$** |
|---|---|---|---|
| Rising edge | 0 | 0 | X |
| Rising edge | 1 | 1 | X |
| Non-Rising | X | $Q_{prev}$ | |

('X' denotes a *Don't care* condition, meaning the signal is irrelevant)

Most D-type flip-flops in ICs have the capability to be forced to the set or reset state (which ignores the D and clock inputs), much like an SR flip-flop. Usually, the illegal S = R = 1 condition is resolved in D-type flip-flops. By setting S = R = 0, the flip-flop can be used as described above.

| Inputs | | | | Outputs | |
|---|---|---|---|---|---|
| **S** | **R** | **D** | **>** | **Q** | **Q'** |
| 0 | 1 | X | X | 0 | 1 |
| 1 | 0 | X | X | 1 | 0 |
| 1 | 1 | X | X | 1 | 1 |

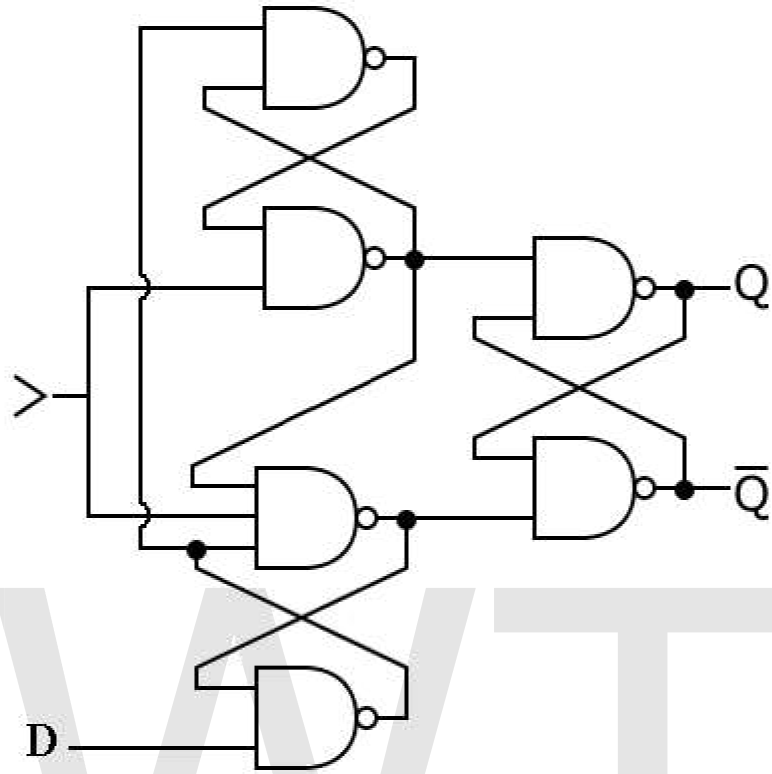



4-bit serial-in, serial-out (SISO) shift register

These flip-flops are very useful, as they form the basis for shift registers, which are an essential part of many electronic devices. The advantage of the D flip-flop over the D-type "transparent latch" is that the signal on the D input pin is captured the moment the flip-flop is clocked, and subsequent changes on the D input will be ignored until the next clock event. An exception is that some flip-flops have a "reset" signal input, which will reset Q (to zero), and may be either asynchronous or synchronous with the clock.

The above circuit shifts the contents of the register to the right, one bit position on each active transition of the clock. The input X is shifted into the leftmost bit position.

## Classical positive-edge-triggered D flip-flop




A positive-edge-triggered D flip-flop

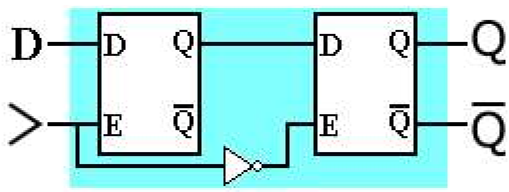
This clever circuit consists of two stages implemented by SR NAND latches. The input stage (the two latches on the left) processes the clock and data signals to ensure correct input signals for the output stage (the single latch on the right). If the clock is low, both the output signals of the input stage are high regardless of the data input; the output latch is unaffected and it stores the previous state. When the clock signal changes from low to high, only one of the output voltages (depending on the data signal) goes low and sets/resets the output latch: if D = 0, the lower output becomes low; if D = 1, the upper output becomes low. If the clock signal continues staying high, the outputs keep their states regardless of the data input and force the output latch to stay in the corresponding state as the input logical zero remains active while the clock is high. Hence the role of the output latch is to store the data only while the clock is low.

The circuit is closely related to the gated D latch as both the circuits convert the two D input states (0 and 1) to two input combinations (01 and 10) for the output SR latch by inverting the data input signal (both the circuits split the single D signal in two complementary S and R signals). The difference is that in the gated D latch simple NAND logical gates are used while in the positive-edge-triggered D flip-flop SR NAND latches are used for this purpose. The role of these latches is to "lock" the active output producing low voltage (a logical zero); thus the positive-edge-triggered D flip-flop can be thought as of a gated D latch with latched input gates.

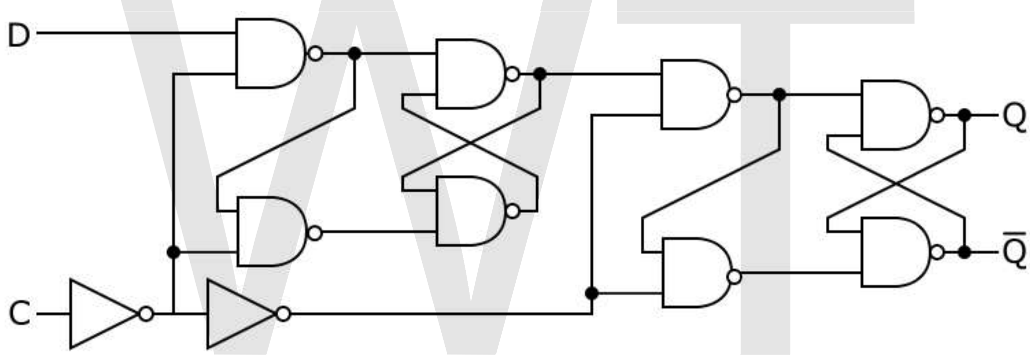
## Master–slave pulse-triggered D flip-flop

A master–slave D flip-flop is created by connecting two gated D latches in series, and inverting the *enable* input to one of them. It is called master–slave because the second latch in the series only changes in response to a change in the first (master) latch.

The term *pulse-triggered* means that data is entered on the rising edge of the clock pulse, but the output does not reflect the change until the falling edge of the clock pulse.




A master–slave D flip-flop. It responds on the negative edge of the *enable* input (usually a clock)

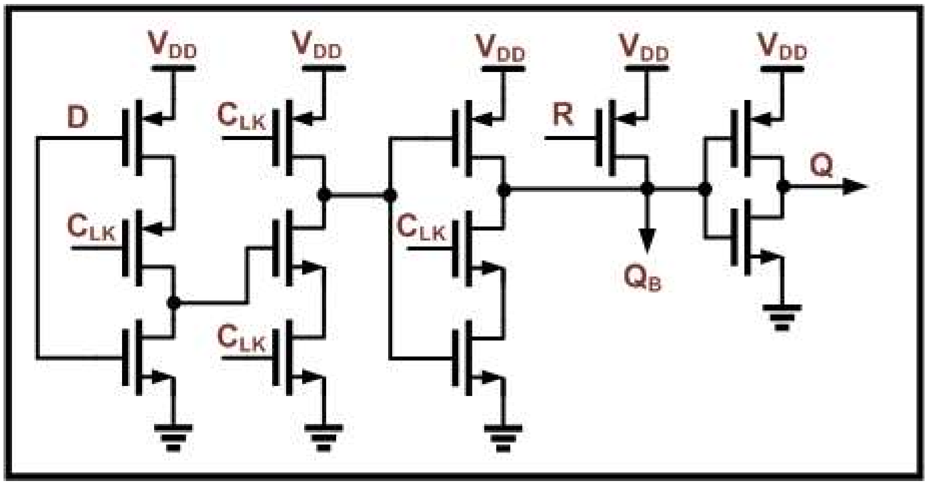



An implementation of a master–slave D flip-flop that is triggered on the positive edge of the clock

For a positive-edge triggered master–slave D flip-flop, when the clock signal is low (logical 0) the "enable" seen by the first or "master" D latch (the inverted clock signal) is high (logical 1). This allows the "master" latch to store the input value when the clock signal transitions from low to high. As the clock signal goes high (0 to 1) the inverted "enable" of the first latch goes low (1 to 0) and the value seen at the input to the master latch is "locked". Nearly simultaneously, the twice inverted "enable" of the second or "slave" D latch transitions from low to high (0 to 1) with the clock signal. This allows the signal captured at the rising edge of the clock by the now "locked" master latch to pass through the "slave" latch. When the clock signal returns to low (1 to 0), the output of the "slave" latch is "locked", and the value seen at the last rising edge of the clock is held while the "master" latch begins to accept new values in preparation for the next rising clock edge.

By removing the leftmost inverter in the circuit at side, a D-type flip flop that strobes on the *falling edge* of a clock signal can be obtained. This has a truth table like this:

| D | Q | > | $Q_{next}$ |
|---|---|---|---|
| 0 | X | Falling | 0 |
| 1 | X | Falling | 1 |



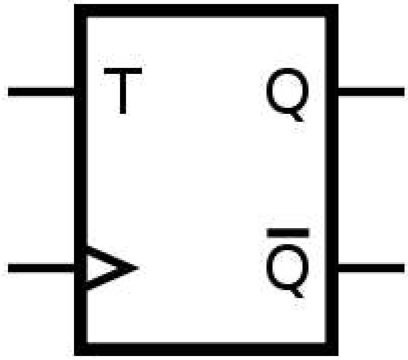

A CMOS IC implementation of a True Single Phase edge-triggered flip-flop with reset

## Edge-triggered dynamic D flip-flop

A more efficient way to make a D flip-flop is not so easy to understand, but it works the same way. While the master–slave D flip-flop is also triggered on the edge of a clock, its components are each triggered by clock levels. The "edge-triggered D flip-flop" does not have the master–slave properties.

Edge-triggered D flip-flops are often implemented in integrated high-speed operations using dynamic logic. This means that the digital output is stored on parasitic device capacitance while the device is not transitioning. This design of dynamic flip flops also enable simple resetting since the reset operation can be performed by simply discharging one or more internal nodes. A common dynamic flip-flop variety is the True Single Phase Clock (TSPC) which performs the flip flop operation with little power and at high speeds. However these types of dynamic flip-flops will not work at static or low clocked speeds, since given enough time the parasitic capacitance will discharge through leakage paths and will cause the logic levels to enter invalid states.

### T flip-flop




A circuit symbol for a T-type flip-flop

If the T input is high, the T flip-flop changes state ("toggles") whenever the clock input is strobed. If the T input is low, the flip-flop holds the previous value. This behavior is described by the characteristic equation:

$$Q_{next} = T \oplus Q = T\overline{Q} + \overline{T}Q$$ (expanding the XOR operator)

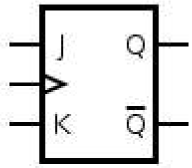
and can be described in a truth table:

**T flip-flop operation**

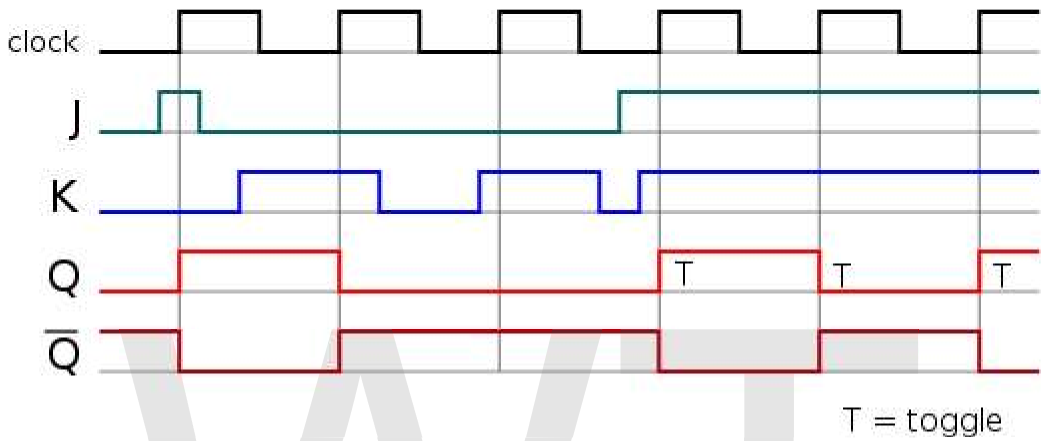
| **Characteristic table** | | | | **Excitation table** | | | |
|---|---|---|---|---|---|---|---|
| ***T*** | ***Q*** | $Q_{next}$ | **Comment** | ***Q*** | $Q_{next}$ | ***T*** | **Comment** |
| 0 | 0 | 0 | hold state (no clk) | 0 | 0 | 0 | No change |
| 0 | 1 | 1 | hold state (no clk) | 1 | 1 | 0 | No change |
| 1 | 0 | 1 | toggle | 0 | 1 | 1 | Complement |
| 1 | 1 | 0 | toggle | 1 | 0 | 1 | Complement |

When T is held high, the toggle flip-flop divides the clock frequency by two; that is, if clock frequency is 4 MHz, the output frequency obtained from the flip-flop will be 2 MHz. This "divide by" feature has application in various types of digital counters. A T flip-flop can also be built using a JK flip-flop (J & K pins are connected together and act as T) or D flip-flop (T input and $Q_{previous}$ is connected to the D input through an XOR gate). A T flip-flop can also be built using an edge-triggered D flip-flop with its D input fed from its own inverted output.

## JK flip-flop




A circuit symbol for a positive-edge-triggered JK flip-flop




JK flip-flop timing diagram

The JK flip-flop augments the behavior of the SR flip-flop (J=Set, K=Reset) by interpreting the S = R = 1 condition as a "flip" or toggle command. Specifically, the combination J = 1, K = 0 is a command to set the flip-flop; the combination J = 0, K = 1 is a command to reset the flip-flop; and the combination J = K = 1 is a command to toggle the flip-flop, i.e., change its output to the logical complement of its current value. Setting J = K = 0 does NOT result in a D flip-flop, but rather, will hold the current state. To synthesize a D flip-flop, simply set K equal to the complement of J. The JK flip-flop is therefore a universal flip-flop, because it can be configured to work as an SR flip-flop, a D flip-flop, or a T flip-flop.

NOTE: The flip-flop is positive-edge triggered (rising clock pulse) as seen in the timing diagram.

The characteristic equation of the JK flip-flop is:

$$Q_{next} = J\overline{Q} + \overline{K}Q$$

and the corresponding truth table is:

**JK Flip Flop operation**

**Characteristic table** **Excitation table**

**J K $Q_{next}$ Comment Q $Q_{next}$ J K Comment**

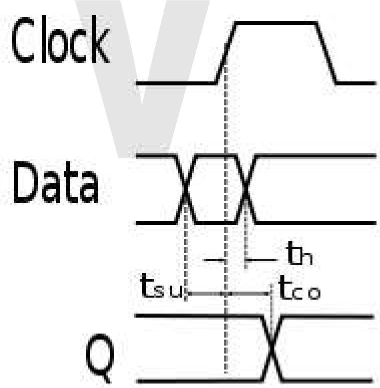
| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | Q | hold state | 0 | 0 | 0 | X | No change |
| 0 | 1 | 0 | reset | 0 | 1 | 1 | X | Set |
| 1 | 0 | 1 | set | 1 | 0 | X | 1 | Reset |
| 1 | 1 | Q | toggle | 1 | 1 | X | 0 | No change |

## *Metastability*

Flip-flops are prone to a problem called metastability, which can happen when two inputs, such as data and clock or clock and reset, are changing at about the same time, such that the resulting state would depend on the order of the input events. When the order is not clear, within appropriate timing constraints, the result is that the output may behave unpredictably, taking many times longer than normal to settle to one state or the other, or even oscillating several times before settling. Theoretically, the time to settle down is not bounded. In a computer system, this metastability can cause corruption of data or a program crash, if the state is not stable before another circuit uses its value; in particular, if two different logical paths use the output of a flip-flop, one path can interpret it as a 0 and the other as a 1 when it has not resolved to stable state, putting the machine into an inconsistent state.

## *Timing considerations*

### Setup and hold times




Flip-flop setup, hold and clock-to-output timing parameters

**Setup time** is the minimum amount of time the data signal should be held steady **before** the clock event so that the data are reliably sampled by the clock. This applies to synchronous circuits such as the flip-flop.

**Hold time** is the minimum amount of time the data signal should be held steady **after** the clock event so that the data are reliably sampled. This applies to synchronous circuits such as the flip-flop.

To summarize: Setup time -> Clock flank -> Hold time.

The metastability in flip-flops can be avoided by ensuring that the data and control inputs are held valid and constant for specified periods before and after the clock pulse, called the **setup time** ($t_{su}$) and the **hold time** ($t_h$) respectively. These times are specified in the data sheet for the device, and are typically between a few nanoseconds and a few hundred picoseconds for modern devices.

Unfortunately, it is not always possible to meet the setup and hold criteria, because the flip-flop may be connected to a real-time signal that could change at any time, outside the control of the designer. In this case, the best the designer can do is to reduce the probability of error to a certain level, depending on the required reliability of the circuit. One technique for suppressing metastability is to connect two or more flip-flops in a chain, so that the output of each one feeds the data input of the next, and all devices share a common clock. With this method, the probability of a metastable event can be reduced to a negligible value, but never to zero. The probability of metastability gets closer and closer to zero as the number of flip-flops connected in series is increased.

So-called metastable-hardened flip-flops are available, which work by reducing the setup and hold times as much as possible, but even these cannot eliminate the problem entirely. This is because metastability is more than simply a matter of circuit design. When the transitions in the clock and the data are close together in time, the flip-flop is forced to decide which event happened first. However fast we make the device, there is always the possibility that the input events will be so close together that it cannot detect which one happened first. It is therefore logically impossible to build a perfectly metastable-proof flip-flop.

### Propagation delay

Another important timing value for a flip-flop (F/F) is the clock-to-output delay (common symbol in data sheets: $t_{CO}$) or propagation delay ($t_P$), which is the time the flip-flop takes to change its output after the clock edge. The time for a high-to-low transition ($t_{PHL}$) is sometimes different from the time for a low-to-high transition ($t_{PLH}$).

When cascading F/Fs which share the same clock (as in a shift register), it is important to ensure that the $t_{CO}$ of a preceding F/F is longer than the hold time ($t_h$) of the following flip-flop, so data present at the input of the succeeding F/F is properly "shifted in" following the active edge of the clock. This relationship between $t_{CO}$ and $t_h$ is normally guaranteed if the F/Fs are physically identical. Furthermore, for correct operation, it is easy to verify that the clock period has to be greater than the sum $t_{su} + t_h$.

## *Generalizations*

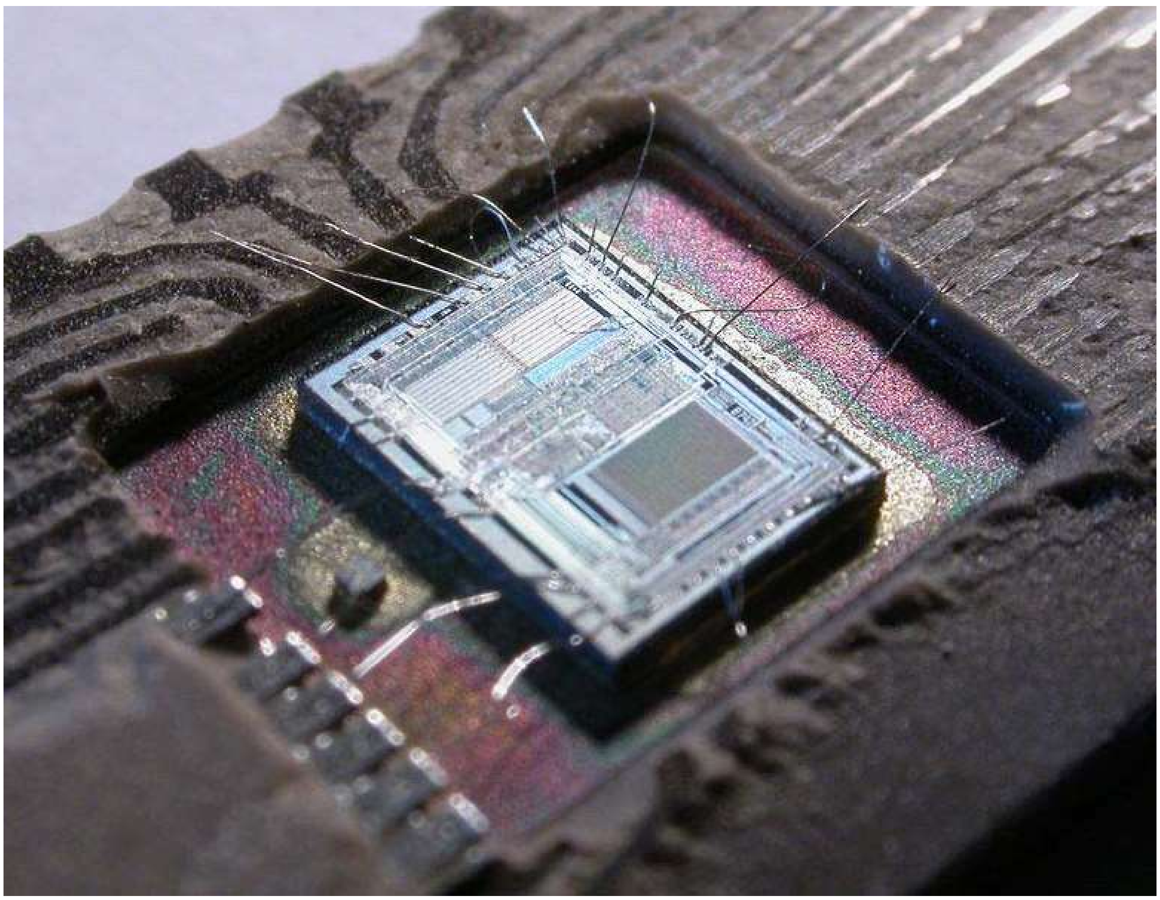
Flip-flops can be generalized in at least two ways: by making them 1-of-N instead of 1-of-2, and by adapting them to logic with more than two states. In the special cases of 1-of-3 encoding, or multi-valued ternary logic, these elements may be referred to as *flip-flap-flops*.

In a conventional flip-flop, exactly one of the two complementary outputs is high. This can be generalized to a memory element with N outputs, exactly one of which is high (alternatively, where exactly one of N is low). The output is therefore always a one-hot (respectively *one-cold*) representation. The construction is similar to a conventional cross-coupled flip-flop; each output, when high, inhibits all the other outputs. Alternatively, more or less conventional flip-flops can be used, one per output, with additional circuitry to make sure only one at a time can be true.

Another generalization of the conventional flip-flop is a memory element for multi-valued logic. In this case the memory element retains exactly one of the logic states until the control inputs induce a change. In addition, a multiple-valued clock can also be used, leading to new possible clock transitions.

Chapter-10

# Electronic Circuit



The die from an Intel 8742, an 8-bit microcontroller that includes a CPU, 128 bytes of RAM, 2048 bytes of EPROM, and I/O in the same chip.

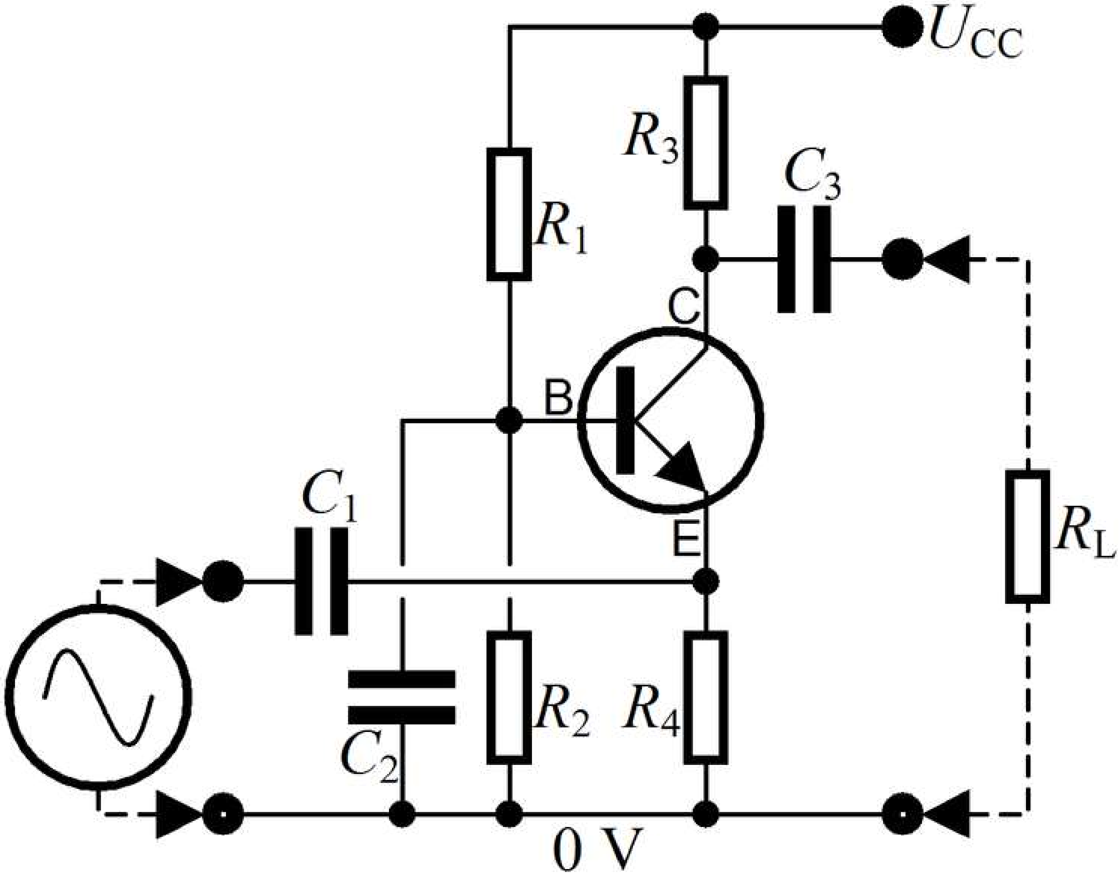

A circuit built on a printed circuit board (PCB).

An **electronic circuit** is composed of individual electronic components, such as resistors, transistors, capacitors, inductors and diodes, connected by conductive wires or traces through which electric current can flow. The combination of components and wires allows various simple and complex operations to be performed: signals can be amplified, computations can be performed, and data can be moved from one place to another. Circuits can be constructed of discrete components connected by individual pieces of wire, but today it is much more common to create interconnections by photolithographic techniques on a laminated substrate (a printed circuit board or PCB) and solder the components to these interconnections to create a finished circuit. In an Integrated Circuit or IC, the components and interconnections are formed on the same substrate, typically a semiconductor such as silicon or (less commonly) gallium arsenide.

Breadboards, perfboards or stripboards are common for testing new designs. They allow the designer to make quick changes to the circuit during development.

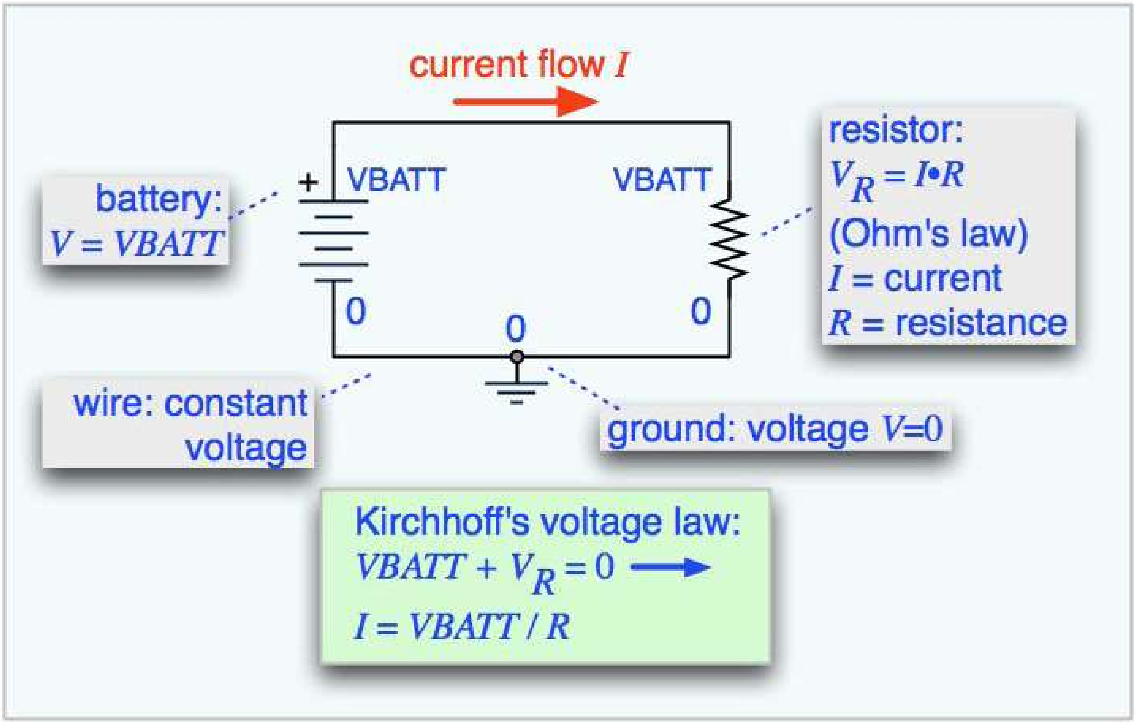
An electronic circuit can usually be categorized as an analog circuit, a digital circuit or a mixed-signal circuit (a combination of analog circuits and digital circuits).

### *Analog circuits*




A circuit diagram representing an analog circuit, in this case a simple amplifier.

Analog electronic circuits are those in which current or voltage may vary continuously with time to correspond to the information being represented. Analog circuitry is constructed from two fundamental building blocks: series and parallel circuits. In a series circuit, the same current passes through a series of components. A string of Christmas lights is a good example of a series circuit: if one goes out, they all do. In a parallel circuit, all the components are connected to the same voltage, and the current divides between the various components according to their resistance.




A simple schematic showing wires, a resistor, and a battery.

The basic components of analog circuits are wires, resistors, capacitors, inductors, diodes, and transistors. (Recently, memristors have been added to the list of available components.) Analog circuits are very commonly represented in schematic diagrams, in which wires are shown as lines, and each component has a unique symbol. Analog circuit analysis employs Kirchhoff's circuit laws: all the currents at a node (a place where wires meet) must add to 0, and the voltage around a closed loop of wires is 0. Wires are usually treated as ideal zero-voltage interconnections; any resistance or reactance is captured by explicitly adding a parasitic element, such as a discrete resistor or inductor. Active components such as transistors are often treated as controlled current or voltage sources: for example, a field-effect transistor can be modeled as a current source from the source to the drain, with the current controlled by the gate-source voltage.

When the circuit size is comparable to a wavelength of the relevant signal frequency, a more sophisticated approach must be used. Wires are treated as transmission lines, with (hopefully) constant characteristic impedance, and the impedances at the start and end determine transmitted and reflected waves on the line. Such considerations typically become important for circuit boards at frequencies above a GHz; integrated circuits are smaller and can be treated as lumped elements for frequencies less than 10 GHz or so.

An alternative model is to take independent power sources and induction as basic electronic units; this allows modeling frequency dependent negative resistors, gyrators, negative impedance converters, and dependent sources as secondary electronic components.

## Digital circuits

In digital electronic circuits, electric signals take on discrete values, to represent logical and numeric values. These values represent the information that is being processed. In the vast majority of cases, binary encoding is used: one voltage (typically the more positive value) represents a binary '1' and another voltage (usually a value near the ground potential, 0 V) represents a binary '0'. Digital circuits make extensive use of transistors, interconnected to create logic gates that provide the functions of Boolean logic: AND, OR, NOT, and all possible combinations there of. Transistors interconnected so as to provide positive feedback are used as latches and flip flops, circuits that have two or more metastable states, and remain in one of these states until changed by an external input. Digital circuits therefore can provide both logic and memory, enabling them to perform arbitrary computational functions. (Memory based on flip-flops is known as SRAM (static random access memory). Memory based on the storage of charge in a capacitor, DRAM (dynamic random access memory) is also widely used.)

Digital circuits are fundamentally easier to design than analog circuits for the same level of complexity, because each logic gate regenerates the binary signal, so the designer need not account for distortion, gain control, offset voltages, and other concerns faced in an analog design. As a consequence, extremely complex digital circuits, with billions of logic elements integrated on a single silicon chip, can be fabricated at low cost. Such digital integrated circuits are ubiquitous in modern electronic devices, such as calculators, mobile phone handsets, and computers.

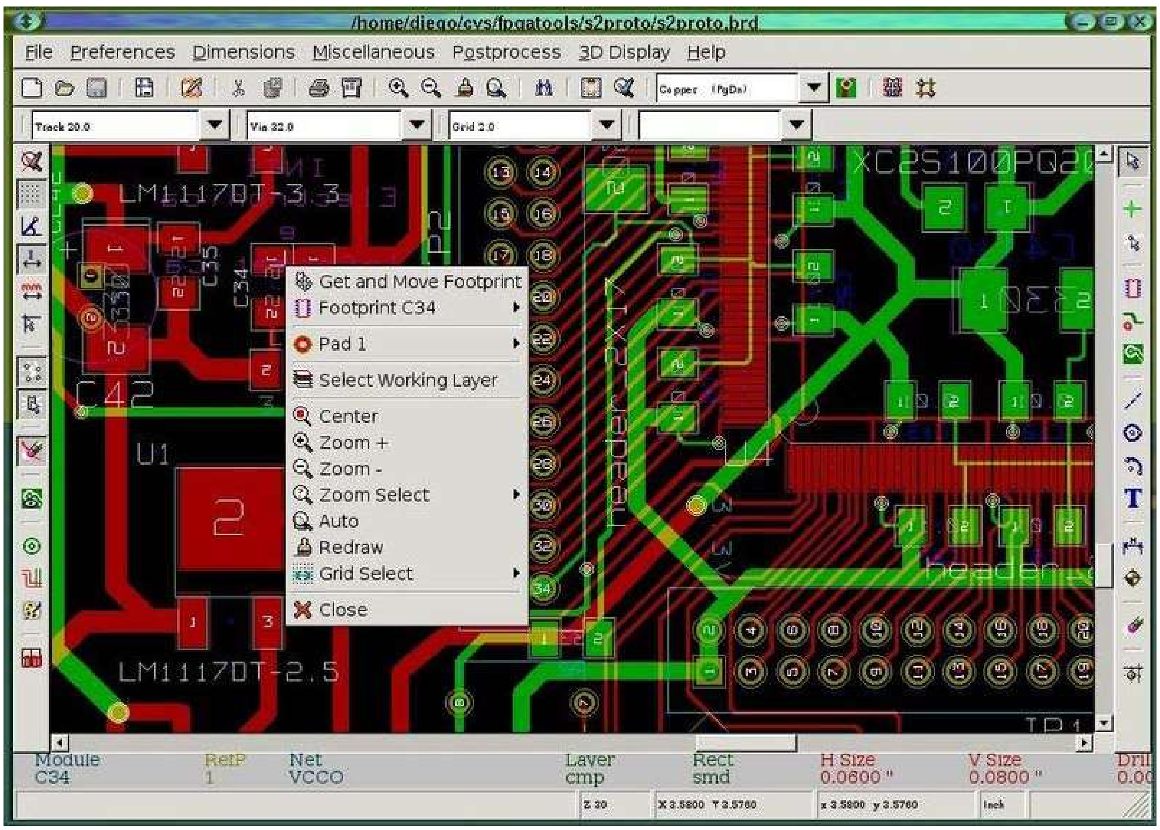
Digital circuitry is used to create general purpose computing chips, such as microprocessors, and custom-designed logic circuits, known as Application Specific Integrated Circuits (ASICs). Field Programmable Gate Arrays (FPGAs), chips with logic circuitry whose configuration can be modified after fabrication, are also widely used in prototyping and development.

## Mixed-signal circuits

Mixed-signal or hybrid circuits contain elements of both analog and digital circuits. Examples include comparators, timers, PLLs, ADCs (analog-to-digital converters), and DACs (digital-to-analog converters). Most modern radio and communications circuitry uses mixed signal circuits. For example, in a receiver, analog circuitry is used to amplify and frequency-convert signals so that they reach a suitable state to be converted into digital values, after which further signal processing can be performed in the digital domain.

**Chapter-11**

# Electronic Design Automation




PCB layout program

**Electronic design automation** (**EDA** or **ECAD**) is a category of software tools for designing electronic systems such as printed circuit boards and integrated circuits. The tools work together in a design flow that chip designers use to design and analyze entire semiconductor chips.

## *History*

### Early days

Before EDA, integrated circuits were designed by hand, and manually laid out. Some advanced shops used geometric software to generate the tapes for the Gerber photoplotter, but even those copied digital recordings of mechanically-drawn components. The process was fundamentally graphic, with the translation from electronics to graphics done manually. The best known company from this era was Calma, whose GDSII format survives.

By the mid-70s, developers started to automate the design, and not just the drafting. The first placement and routing (Place and route) tools were developed. The proceedings of the Design Automation Conference cover much of this era.

The next era began about the time of the publication of "Introduction to VLSI Systems" by Carver Mead and Lynn Conway in 1980. This ground breaking text advocated chip design with programming languages that compiled to silicon. The immediate result was a considerable increase in the complexity of the chips that could be designed, with improved access to design verification tools that used logic simulation. Often the chips were easier to lay out and more likely to function correctly, since their designs could be simulated more thoroughly prior to construction. Although the languages and tools have evolved, this general approach of specifying the desired behavior in a textual programming language and letting the tools derive the detailed physical design remains the basis of digital IC design today.

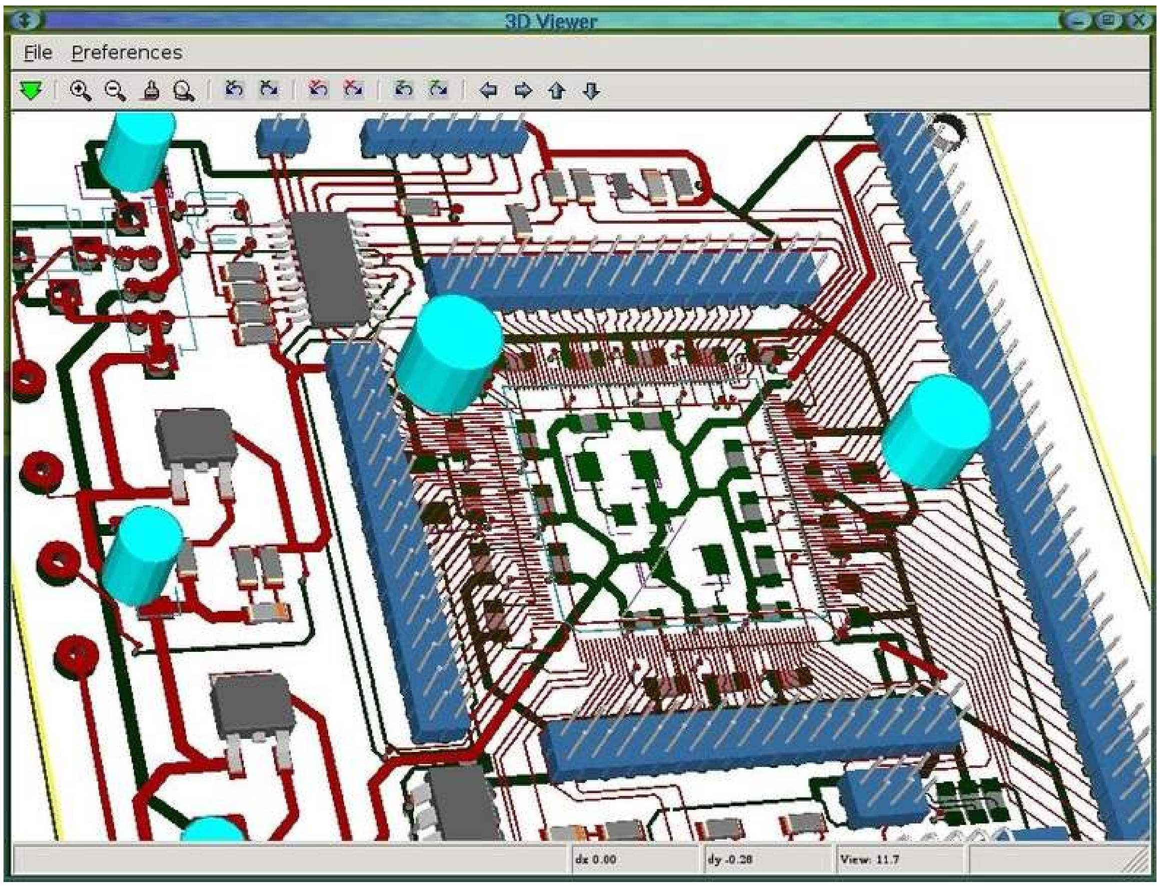
The earliest EDA tools were produced academically. One of the most famous was the "Berkeley VLSI Tools Tarball", a set of UNIX utilities used to design early VLSI systems. Still widely used is the Espresso heuristic logic minimizer and Magic.

Another crucial development was the formation of MOSIS, a consortium of universities and fabricators that developed an inexpensive way to train student chip designers by producing real integrated circuits. The basic concept was to use reliable, low-cost, relatively low-technology IC processes, and pack a large number of projects per wafer, with just a few copies of each projects' chips. Cooperating fabricators either donated the processed wafers, or sold them at cost, seeing the program as helpful to their own long-term growth.

### Birth of commercial EDA

1981 marks the beginning of EDA as an industry. For many years, the larger electronic companies, such as Hewlett Packard, Tektronix, and Intel, had pursued EDA internally. In 1981, managers and developers spun out of these companies to concentrate on EDA as a business. Daisy Systems, Mentor Graphics, and Valid Logic Systems were all founded around this time, and collectively referred to as **DMV**. Within a few years there were many companies specializing in EDA, each with a slightly different emphasis. The first trade show for EDA was held at the Design Automation Conference in 1984.

In 1986, Verilog, a popular high-level design language, was first introduced as a hardware description language by Gateway Design Automation. In 1987, the U.S. Department of Defense funded creation of VHDL as a specification language. Simulators quickly followed these introductions, permitting direct simulation of chip designs: executable specifications. In a few more years, back-ends were developed to perform logic synthesis.




3D PCB layout

## Current status

Current digital flows are extremely modular. The front ends produce standardized design descriptions that compile into invocations of "cells,", without regard to the cell technology. Cells implement logic or other electronic functions using a particular integrated circuit technology. Fabricators generally provide libraries of components for their production processes, with simulation models that fit standard simulation tools. Analog EDA tools are far less modular, since many more functions are required, they interact more strongly, and the components are (in general) less ideal.

EDA for electronics has rapidly increased in importance with the continuous scaling of semiconductor technology. Some users are foundry operators, who operate the semiconductor fabrication facilities, or "fabs", and design-service companies who use EDA software to evaluate an incoming design for manufacturing readiness. EDA tools are also used for programming design functionality into FPGAs.

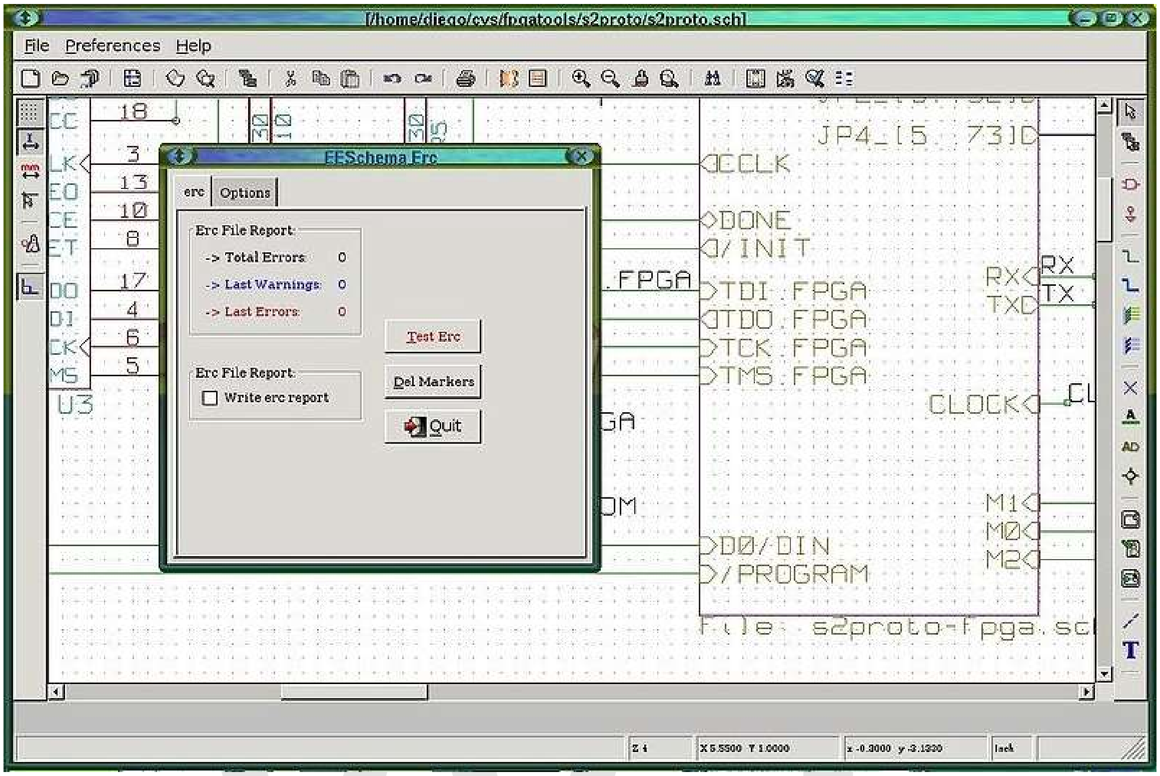
## *Software focuses*

### Design

- High-level synthesis(syn. behavioural synthesis, algorithmic synthesis) For digital chips
- Logic synthesis translation of abstract, logical language such as Verilog or VHDL into a discrete netlist of logic-gates
- Schematic Capture For standard cell digital, analog, rf like Capture CIS in Orcad by CADENCE and ISIS in Proteus
- Layout like Layout in Orcad by Cadence, ARES in Proteus

### Simulation

- Transistor simulation – low-level transistor-simulation of a schematic/layout's behavior, accurate at device-level.
- Logic simulation – digital-simulation of an RTL or gate-netlist's digital (boolean 0/1) behavior, accurate at boolean-level.
- **Behavioral Simulation** – high-level simulation of a design's architectural operation, accurate at cycle-level or interface-level.
- Hardware emulation – Use of special purpose hardware to emulate the logic of a proposed design. Can sometimes be plugged into a system in place of a yet-to-be-built chip; this is called **in-circuit emulation**.
- Technology CAD simulate and analyze the underlying process technology. Electrical properties of devices are derived directly from device physics.
- Electromagnetic field solvers, or just field solvers, solve Maxwell's equations directly for cases of interest in IC and PCB design. They are known for being slower but more accurate than the layout extraction above.




Schematic capture program

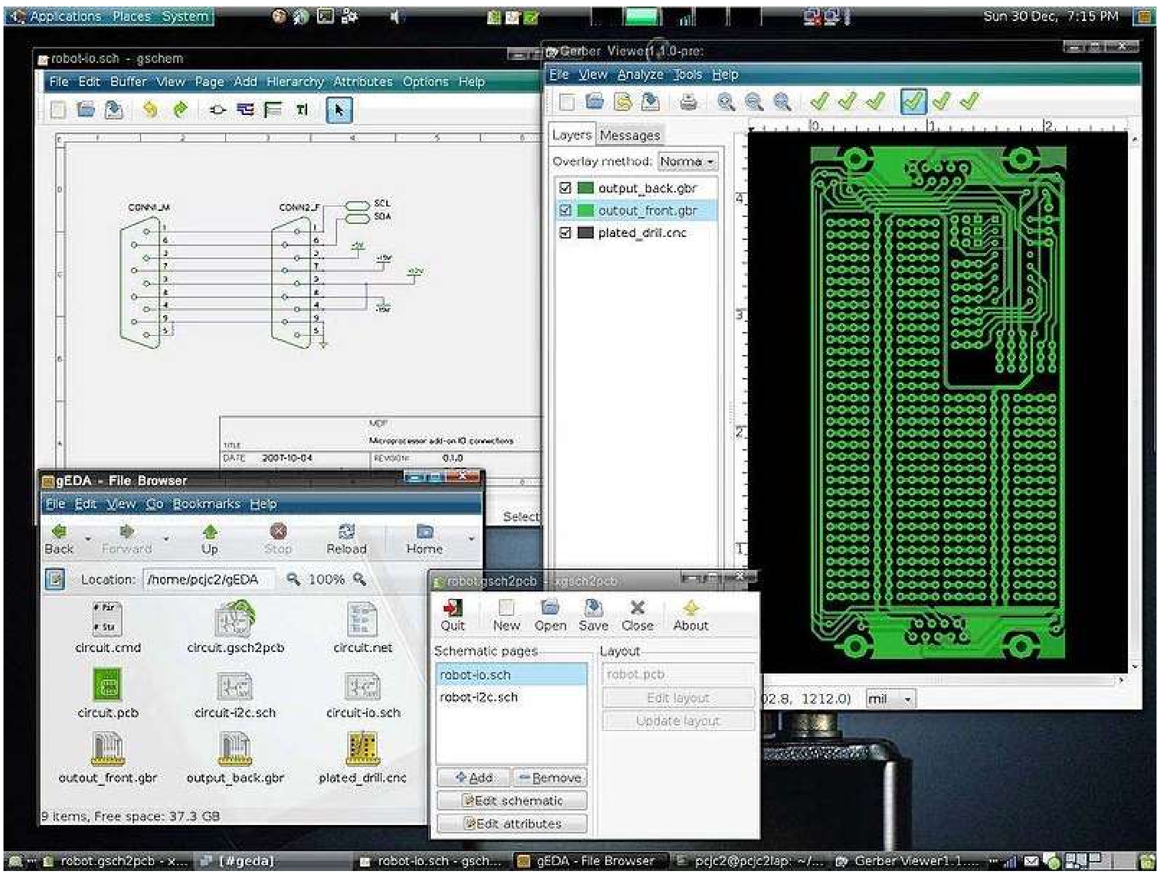
## Analysis and verification

- Functional verification
- Clock Domain Crossing Verification (CDC check): Similar to linting, but these checks/tools specialize in detecting and reporting potential issues like data loss, meta-stability due to use of multiple clock domains in the design.
- Formal verification, also model checking: Attempts to prove, by mathematical methods, that the system has certain desired properties, and that certain undesired effects (such as deadlock) cannot occur.
- Equivalence checking: algorithmic comparison between a chip's RTL-description and synthesized gate-netlist, to ensure functional equivalence at the *logical* level.
- Static timing analysis: Analysis of the timing of a circuit in an input-independent manner, hence finding a worst case over all possible inputs.
- Physical verification, PV: checking if a design is physically manufacturable, and that the resulting chips will not have any function-preventing physical defects, and will meet original specifications.

## Manufacturing preparation

- Mask data preparation, MDP: generation of actual lithography photomask used to physically manufacture the chip.

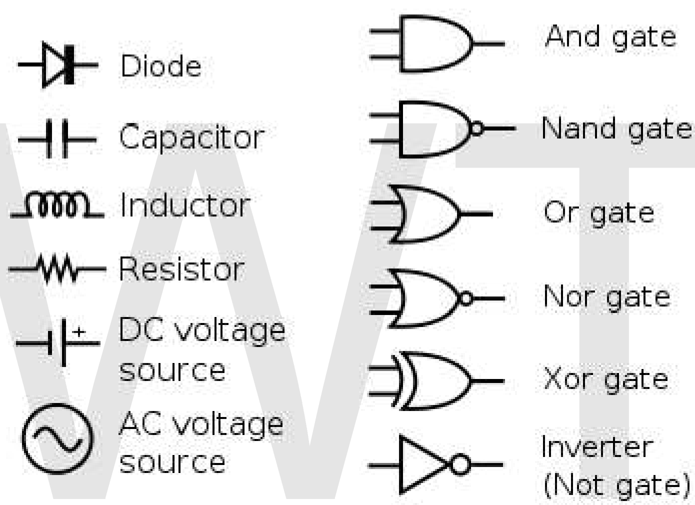
- Resolution enhancement techniques, RET – methods of increasing of quality of final photomask.
- Optical proximity correction, OPC – up-front compensation for diffraction and interference effects occurring later when chip is manufactured using this mask.
- Mask generation – generation of flat mask image from hierarchical design.
- Automatic test pattern generation, ATPG – generates pattern-data to systematically exercise as many logic-gates, and other components, as possible.
- Built-in self-test, or BIST – installs self-contained test-controllers to automatically test a logic (or memory) structure in the design




PCB layout and schematic for connector design

**Chapter-12**

# Electronic Symbol




Common circuit diagram symbols (US symbols)

An **electronic symbol** is a pictogram used to represent various electrical and electronic devices (such as wires, batteries, resistors, and transistors) in a schematic diagram of an electrical or electronic circuit. These symbols can (because of remaining traditions) vary from country to country, but are today to a large extent internationally standardized. Some symbols (such as those of vacuum tubes) became virtually extinct with the development of new technologies.

## *Standards for symbols*

There are several national and international standards for graphical symbols in circuit diagrams, in particular:

- IEC 60617 (also known as British Standard BS 3939)
- ANSI standard Y32 (also known as IEEE Std 315)
- Australian Standard AS 1102

Different symbols may be used depending on the discipline using the drawing. For example, lighting and power symbols used as part of architectural drawings may be different from symbols for devices used in electronics. National and local variations to international standards also exist.

## *Reference designations*

A **reference designator** unambiguously identifies a component in an electrical schematic (circuit diagram) or on a printed circuit board (PCB). The reference designator usually consists of one or two letters followed by a number, e.g. R13, C1002. The number is sometimes followed by a letter, indicating that components are grouped or matched with each other, e.g. R17A, R17B. IEEE 315 contains a list of Class Designation Letters to use for electrical and electronic assemblies. For example, the letter R is a reference prefix for the resistors of an assembly, C for capacitors, K for relays.

IEEE 200-1975 or "Standard Reference Designations for Electrical and Electronics Parts and Equipments" is a standard that was used to define referencing naming systems for collections of electronic equipment. IEEE 200 was ratified in 1975. The IEEE renewed the standard in the 1990s, but withdrew it from active support shortly thereafter. This standard also has an ANSI document number ANSI Y32.16-1975. They are the same document.

This standard codified information from, among other sources, a United States military standard MIL-STD-16 which dates back to at least the 1950s in American industry.

To replace IEEE 200-1975, a standards body for Mechanical Engineers, ASME, initiated the new standard ASME Y14.44-2008.

This standard along with IEEE 315-1975 provide the electrical designer with guidance on how to properly reference and annotate everything from a simple circuit board to a complete enclosure all the way to a collection of these assemblies.

It breaks down a system into units, and then any number of sub-assemblies. The Unit is the highest level of demarcation in a system and is always a numeral. Subsequent demarcation are called assemblies and always have the Class Letter "A" as a prefix following by a sequential number starting with 1. Any number of sub-assemblies may be defined until finally reaching the component.

Especially valuable is the method of referencing and annotating cables plus their connectors within and outside assemblies. Examples:

- 1A1A44J5 - Unit 1, Assembly 1, Sub-Assembly 44, Jack 5 (J5 is a connector on a box referenced as A44)
- 1A1A45J333 - Unit 1, Assembly 1, Sub-Assembly 45, Jack 333 (J333 is a connector on a box referenced as A45)

A cable connecting these two might be:

- 1A1W35 - In the assembly A1 is a cable called W35.

Connectors on this cable would be designated:

- 1A1W35P1
- 1A1W35P2

ASME Y14.44-2008 continues the convention of Plug P and Jack J when assigning references for connectors in electrical assemblies where a J (or jack) is the more fixed and P (or plug) is the less fixed of a connector pair without regard to the gender of the connector contacts.

The construction of reference designators is covered by IEEE 200-1975/ANSI Y32.16-1975 (replaced by ASME Y14.44-2008) and IEEE-315-1975. The table below lists designators commonly used, and does not comply with the standard.

| Designator | Component Type |
|---|---|
| AT | Attenuator |
| BR | Bridge rectifier |
| BT | Battery |
| C | Capacitor |
| CN | Capacitor network |
| D | Diode (including zeners, thyristors and LEDs) |
| DL | Delay line |
| DS | Display |
| F | Fuse |
| FB or FEB | Ferrite bead |
| FD | Fiducial |
| J | Jack connector (female) |
| JP | Link (Jumper) |
| K | Relay |
| L | Inductor |
| LS | Loudspeaker or buzzer |
| M | Motor |
| MK | Microphone |
| MP | Mechanical part (including screws and fasteners) |
| P | Plug connector (male) |
| PS | Power supply |
| Q | Transistor (all types) |

| | |
|---|---|
| R | Resistor |
| RN | Resistor network |
| RT | Thermistor |
| RV | Varistor |
| S | Switch (all types, including push-buttons) |
| T | Transformer |
| TC | Thermocouple |
| TUN | Tuner |
| TP | Test point |
| U | Integrated circuit |
| V | Vacuum Tube |
| VR | Variable Resistor (potentiometer or rheostat) |
| X | Transducer not matching any other category |
| Y | Crystal or oscillator |
| Z | Zener Diode |

Component name abbreviations widely used in industry:

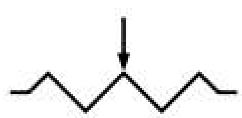
- AE: aerial, antenna
- B: battery
- BR: bridge rectifier
- C: capacitor
- CRT:cathode ray tube
- D or CR: diode
- DSP:digital signal processor
- F: fuse
- FET:field effect transistor
- GDT: gas discharge tube
- IC: integrated circuit
- J: wire link ("jumper")
- JFET: junction gate field-effect transistor
- L: inductor
- LCD:Liquid crystal display
- LDR: light dependent resistor
- LED: light emitting diode
- LS: speaker
- M: motor
- MCB: circuit breaker
- Mic: microphone
- MOSFET:Metal oxide semiconductor field effect transistor
- Ne: neon lamp
- OP: Operational Amplifier
- PCB: printed circuit board

- PU: pickup
- Q: transistor
- R: resistor
- RLA: RY: relay
- SCR: silicon controlled rectifier
- SW: switch
- T: transformer
- TFT:thin film transistor(display)
- TH: thermistor
- TP: test point
- Tr: transistor
- U: integrated circuit
- V: valve (tube)
- VC: variable capacitor
- VFD: vacuum fluorescent display
- VLSI:very large scale integration
- VR: variable resistor
- X: crystal, ceramic resonator
- XMER: transformer
- XTAL: crystal
- Z or ZD: Zener diode

## *Gallery of common electronic symbols*

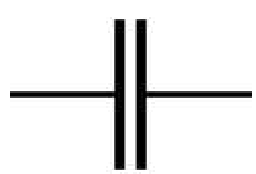
### Resistors

Resistor: American (top) and IEC (bottom)

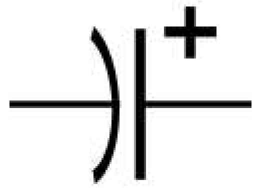


Potentiometer: American

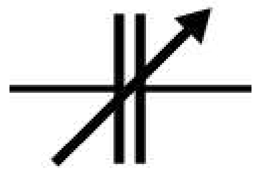
### Capacitors



Capacitor



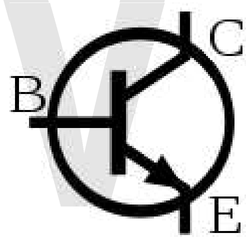
Capacitor, polarized (American)



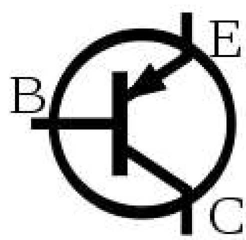
Capacitor, variable



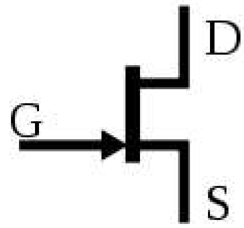
**Transistors**



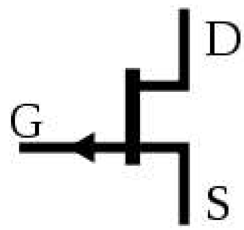

NPN transistor



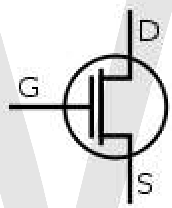

PNP transistor




n-channel junction gate field-effect transistor (JFET)

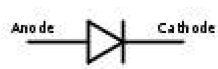



p-channel junction gate field-effect transistor (JFET)

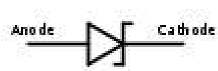



Field-effect transistor

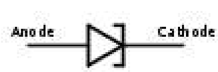
## Diodes



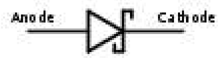

Diode



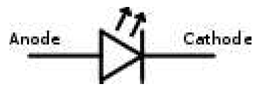

Zener diode



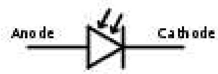

Tunnel diode




Schottky diode




Light Emitting Diode (LED)

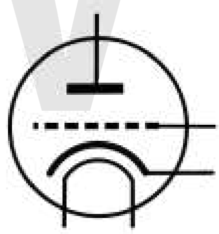



Photodiode

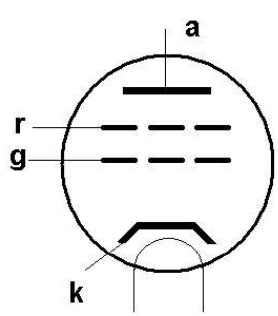

Varicap



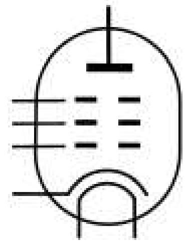
## Vacuum tubes



Vacuum tube triode



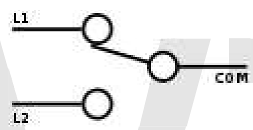

Vacuum tube tetrode



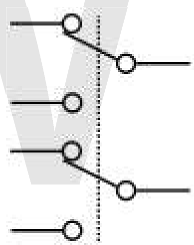
Vacuum tube pentode

## Switches

Switch, Single Pole/Single Throw (SPST)




Switch, Single Pole/Double Throw (SPDT)

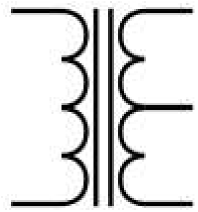


Switch, Double Pole/Double Throw (DPDT)

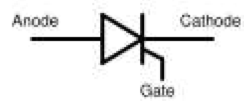
## Miscellaneous


Batteries, single cell, multi-cell

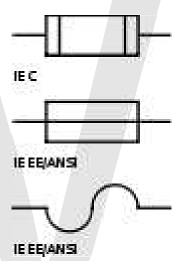
Inductor



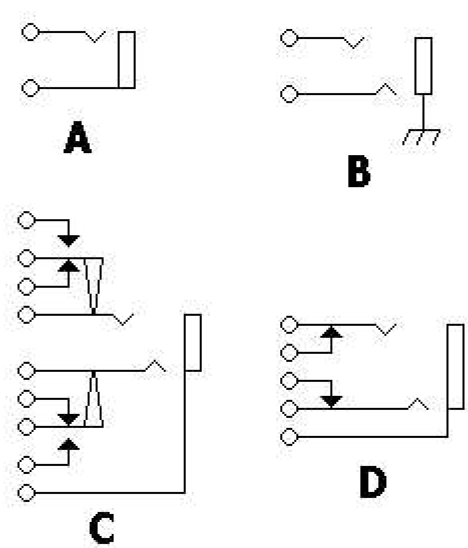
Transformer, with center tap




Silicon-controlled rectifier




Fuse: IEC (top) and American (lower two)




Phone jacks

**Chapter-13**

# Circuit Design and Current–voltage Characteristic

## Circuit design

The process of **circuit design** can cover systems ranging from complex electronic systems all the way down to the individual transistors within an integrated circuit. For simple circuits the design process can often be done by one person without needing a planned or structured design process, but for more complex designs, teams of designers following a systematic approach with intelligently guided computer simulation are becoming increasingly common.

In integrated circuit design automation, the term "circuit design" often refers the step of the design cycle which outputs the schematics of the integrated circuit. Typically this is the step between logic design and physical design.

Formal circuit design usually involves the following stages:

- sometimes, writing the requirement specification after liaising with the customer
- writing a technical proposal to meet the requirements of the customer specification
- synthesising on paper a schematic circuit diagram, an abstract electrical or electronic circuit that will meet the specifications
- calculating the component values to meet the operating specifications under specified conditions
- performing simulations to verify the correctness of the design
- building a breadboard or other prototype version of the design and testing against specification
- making any alterations to the circuit to achieve compliance
- choosing a method of construction as well as all the parts and materials to be used
- presenting component and layout information to draughtspersons, and layout and mechanical engineers, for prototype production
- testing or type-testing a number of prototypes to ensure compliance with customer requirements

- signing and approving the final manufacturing drawings
- post-design services (obsolescence of components etc.)

## *Specification*

The process of circuit design begins with the specification, which states the functionality that the finished design must provide, but does not indicate how it is to be achieved . The initial specification is basically a technically detailed description of what the customer wants the finished circuit to achieve and can include a variety of electrical requirements, such as what signals the circuit will receive, what signals it must output, what power supplies are available and how much power it is permitted to consume. The specification can (and normally does) also set some of the physical parameters that the design must meet, such as size, weight, moisture resistance, temperature range, thermal output, vibration tolerance and acceleration tolerance.

As the design process progresses the designer(s) will frequently return to the specification and alter it to take account of the progress of the design. This can involve tightening specifications that the customer has supplied, and adding tests that the circuit must pass in order to be accepted. These additional specifications will often be used in the verification of a design. Changes that conflict with or modify the customer's original specifications will almost always have to be approved by the customer before they can be acted upon.

Correctly identifying the customer needs can avoid a condition known as 'design creep' which occurs in the absence of realistic initial expectations, and later by failing to communicate fully with the client during the design process. It can be defined in terms of its results; "at one extreme is a circuit with more functionality than necessary, and at the other is a circuit having an incorrect functionality". (DeMers, 1997) Nevertheless some changes can be expected and it is good practice to keep options open for as long as possible because it's easier to remove spare elements from the circuit later on than it is to put them in.

## *Design*

The design process involves moving from the specification at the start, to a plan that contains all the information needed to be physically constructed at the end, this normally happens by passing through a number of stages, although in very simple circuit it may be done in a single step.  The process normally begins with the conversion of the specification into a block diagram of the various functions that the circuit must perform, at this stage the contents of each block are not considered, only what each block must do, this is sometimes referred to as a "black box" design. This approach allows the possibly very complicated task to be broken into smaller tasks which may either by tackled in sequence or divided amongst members of a design team.

Each block is then considered in more detail, still at an abstract stage, but with a lot more focus on the details of the electrical functions to be provided. At this or later stages it is common to require a large amount of research or mathematical modeling into what is and is not feasible to achieve. The results of this research may be fed back into earlier stages of the design process, for example if it turns out one of the blocks cannot be designed within the parameters set for it, it may be necessary to alter other blocks instead. At this point it is also common to start considering both how to demonstrate that the design does meet the specifications, and how it is to be tested ( which can include self diagnostic tools ).

Finally the individual circuit components are chosen to carry out each function in the overall design, at this stage the physical layout and electrical connections of each component are also decided, this layout commonly taking the form of artwork for the production of a printed circuit board or Integrated circuit. This stage is typically extremely time consuming because of the vast array of choices available. A practical constraint on the design at this stage is that of standardization, while a certain value of component may be calculated for use in some location in a circuit, if that value cannot be purchased from a supplier, then the problem has still not been solved. To avoid this a certain amount of 'catalog engineering' can be applied to solve the more mundane tasks within an overall design.

### Costs

Proper design philosophy and structure incorporates economic and technical considerations and keeps them in balance at all times, and right from the start. Balance is the key concept here; just as many delays and pitfalls can come from ill considered cost cutting as with cost overruns. Good accounting tools (and a design culture that fosters their use) is imperative for a successful project. "Manufacturing costs shrink as design costs soar," is often quoted as a truism in circuit design, particularly for ICs.

## *Verification and testing*

Once a circuit has been designed, it must be both verified and tested. Verification is the process of going through each stage of a design and ensuring that it will do what the specification requires it to do. This is frequently a highly mathematical process and can involve large-scale computer simulations of the design. In any complicated design it is very likely that problems will be found at this stage and may involve a large amount of the design work be redone in order to fix them.

Testing is the real-world counterpart to verification, testing involves physically building at least a prototype of the design and then (in combination with the test procedures in the specification or added to it) checking the circuit really does do what it was designed to.

## *Prototyping*

Prototyping is a means of exploring ideas before an investment is made in them. Depending on the scope of the prototype and the level of detail required, prototypes can be built at any time during the project. Sometimes they are created early in the project, during the planning and specification phase, commonly using a process known as breadboarding; that's when the need for exploration is greatest, and when the time investment needed is most viable. Later in the cycle packaging mock-ups are used to explore appearance and usability, and occasionally a circuit will need to be modified to take these factors into account.

## *Results*

As circuit design is the process of working out the physical form that an electronic circuit will take, the result of the circuit design process is the instructions on how to construct the physical electronic circuit. This will normally take the form of blueprints describing the size, shape, connectors, etc., in use, and artwork or CAM file for manufacturing a printed circuit board or Integrated circuit.

### Documentation

Any commercial design will normally also include an element of documentation, the precise nature of this documentation will vary according to the size and complexity of the circuit as well as the country in which it is to be used. As a bare minimum the documentation will normally include at least the specification and testing procedures for the design and a statement of compliance with current regulations. In the EU this last item will normally take the form of a CE Declaration listing the European directives complied with and naming an individual responsible for compliance.

# Current–voltage characteristic

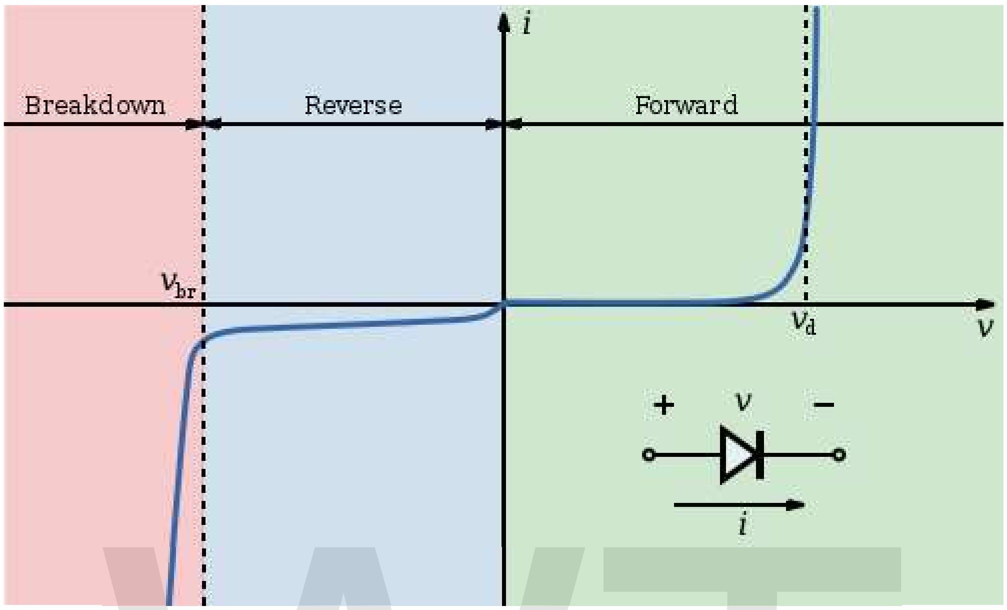



Figure 1. I–V curve of a P–N junction diode (not to scale).

A **current–voltage characteristic** is a relationship, typically represented as a chart or graph, between an electric current and a corresponding voltage, or potential difference.

### *In electronics*

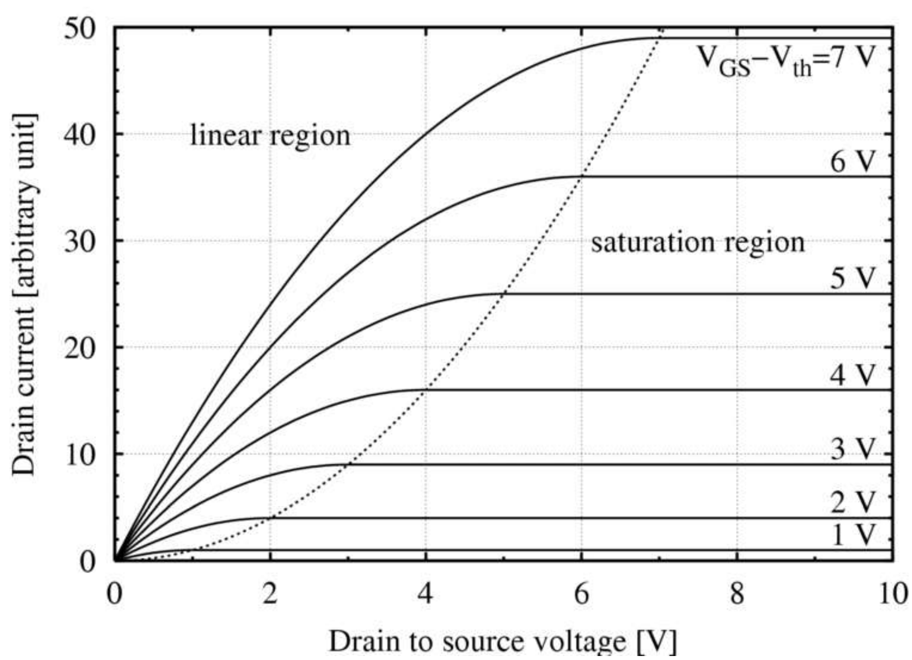


Figure 2. MOSFET drain current vs. drain-to-source voltage for several values of the *overdrive voltage*, $V_{GS} - V_{th}$; the boundary between **linear** (**Ohmic**) and **saturation** (**active**) modes is indicated by the upward curving parabola

In electronics, the relationship between the DC current through an electronic device and the DC voltage across its terminals is called a current–voltage characteristic of the device. Electronic engineers use these charts to determine basic parameters of a device and to model its behavior in an electrical circuit. These characteristics are also known as **I-V curves**, referring to the standard symbols for current and voltage.

A more general form of current–voltage characteristic is one that describes the dependence of a terminal current on more than one terminal voltage difference; electronic devices such as vacuum tubes and transistors are described by such characteristics.

Figure 1 shows an I–V curve for a P-N junction diode. Figure 2 shows a family of I–V curves for a MOSFET as a function of drain voltage with overvoltage ($V_{GS}$-$V_{th}$) as a parameter.

The simplest I–V characteristic involves a resistor, which according to Ohm's Law exhibits a linear relationship between the applied voltage and the resulting electric current. However, even in this case environmental factors such as temperature or material characteristics of the resistor can produce a non-linear curve.

The transconductance and Early voltage of a transistor are examples of parameters traditionally measured with the assistance of an I–V chart, or laboratory equipment that traces the charts in real time on an oscilloscope.

### *In electrophysiology*

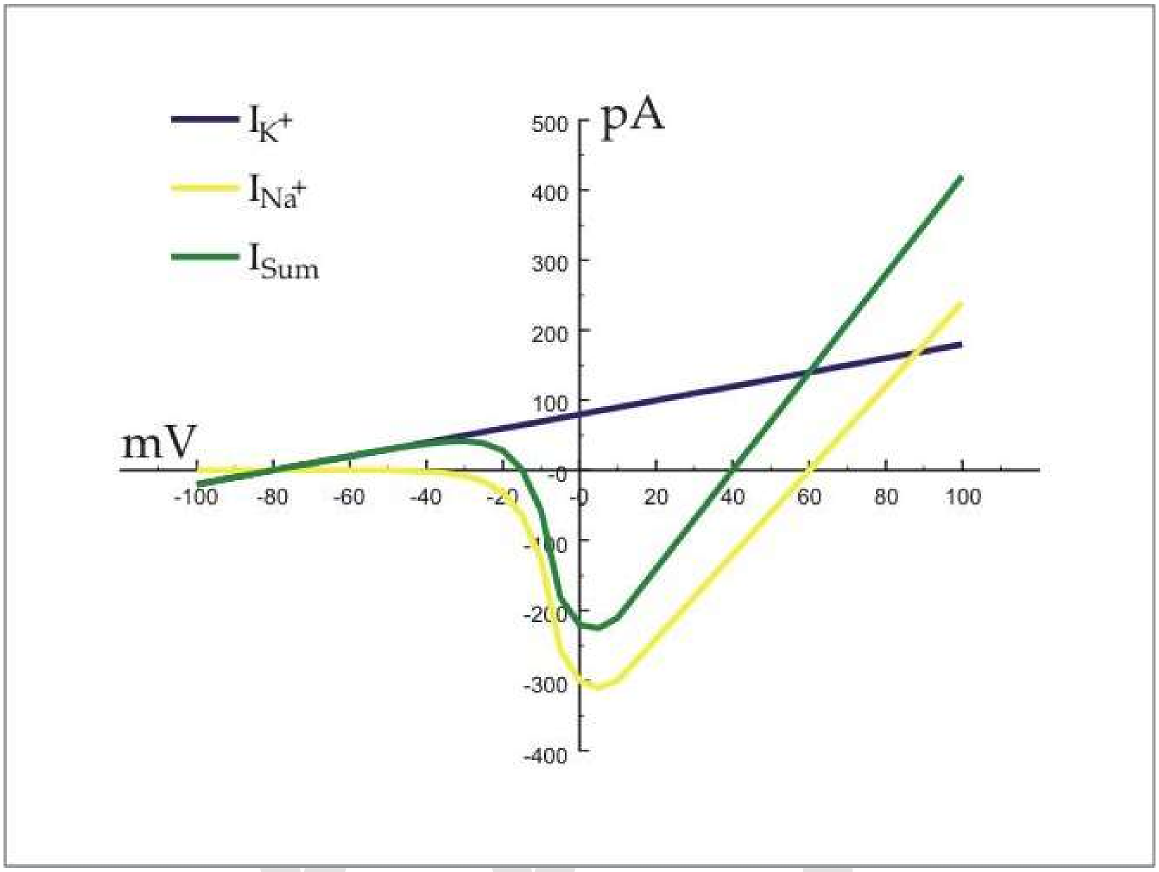



Figure 3. An approximation of the potassium and sodium ion components of a so-called "whole cell" I–V curve of a neuron.

While *I–V* curves are applicable to any electrical system, they find wide use in the field of biological electricity, particularly in the sub-field of electrophysiology. In this case, the voltage refers to the voltage across a biological membrane, a membrane potential, and the current is the flow of charged ions across channels in this membrane. The current is determined by the conductances of these channels.

In the case of ionic current across biological membranes, currents are measured from inside to outside. That is, positive currents, known as "outward current", corresponding to positively charged ions crossing a cell membrane from the inside to the outside, or a negatively charged ion crossing from the outside to the inside. Similarly, currents with a negative value are referred to as "inward current", corresponding to positively charged ions crossing a cell membrane from the outside to the inside, or a negatively charged ion crossing from inside to outside

Figure 3 shows an *I-V* curve that is more relevant to the currents in excitable biological membranes (such as a neuronal axon). The blue line shows the *I–V* relationship for the potassium ion. Note that it is linear, indicating no voltage-dependent gating of the potassium ion channel. The yellow line shows the *I–V* relationship for the sodium ion. Note that it is not linear, indicating that the sodium ion channel is voltage-dependent. The green line indicates the *I–V* relationship derived from summing the sodium and potassium currents. This approximates the actual membrane potential and current relationship of a cell containing both types of channel.

# Chapter-14

# Bipolar Transistor Biasing

Bipolar transistor amplifiers must be properly biased to operate correctly. In circuits made with individual devices (discrete circuits), biasing networks consisting of resistors are commonly employed. Much more elaborate biasing arrangements are used in integrated circuits, for example, bandgap voltage references and current mirrors.

The operating point of a device, also known as *bias point*, *quiescent point*, or *Q-point*, is the point on the output characteristics that shows the DC collector–emitter voltage ($V_{ce}$) and the collector current ($I_c$) with no input signal applied. The term is normally used in connection with devices such as transistors.

## *Bias circuit requirements*

### Signal requirements for Class A amplifiers

For analog circuit operation, the Q-point is placed so the transistor stays in **active mode** (does not shift to operation in the saturation region or cut-off region) when input is applied. For digital operation, the Q-point is placed so the transistor does the contrary - switches from "on" to "off" state. Often, Q-point is established near the center of active region of transistor characteristic to allow similar signal swings in positive and negative directions. Q-point should be stable. In particular, it should be insensitive to variations in transistor parameters (for example, should not shift if transistor is replaced by another of the same type), variations in temperature, variations in power supply voltage and so forth. The circuit must be practical: easily implemented and cost-effective.

### Thermal considerations

At constant current, the voltage across the emitter–base junction $V_{BE}$ of a bipolar transistor *decreases* 2 mV (silicon) and 1.8mV (germanium) for each 1°C rise in temperature (reference being 25°C). By the Ebers–Moll model, if the base–emitter voltage $V_{BE}$ is held constant and the temperature rises, the current through the base–emitter diode $I_B$ will increase, and thus the collector current $I_C$ will also increase.

Depending on the bias point, the power dissipated in the transistor may also increase, which will further increase its temperature and exacerbate the problem. This deleterious positive feedback results in **thermal runaway**. There are several approaches to mitigate bipolar transistor thermal runaway. For example,

- Negative feedback can be built into the biasing circuit so that increased collector current leads to decreased base current. Hence, the increasing collector current throttles its source.
- Heat sinks can be used that carry away extra heat and prevent the base–emitter temperature from rising.
- The transistor can be biased so that its collector is normally less than half of the power supply voltage, which implies that collector–emitter power dissipation is at its maximum value. Runaway is then impossible because increasing collector current leads to a decrease in dissipated power; this notion is known as the *half-voltage principle.*

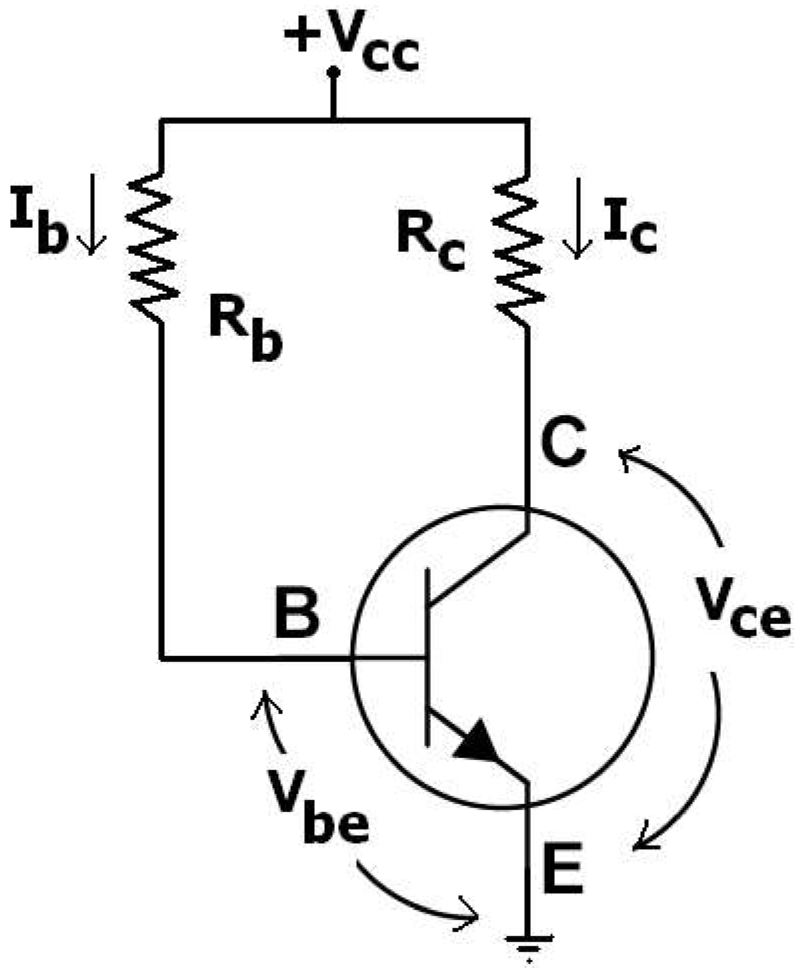
The circuits below primarily demonstrate the use of negative feedback to prevent thermal runaway.

## *Types of bias circuit for Class A amplifiers*

The following discussion treats five common biasing circuits used with Class A bipolar transistor amplifiers:

1. Fixed bias
2. Collector-to-base bias
3. Fixed bias with emitter resistor
4. Voltage divider bias
5. Emitter bias

**Fixed bias (base bias)**




Fixed bias (Base bias)

This form of biasing is also called *base bias*. In the example image on the right, the single power source (for example, a battery) is used for both collector and base of transistor, although separate batteries can also be used.

In the given circuit,

$$V_{cc} = I_B R_B + V_{be}$$

Therefore,

$$I_B = (V_{cc} - V_{be})/R_B$$

For a given transistor, $V_{be}$ does not vary significantly during use. As $V_{cc}$ is of fixed value, on selection of $R_B$, the base current $I_B$ is fixed. Therefore this type is called *fixed bias* type of circuit.

Also for given circuit,

$$V_{cc} = I_C R_C + V_{ce}$$

Therefore,

$$V_{ce} = V_{cc} - I_C R_C$$

The common-emitter current gain of a transistor is an important parameter in circuit design, and is specified on the data sheet for a particular transistor. It is denoted as β on this page.

Because

$$I_C = \beta I_B$$

we can obtain $I_C$ as well. In this manner, operating point given as $(V_{ce}, I_C)$ can be set for given transistor.

**Merits:**

- It is simple to shift the operating point anywhere in the active region by merely changing the base resistor ($R_B$).
- A very small number of components are required.

**Demerits:**

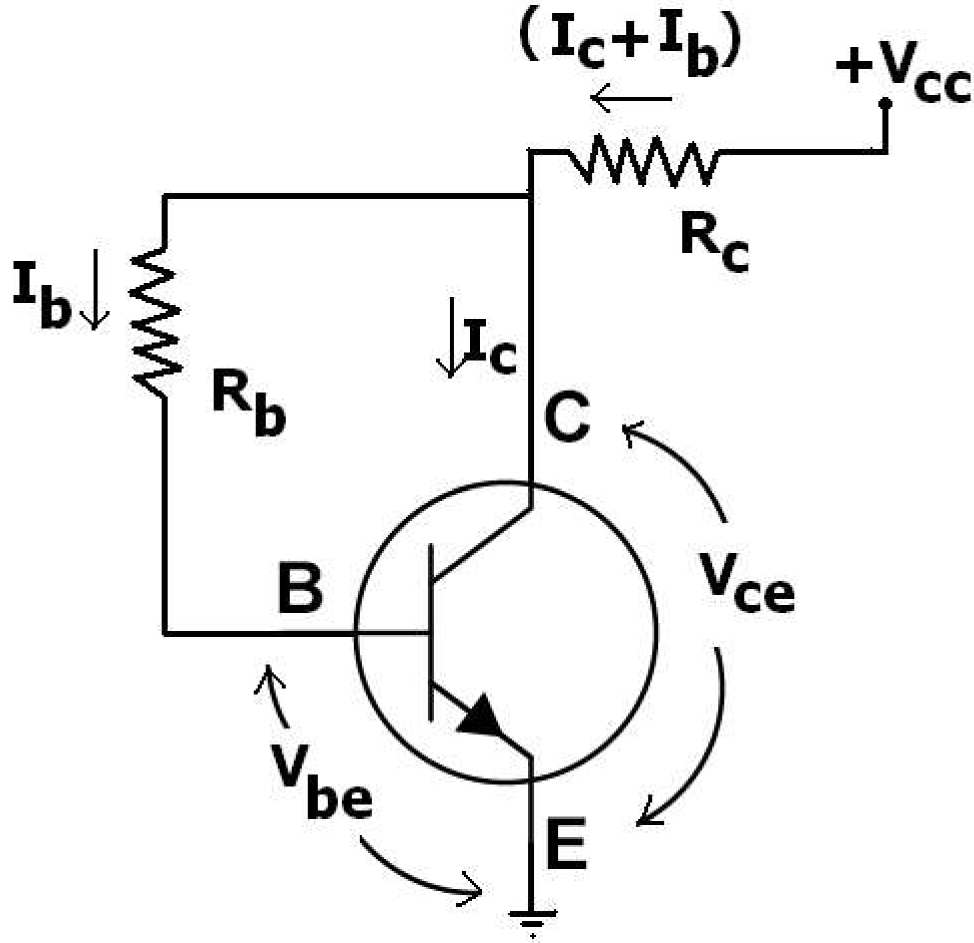
- The collector current does not remain constant with variation in temperature or power supply voltage. Therefore the operating point is unstable.

- Changes in $V_{be}$ will change $I_B$ and thus cause $R_E$ to change. This in turn will alter the gain of the stage.
- When the transistor is replaced with another one, considerable change in the value of β can be expected. Due to this change the operating point will shift.
- For small-signal transistors (e.g., not power transistors) with relatively high values of β (i.e., between 100 and 200), this configuration will be prone to thermal runaway. In particular, the stability factor, which is a measure of the change in collector current with changes in reverse saturation current, is approximately β+1. To ensure absolute stability of the amplifier, a stability factor of less than 25 is preferred, and so small-signal transistors have large stability factors.

**Usage:**

Due to the above inherent drawbacks, fixed bias is rarely used in linear circuits (i.e., those circuits which use the transistor as a current source). Instead, it is often used in circuits where transistor is used as a switch. However, one application of fixed bias is to achieve crude automatic gain control in the transistor by feeding the base resistor from a DC signal derived from the AC output of a later stage.

### Collector-to-base bias




Collector-to-base bias

This configuration employs negative feedback to prevent thermal runaway and stabilize the operating point. In this form of biasing, the base resistor $R_B$ is connected to the collector instead of connecting it to the DC source $V_{cc}$. So any thermal runaway will induce a voltage drop across the $R_C$ resistor that will throttle the transistor's base current.

From Kirchhoff's voltage law, the voltage $V_{R_b}$ across the base resistor $R_b$ is

$$V_{R_b} = V_{cc} - \overbrace{(I_c + I_b)R_c}^{\text{Voltage drop across } R_c} - \overbrace{V_{be}}^{\text{Voltage at base}} .$$

By the Ebers–Moll model, $I_c = \beta I_b$, and so

$$V_{\mathrm{R_b}} = V_{\mathrm{cc}} - (\overbrace{\beta I_{\mathrm{b}}}^{I_{\mathrm{c}}} + I_{\mathrm{b}}) R_{\mathrm{c}} - V_{\mathrm{be}} = V_{\mathrm{cc}} - I_{\mathrm{b}}(\beta + 1) R_{\mathrm{c}} - V_{\mathrm{be}}.$$

From Ohm's law, the base current $I_{\mathrm{b}} = V_{\mathrm{R_b}} / R_{\mathrm{b}}$, and so

$$\overbrace{I_{\mathrm{b}} R_{\mathrm{b}}}^{V_{\mathrm{R_b}}} = V_{\mathrm{cc}} - I_{\mathrm{b}}(\beta + 1) R_{\mathrm{c}} - V_{\mathrm{be}}.$$

Hence, the base current $I_b$ is

$$I_{\mathrm{b}} = \frac{V_{\mathrm{cc}} - V_{\mathrm{be}}}{R_{\mathrm{b}} + (\beta + 1) R_{\mathrm{c}}}$$

If $V_{be}$ is held constant and temperature increases, then the collector current $I_c$ increases. However, a larger $I_c$ causes the voltage drop across resistor $R_c$ to increase, which in turn reduces the voltage $V_{\mathrm{R_b}}$ across the base resistor $R_b$. A lower base-resistor voltage drop reduces the base current $I_b$, which results in less collector current $I_c$. Because an increase in collector current with temperature is opposed, the operating point is kept stable.

**Merits:**

- Circuit stabilizes the operating point against variations in temperature and β (i.e. replacement of transistor)

**Demerits:**

- In this circuit, to keep $I_c$ independent of β, the following condition must be met:

$$I_{\mathrm{c}} = \beta I_{\mathrm{b}} = \frac{\beta (V_{\mathrm{cc}} - V_{\mathrm{be}})}{R_{\mathrm{b}} + R_{\mathrm{c}} + \beta R_{\mathrm{c}}} \approx \frac{(V_{\mathrm{cc}} - V_{\mathrm{be}})}{R_{\mathrm{c}}}$$

which is the case when

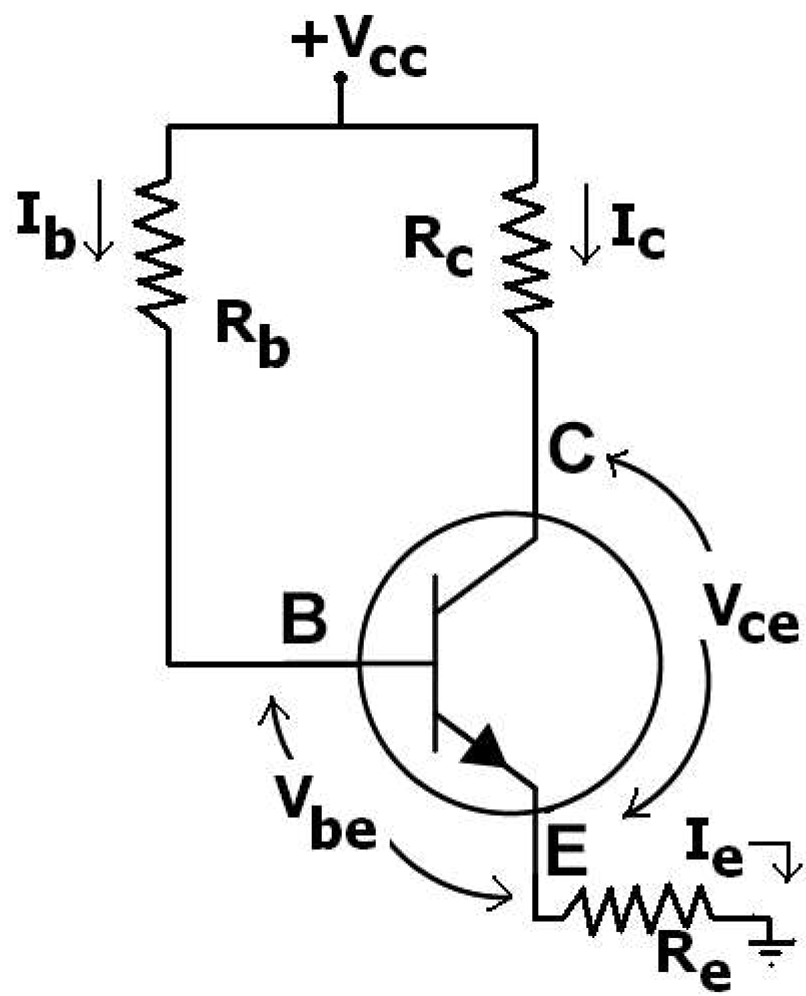
$$\beta R_{\mathrm{c}} \gg R_{\mathrm{b}}.$$

- As β-value is fixed (and generally unknown) for a given transistor, this relation can be satisfied either by keeping $R_c$ fairly large or making $R_b$ very low.
  - If $R_c$ is large, a high $V_{cc}$ is necessary, which increases cost as well as precautions necessary while handling.
  - If $R_b$ is low, the reverse bias of the collector–base region is small, which limits the range of collector voltage swing that leaves the transistor in active mode.

- The resistor $R_b$ causes an AC feedback, reducing the voltage gain of the amplifier. This undesirable effect is a trade-off for greater Q-point stability.

**Usage:** The feedback also decreases the input impedance of the amplifier as seen from the base, which can be advantageous. Due to the gain reduction from feedback, this biasing form is used only when the trade-off for stability is warranted.

## Fixed bias with emitter resistor




Fixed bias with emitter resistor

The fixed bias circuit is modified by attaching an external resistor to the emitter. This resistor introduces negative feedback that stabilizes the Q-point. From Kirchhoff's voltage law, the voltage across the base resistor is

$V_{Rb} = V_{CC} - I_e R_e - V_{be}$.

From Ohm's law, the base current is

$I_b = V_{Rb} / R_b$.

The way feedback controls the bias point is as follows. If $V_{be}$ is held constant and temperature increases, emitter current increases. However, a larger $I_e$ increases the emitter voltage $V_e = I_eR_e$, which in turn reduces the voltage $V_{Rb}$ across the base resistor. A lower base-resistor voltage drop reduces the base current, which results in less collector current because $I_c = ß\ I_B$. Collector current and emitter current are related by $I_c = \alpha\ I_e$ with $\alpha \approx 1$, so increase in emitter current with temperature is opposed, and operating point is kept stable.

Similarly, if the transistor is replaced by another, there may be a change in $I_C$ (corresponding to change in β-value, for example). By similar process as above, the change is negated and operating point kept stable.

For the given circuit,

$I_B = (V_{CC} - V_{be})/(R_B + (\beta+1)R_E)$.

**Merits:**

The circuit has the tendency to stabilize operating point against changes in temperature and β-value.

**Demerits:**

- In this circuit, to keep $I_C$ independent of β the following condition must be met:

$$I_C = \beta I_B = \frac{\beta(V_{CC} - V_{be})}{R_B + (\beta+1)R_E} \approx \frac{(V_{CC} - V_{be})}{R_E}$$

which is approximately the case if

$$( \beta + 1 )R_E >> R_B.$$

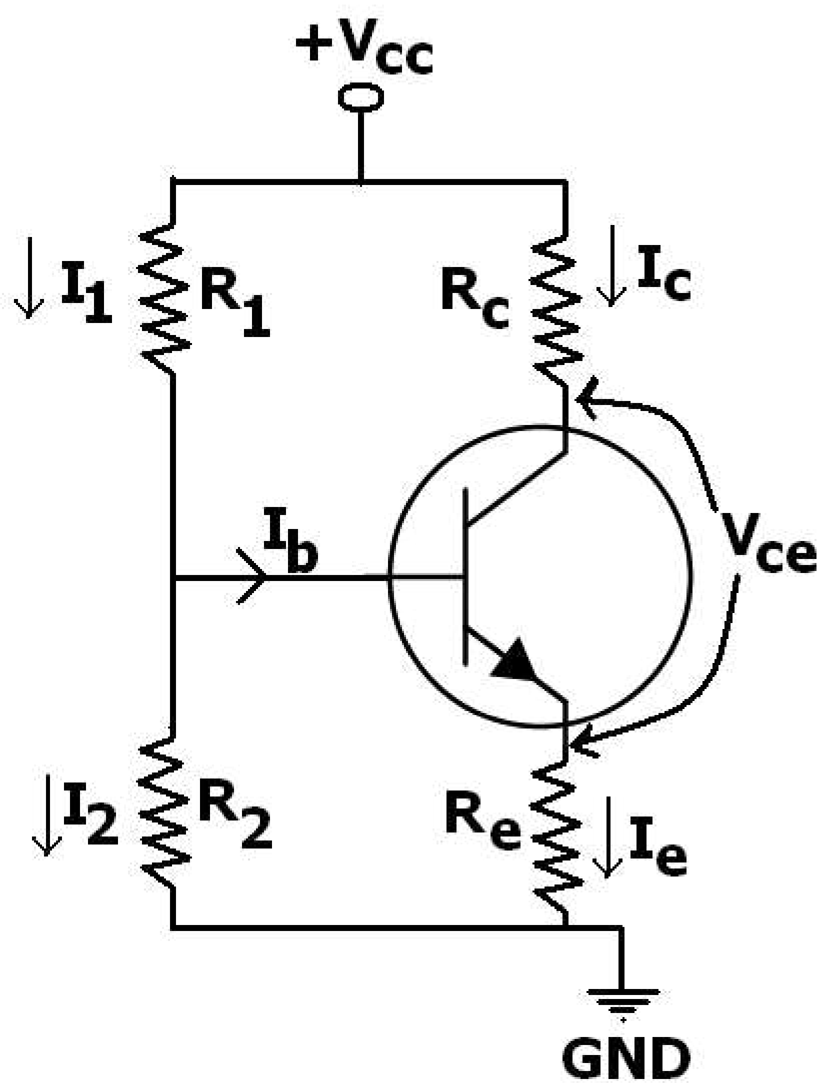
- As β-value is fixed for a given transistor, this relation can be satisfied either by keeping $R_E$ very large, or making $R_B$ very low.
    - If $R_E$ is of large value, high $V_{CC}$ is necessary. This increases cost as well as precautions necessary while handling.
    - If $R_B$ is low, a separate low voltage supply should be used in the base circuit. Using two supplies of different voltages is impractical.
- In addition to the above, $R_E$ causes ac feedback which reduces the voltage gain of the amplifier.

**Usage:**

The feedback also increases the input impedance of the amplifier when seen from the base, which can be advantageous. Due to the above disadvantages, this type of biasing circuit is used only with careful consideration of the trade-offs involved.

## Collector-Stabilized Biasing

## Voltage divider bias




Voltage divider bias

The voltage divider is formed using external resistors $R_1$ and $R_2$. The voltage across $R_2$ forward biases the emitter junction. By proper selection of resistors $R_1$ and $R_2$, the operating point of the transistor can be made independent of β. In this circuit, the voltage divider holds the base voltage fixed independent of base current provided the divider current is large compared to the base current. However, even with a fixed base voltage, collector current varies with temperature (for example) so an emitter resistor is added to stabilize the Q-point, similar to the above circuits with emitter resistor.

In this circuit the base voltage is given by:

$V_B =$ voltage across $R_2 = V_{cc}\frac{R_2}{(R_1+R_2)} - I_B\frac{R_1 R_2}{(R_1+R_2)}$

$\approx V_{cc}\frac{R_2}{(R_1+R_2)}$ provided $I_B << I_2 = V_B/R_2$.

Also $V_B = V_{be} + I_E R_E$

For the given circuit,

$$I_B = \frac{\frac{V_{CC}}{1+R_1/R_2} - V_{be}}{(\beta+1)R_E + R_1 \parallel R_2}.$$

**Merits:**

- Unlike above circuits, only one dc supply is necessary.
- Operating point is almost independent of β variation.
- Operating point stabilized against shift in temperature.

**Demerits:**

- In this circuit, to keep $I_C$ independent of β the following condition must be met:

$$I_C = \beta I_B = \beta\frac{\frac{V_{CC}}{1+R_1/R_2} - V_{be}}{(\beta+1)R_E + R_1 \parallel R_2} \approx \frac{\frac{V_{CC}}{1+R_1/R_2} - V_{be}}{R_E},$$

which is approximately the case if

$$(\beta+1)R_E >> R_1 \parallel R_2$$

where $R_1 \parallel R_2$ denotes the equivalent resistance of $R_1$ and $R_2$ connected in parallel.

- As β-value is fixed for a given transistor, this relation can be satisfied either by keeping $R_E$ fairly large, or making $R_1 \| R_2$ very low.

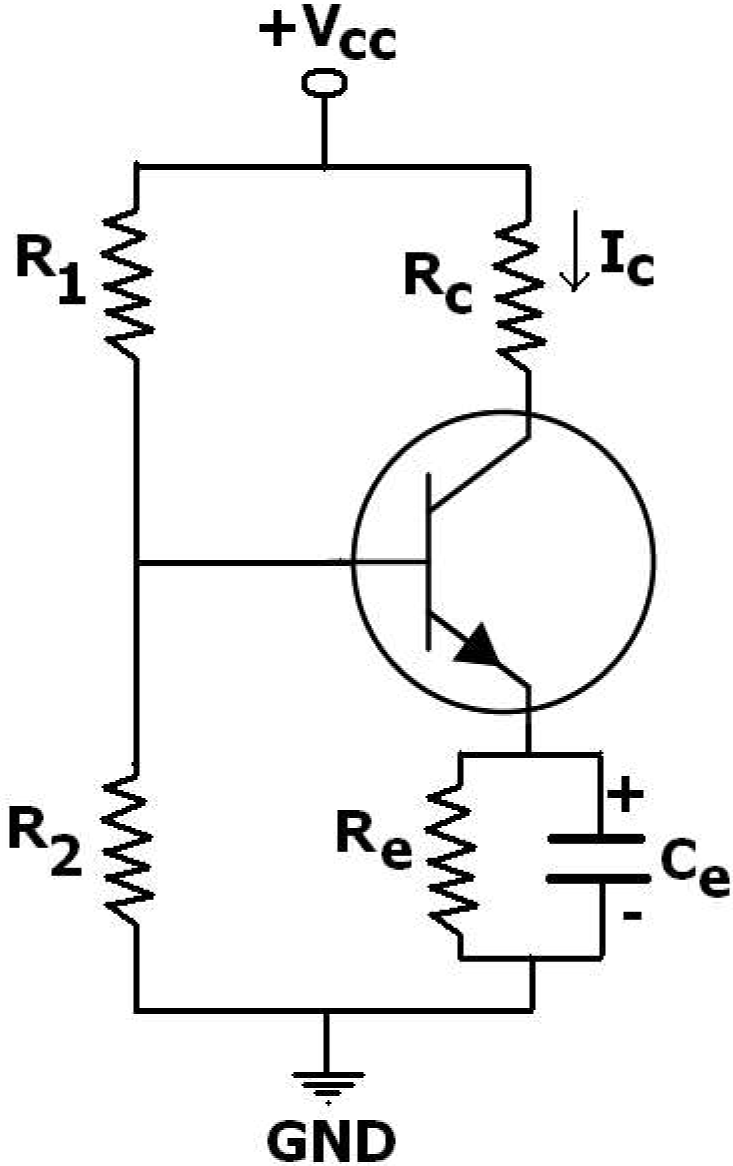
- If $R_E$ is of large value, high $V_{CC}$ is necessary. This increases cost as well as precautions necessary while handling.
- If $R_1 \parallel R_2$ is low, either $R_1$ is low, or $R_2$ is low, or both are low. A low $R_1$ raises $V_B$ closer to $V_C$, reducing the available swing in collector voltage, and limiting how large $R_C$ can be made without driving the transistor out of active mode. A low $R_2$ lowers $V_{be}$, reducing the allowed collector current. Lowering both resistor values draws more current from the power supply and lowers the input resistance of the amplifier as seen from the base.

- AC as well as DC feedback is caused by $R_E$, which reduces the AC voltage gain of the amplifier. A method to avoid AC feedback while retaining DC feedback is discussed below.

**Usage:**

The circuit's stability and merits as above make it widely used for linear circuits.

**Voltage divider with AC bypass capacitor**



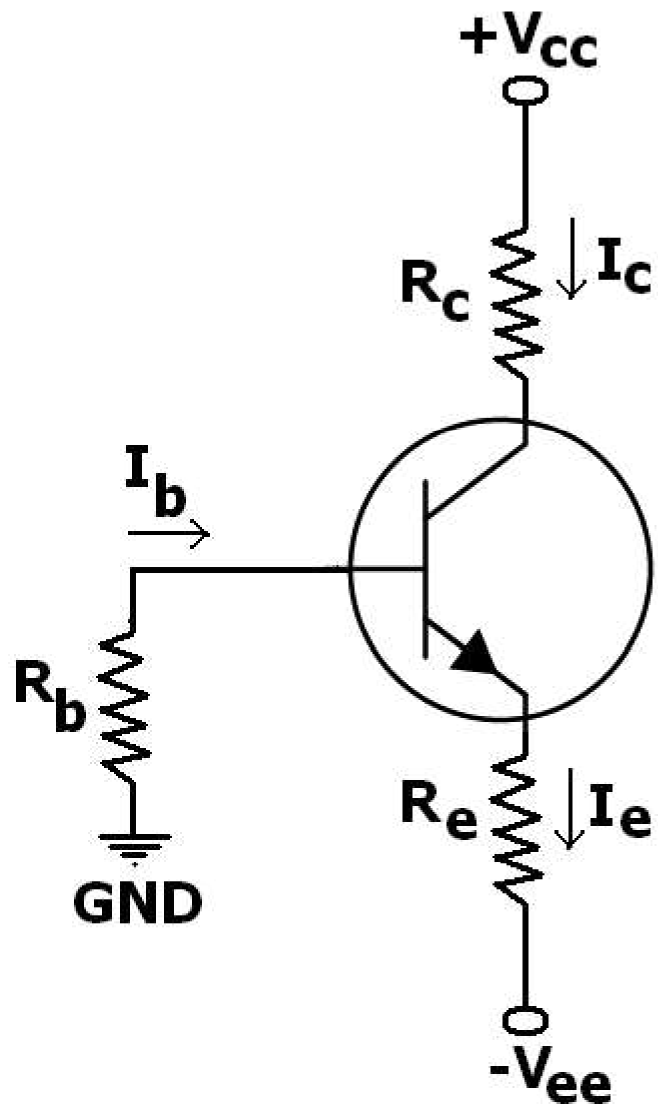

Voltage divider with capacitor

The standard voltage divider circuit discussed above faces a drawback - AC feedback caused by resistor $R_E$ reduces the gain. This can be avoided by placing a capacitor ($C_E$) in parallel with $R_E$, as shown in circuit diagram.

This capacitor is usually chosen to have a low enough reactance at the signal frequencies of interest such that $R_E$ is essentially shorted at AC, thus grounding the emitter. Feedback is therefore only present at DC to stabilize the operating point, in which case any AC advantages of feedback are lost.

Of course, this idea can be used to shunt only a portion of $R_E$, thereby retaining some AC feedback.

### Emitter bias




Emitter bias

When a split supply (dual power supply) is available, this biasing circuit is the most effective, and provides zero bias voltage at the emitter or collector for load. The negative supply $V_{EE}$ is used to forward-bias the emitter junction through $R_E$. The positive supply $V_{CC}$ is used to reverse-bias the collector junction. Only two resistors are necessary for the common collector stage and four resistors for the common emitter or common base stage.

We know that,

$V_B - V_E = V_{be}$

If $R_B$ is small enough, base voltage will be approximately zero. Therefore emitter current is,

$I_E = (V_{EE} - V_{be})/R_E$

The operating point is independent of β if $R_E >> R_B/\beta$

**Merit:**

Good stability of operating point similar to voltage divider bias.

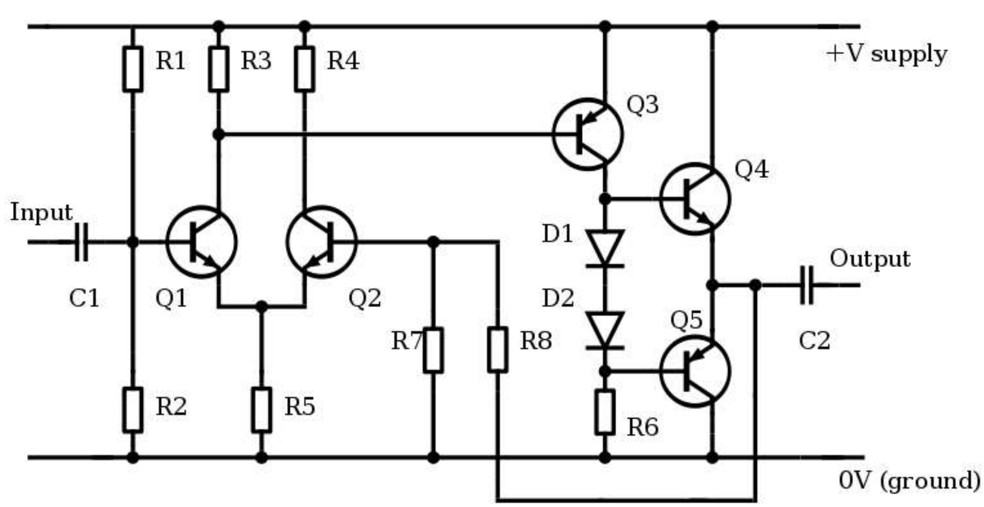
**Demerit:**

This type can only be used when a split (dual) power supply is available.

## *Class B and AB amplifiers*

### Signal requirements

Class B and AB amplifiers employ 2 active devices to cover the complete 360 deg of input signal flow. Each transistor is therefore biased to perform over approximately 180 deg of the input signal. Class B bias is when the collector current $I_c$ with no signal is just conducting (about 1 % of maximum possible value). Class AB bias is when the collector current $I_c$ is about 1/4 of maximum possible value. The class AB push–pull output amplifier circuit below could be the basis for a moderate-power audio amplifier.




A practical amplifier circuit

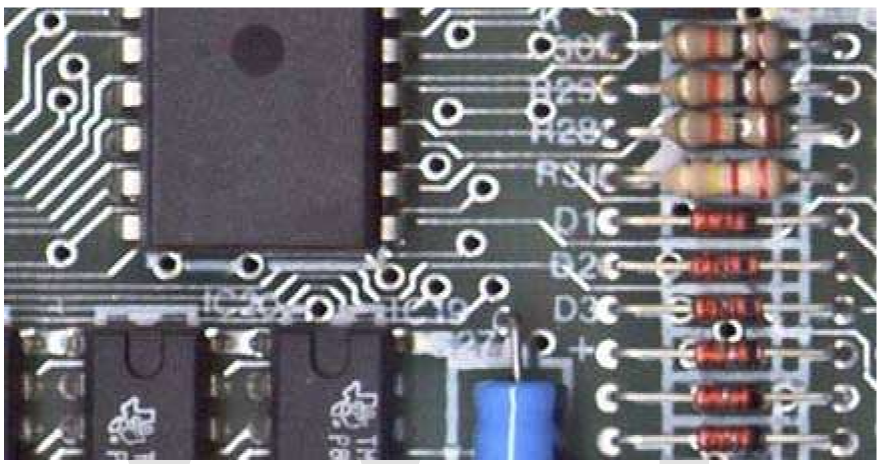
Q3 is a common emitter stage that provides amplification of the signal and the DC bias current through D1 and D2 to generate a bias voltage for the output devices. The output pair are arranged in Class AB push–pull, also called a complementary pair. The diodes D1 and D2 provide a small amount of constant voltage bias for the output pair, just biasing them into the conducting state so that crossover distortion is minimized. That is, the diodes push the output stage into class-AB mode (assuming that the base-emitter drop of the output transistors is reduced by heat dissipation).

This design automatically stabilizes its operating point, since overall feedback internally operates from DC up through the audio range and beyond. The use of fixed diode bias requires the diodes to be both electrically and thermally matched to the output transistors. If the output transistors conduct too much, they can easily overheat and destroy themselves, as the full current from the power supply is not limited at this stage.

A common solution to help stabilize the output device operating point is to include some emitter resistors, typically an ohm or so. Calculating the values of the circuit's resistors and capacitors is done based on the components employed and the intended use of the amplifier.

**Chapter-15**

# Printed Circuit Board



Part of a 1983 Sinclair ZX Spectrum computer board; a populated PCB, showing the conductive traces, vias (the through-hole paths to the other surface), and some mounted electrical components

A **printed circuit board**, or **PCB**, is used to mechanically support and electrically connect electronic components using conductive pathways, tracks or signal traces etched from copper sheets laminated onto a non-conductive *substrate*. It is also referred to as **printed wiring board** (**PWB**) or **etched wiring board**. A PCB populated with electronic components is a **printed circuit assembly** (**PCA**), also known as a **printed circuit board assembly** (**PCBA**). Printed circuit boards are used in virtually all but the simplest commercially-produced electronic devices.

PCBs are inexpensive, and can be highly reliable. They require much more layout effort and higher initial cost than either wire wrap or point-to-point construction, but are much cheaper and faster for high-volume production; the production and soldering of PCBs can be done by totally automated equipment. Much of the electronics industry's PCB design, assembly, and quality control needs are set by standards that are published by the IPC organization.

## *History*

The inventor of the printed circuit was the Austrian engineer Paul Eisler who, while working in England, made one circa 1936 as part of a radio set. Around 1943 the USA began to use the technology on a large scale to make rugged radios for use in World War II. After the war, in 1948, the USA released the invention for commercial use. Printed circuits did not become commonplace in consumer electronics until the mid-1950s, after the *Auto-Sembly* process was developed by the United States Army.

Before printed circuits (and for a while after their invention), point-to-point construction was used. For prototypes, or small production runs, wire wrap or turret board can be more efficient. Predating the printed circuit invention, and similar in spirit, was John Sargrove's 1936-1947 Electronic Circuit Making Equipment (ECME) which sprayed metal onto a Bakelite plastic board. The ECME could produce 3 radios per minute.

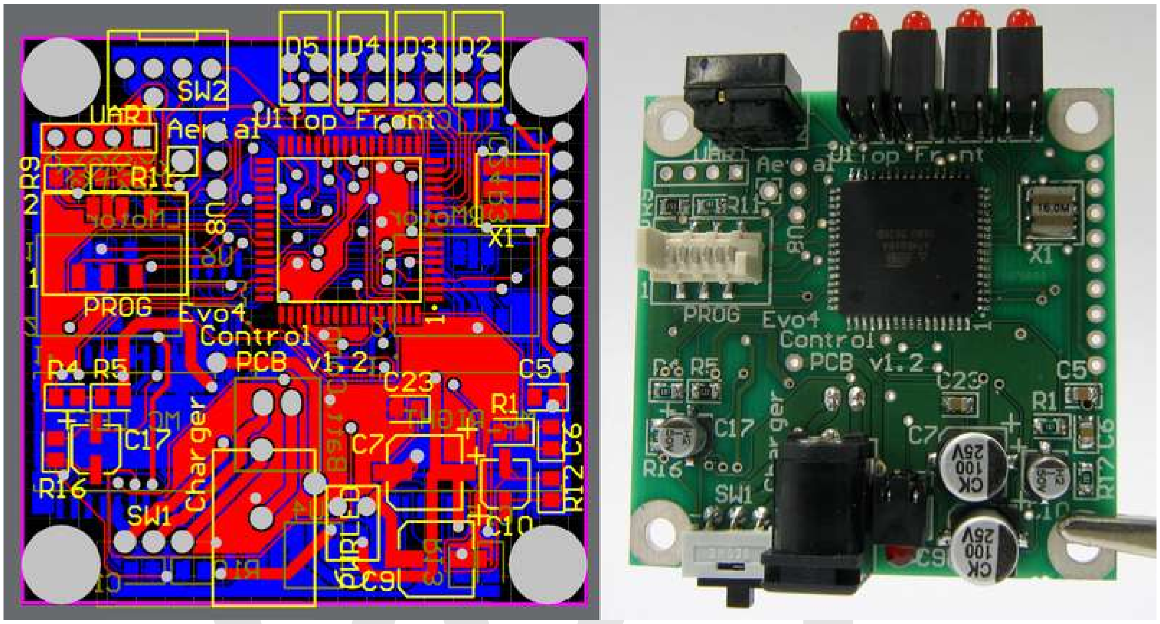
During World War II, the development of the anti-aircraft proximity fuse required an electronic circuit that could withstand being fired from a gun, and could be produced in quantity. The Centralab Division of Globe Union submitted a proposal which met the requirements: a ceramic plate would be screenprinted with metallic paint for conductors and carbon material for resistors, with ceramic disc capacitors and subminiature vacuum tubes soldered in place.

Originally, every electronic component had wire leads, and the PCB had holes drilled for each wire of each component. The components' leads were then passed through the holes and soldered to the PCB trace. This method of assembly is called *through-hole* construction. In 1949, Moe Abramson and Stanislaus F. Danko of the United States Army Signal Corps developed the Auto-Sembly process in which component leads were inserted into a copper foil interconnection pattern and dip soldered. With the development of board lamination and etching techniques, this concept evolved into the standard printed circuit board fabrication process in use today. Soldering could be done automatically by passing the board over a ripple, or wave, of molten solder in a wave-soldering machine. However, the wires and holes are wasteful since drilling holes is expensive and the protruding wires are merely cut off.

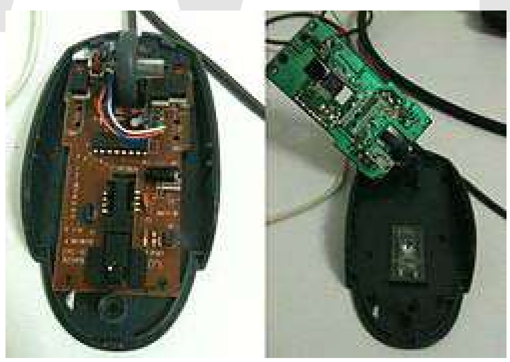
In recent years, the use of surface mount parts has gained popularity as the demand for smaller electronics packaging and greater functionality has grown.

## *Manufacturing*

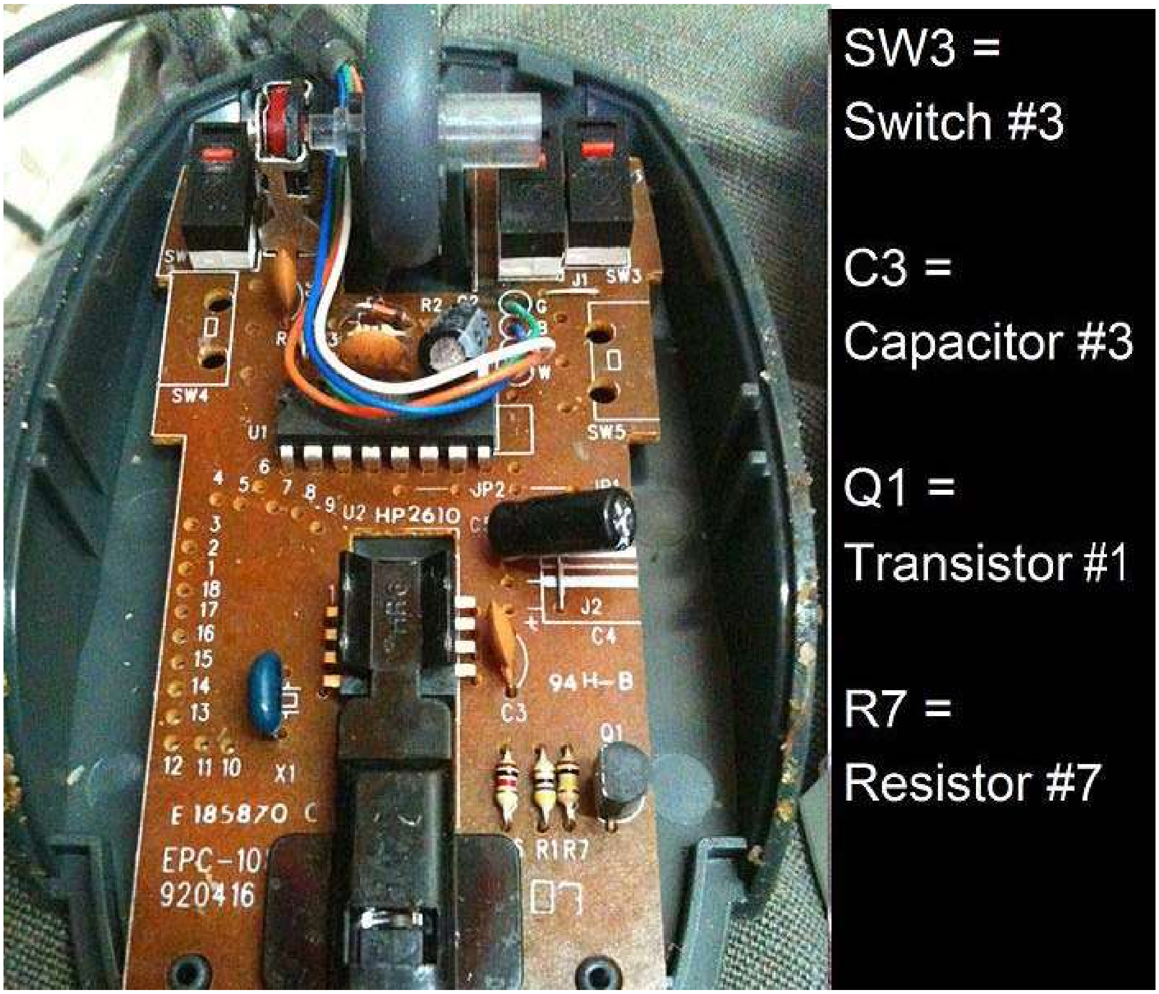
### Materials




A PCB as a design on a computer (left) and realized as a board assembly populated with components (right). The board is double sided, with through-hole plating, green solder resist, and white silkscreen printing. Both surface mount and through-hole components have been used.



A PCB in a computer mouse. The Component Side (left) and the printed side (right).




The Component Side of a PCB in a computer mouse; some examples for common components and their reference designations on the silk screen.

Conducting layers are typically made of thin copper foil. Insulating layers dielectric are typically laminated together with epoxy resin prepreg. The board is typically coated with a solder mask that is green in color. Other colors that are normally available are blue, black, white and red. There are quite a few different dielectrics that can be chosen to provide different insulating values depending on the requirements of the circuit. Some of these dielectrics are polytetrafluoroethylene (Teflon), FR-4, FR-1, CEM-1 or CEM-3. Well known prepreg materials used in the PCB industry are FR-2 (Phenolic cotton paper), FR-3 (Cotton paper and epoxy), FR-4 (Woven glass and epoxy), FR-5 (Woven glass and epoxy), FR-6 (Matte glass and polyester), G-10 (Woven glass and epoxy), CEM-1 (Cotton paper and epoxy), CEM-2 (Cotton paper and epoxy), CEM-3 (Woven glass and epoxy), CEM-4 (Woven glass and epoxy), CEM-5 (Woven glass and polyester). Thermal expansion is an important consideration especially with BGA and naked die technologies, and glass fiber offers the best dimensional stability.

FR-4 is by far the most common material used today. The board with copper on it is called "copper-clad laminate".

Copper foil thickness can be specified in ounces per square foot or micrometres. One ounce per square foot is 1.344 mils or 34 micrometres.

## Patterning (etching)

The vast majority of printed circuit boards are made by bonding a layer of copper over the entire substrate, sometimes on both sides, (creating a "blank PCB") then removing unwanted copper after applying a temporary mask (e.g. by etching), leaving only the desired copper traces. A few PCBs are made by *adding* traces to the bare substrate (or a substrate with a very thin layer of copper) usually by a complex process of multiple electroplating steps. The PCB manufacturing method primarily depends on whether it is for production volume or sample/prototype quantities.

### Commercial (production quantities, usually PTH)

- silk screen printing -the main commercial method.
- Photographic methods. Used when fine linewidths are required.

### Hobbyist/prototype (small quantities, usually not PTH)

- Laser-printed resist: Laser-print onto paper (or wax paper), heat-transfer with an iron or modified laminator onto bare laminate, then etch.
- Print onto transparent film and use as photomask along with photo-sensitized boards. (i.e. pre-sensitized boards), Then etch. (Alternatively, use a film photoplotter).
- Laser resist ablation: Spray black paint onto copper clad laminate, place into CNC laser plotter. The laser raster-scans the PCB and ablates (vaporizes) the paint where no resist is wanted. Etch. (Note: laser copper ablation is rarely used and is considered experimental.)
- Use a CNC-mill with a spade-shaped (i.e. 45-degree) cutter or miniature end-mill to route away the undesired copper, leaving only the traces.

There are three common "subtractive" methods (methods that remove copper) used for the production of printed circuit boards:

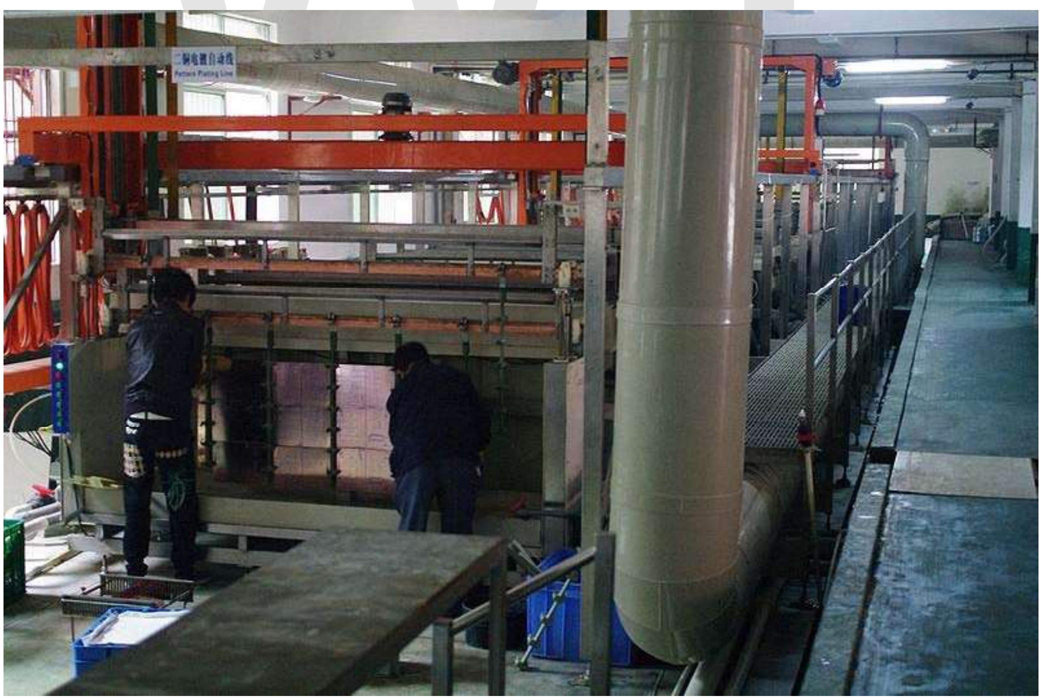
1. **Silk screen printing** uses etch-resistant inks to protect the copper foil. Subsequent etching removes the unwanted copper. Alternatively, the ink may be conductive, printed on a blank (non-conductive) board. The latter technique is also used in the manufacture of hybrid circuits.
2. **Photoengraving** uses a photomask and developer to selectively remove a photoresist coating. The remaining photoresist protects the copper foil. Subsequent etching removes the unwanted copper. The photomask is usually prepared with a photoplotter from data produced by a technician using CAM, or

computer-aided manufacturing software. Laser-printed transparencies are typically employed for *phototools*; however, direct laser imaging techniques are being employed to replace phototools for high-resolution requirements.

3. **PCB milling** uses a two or three-axis mechanical milling system to mill away the copper foil from the substrate. A PCB milling machine (referred to as a 'PCB Prototyper') operates in a similar way to a plotter, receiving commands from the host software that control the position of the milling head in the x, y, and (if relevant) z axis. Data to drive the Prototyper is extracted from files generated in PCB design software and stored in HPGL or Gerber file format.

"Additive" processes also exist. The most common is the "semi-additive" process. In this version, the unpatterned board has a thin layer of copper already on it. A reverse mask is then applied. (Unlike a subtractive process mask, this mask exposes those parts of the substrate that will eventually become the traces.) Additional copper is then plated onto the board in the unmasked areas; copper may be plated to any desired weight. Tin-lead or other surface platings are then applied. The mask is stripped away and a brief etching step removes the now-exposed original copper laminate from the board, isolating the individual traces. Some boards with plated through holes but still single sided were made with a process like this. General Electric made consumer radio sets in the late 1960s using boards like these.

The additive process is commonly used for multi-layer boards as it facilitates the plating-through of the holes (to produce conductive vias) in the circuit board.



PCB copper electroplating machine for adding copper to the in-process PCB

PCB's in process of adding copper via electroplating

The dimensions of the copper conductors of the printed circuit board is related to the amount of current the conductor must carry. Each trace consists of a flat, narrow part of the copper foil that remains after etching. Signal traces are usually narrower than power or ground traces because their current carrying requirements are usually much less. In a multi-layer board one entire layer may be mostly solid copper to act as a ground plane for shielding and power return. For printed circuit boards that contain microwave circuits, transmission lines can be laid out in the form of stripline and microstrip with carefully controlled dimensions to assure a consistent impedance. In radio-frequency circuits the inductance and capacitance of the printed circuit board conductors can be used as a delibrate part of the circuit design, obviating the need for additional discrete components.

## Etching

Chemical etching is done with ferric chloride, ammonium persulfate, or sometimes hydrochloric acid. For PTH (plated-through holes), additional steps of electroless deposition are done after the holes are drilled, then copper is electroplated to build up the thickness, the boards are screened, and plated with tin/lead. The tin/lead becomes the resist leaving the bare copper to be etched away.

## Lamination

Some PCBs have trace layers inside the PCB and are called *multi-layer* PCBs. These are formed by bonding together separately etched thin boards.

## Drilling

Holes through a PCB are typically drilled with tiny drill bits made of solid tungsten carbide. The drilling is performed by automated drilling machines with placement controlled by a *drill tape* or *drill file*. These computer-generated files are also called *numerically controlled drill* (NCD) files or "Excellon files". The drill file describes the location and size of each drilled hole. These holes are often filled with annular rings (hollow rivets) to create vias. Vias allow the electrical and thermal connection of conductors on opposite sides of the PCB.

Most common laminate is epoxy filled fiberglass. Drill bit wear is partly due to embedded glass, which is harder than steel. High drill speed necessary for cost effective drilling of hundreds of holes per board causes very high temperatures at the drill bit tip, and high temperatures (400-700 degrees) soften steel and decompose (oxidize) laminate filler. Copper is softer than epoxy and interior conductors may suffer damage during drilling.

When very small vias are required, drilling with mechanical bits is costly because of high rates of wear and breakage. In this case, the vias may be evaporated by lasers. Laser-drilled vias typically have an inferior surface finish inside the hole. These holes are called *micro vias*.

It is also possible with *controlled-depth* drilling, laser drilling, or by pre-drilling the individual sheets of the PCB before lamination, to produce holes that connect only some of the copper layers, rather than passing through the entire board. These holes are called *blind vias* when they connect an internal copper layer to an outer layer, or *buried vias* when they connect two or more internal copper layers and no outer layers.

The walls of the holes, for boards with 2 or more layers, are made conductive then plated with copper to form *plated-through holes* that electrically connect the conducting layers of the PCB. For multilayer boards, those with 4 layers or more, drilling typically produces a *smear* of the high temperature decomposition products of bonding agent in the laminate system. Before the holes can be plated through, this *smear* must be removed by a chemical *de-smear* process, or by *plasma-etch*. Removing (etching back) the smear also reveals the interior conductors as well.

## Exposed conductor plating and coating

PCBs are plated with solder, tin, or gold over nickel as a resist for etching away the unneeded underlying copper.

After PCBs are etched and then rinsed with water, the soldermask is applied, and then any exposed copper is coated with solder, nickel/gold, or some other anti-corrosion coating.

Matte solder is usually fused to provide a better bonding surface or stripped to bare copper. Treatments, such as benzimidazolethiol, prevent surface oxidation of bare copper. The places to which components will be mounted are typically plated, because untreated bare copper oxidizes quickly, and therefore is not readily solderable. Traditionally, any exposed copper was coated with solder by hot air solder levelling (HASL). The HASL finish prevents oxidation from the underlying copper, thereby guaranteeing a solderable surface. This solder was a tin-lead alloy, however new solder compounds are now used to achieve compliance with the RoHS directive in the EU and US, which restricts the use of lead. One of these lead-free compounds is SN100CL, made up of 99.3% tin, 0.7% copper, 0.05% nickel, and a nominal of 60ppm germanium.

It is important to use solder compatible with both the PCB and the parts used. An example is Ball Grid Array (BGA) using tin-lead solder balls for connections losing their balls on bare copper traces or using lead-free solder paste.

Other platings used are OSP (organic surface protectant), immersion silver (IAg), immersion tin, electroless nickel with immersion gold coating (ENIG), and direct gold plating (over nickel). Edge connectors, placed along one edge of some boards, are often nickel plated then gold plated. Another coating consideration is rapid diffusion of coating metal into Tin solder. Tin forms intermetallics such as $Cu_5Sn_6$ and $Ag_3Cu$ that dissolve into the Tin liquidus or solidus(@50C), stripping surface coating and/or leaving voids.

*Electrochemical migration* (ECM) is the growth of conductive metal filaments on or in a printed circuit board (PCB) under the influence of a DC voltage bias. Silver, zinc, and aluminum are known to grow whiskers under the influence of an electric field. Silver also grows conducting surface paths in the presence of halide and other ions, making it a poor choice for electronics use. Tin will grow "whiskers" due to tension in the plated surface. Tin-Lead or Solder plating also grows whiskers, only reduced by the percentage Tin replaced. Reflow to melt solder or tin plate to relieve surface stress lowers whisker incidence. Another coating issue is tin pest, the transformation of tin to a powdery allotrope at low temperature.

## Solder resist

Areas that should not be soldered may be covered with a polymer *solder resist* (*solder mask*) coating. The solder resist prevents solder from bridging between conductors and creating short circuits. Solder resist also provides some protection from the environment. Solder resist is typically 20-30 micrometres thick.

## Screen printing

Line art and text may be printed onto the outer surfaces of a PCB by screen printing. When space permits, the screen print text can indicate component designators, switch setting requirements, test points, and other features helpful in assembling, testing, and servicing the circuit board.

Screen print is also known as the *silk screen*, or, in one sided PCBs, the *red print*.

Lately some digital printing solutions have been developed to substitute the traditional screen printing process. This technology allows printing variable data onto the PCB, including serialization and barcode information for traceability purposes.

## Test

Unpopulated boards may be subjected to a *bare-board test* where each circuit connection (as defined in a *netlist*) is verified as correct on the finished board. For high-volume production, a Bed of nails tester, a fixture or a Rigid needle adapter is used to make contact with copper lands or holes on one or both sides of the board to facilitate testing. A computer will *instruct* the electrical test unit to apply a small voltage to each contact point on the bed-of-nails as required, and verify that such voltage appears at other appropriate contact points. A "short" on a board would be a connection where there should not be one; an "open" is between two points that should be connected but are not. For small- or medium-volume boards, *flying probe* and *flying-grid* testers use moving test heads to make contact with the copper/silver/gold/solder lands or holes to verify the electrical connectivity of the board under test.

## Printed circuit assembly

After the printed circuit board (PCB) is completed, electronic components must be attached to form a functional *printed circuit assembly*, or PCA (sometimes called a "printed circuit board assembly" PCBA). In *through-hole* construction, component leads are inserted in holes. In *surface-mount* construction, the components are placed on *pads* or *lands* on the outer surfaces of the PCB. In both kinds of construction, component leads are electrically and mechanically fixed to the board with a molten metal solder.

There are a variety of soldering techniques used to attach components to a PCB. High volume production is usually done with machine placement and bulk wave soldering or reflow ovens, but skilled technicians are able to solder very tiny parts (for instance 0201 packages which are 0.02 in. by 0.01 in.) by hand under a microscope, using tweezers and a fine tip soldering iron for small volume prototypes. Some parts are impossible to solder by hand, such as ball grid array (BGA) packages.

Often, through-hole and surface-mount construction must be combined in a single assembly because some required components are available only in surface-mount packages, while others are available only in through-hole packages. Another reason to use both methods is that through-hole mounting can provide needed strength for components likely to endure physical stress, while components that are expected to go untouched will take up less space using surface-mount techniques.

After the board has been populated it may be tested in a variety of ways:

- While the power is off, visual inspection, automated optical inspection. JEDEC guidelines for PCB component placement, soldering, and inspection are commonly used to maintain quality control in this stage of PCB manufacturing.

- While the power is off, analog signature analysis, power-off testing.

- While the power is on, in-circuit test, where physical measurements (i.e. voltage, frequency) can be done.

- While the power is on, functional test, just checking if the PCB does what it had been designed for.

To facilitate these tests, PCBs may be designed with extra pads to make temporary connections. Sometimes these pads must be isolated with resistors. The in-circuit test may also exercise boundary scan test features of some components. In-circuit test systems may also be used to program nonvolatile memory components on the board.

In boundary scan testing, test circuits integrated into various ICs on the board form temporary connections between the PCB traces to test that the ICs are mounted correctly. Boundary scan testing requires that all the ICs to be tested use a standard test configuration procedure, the most common one being the Joint Test Action Group (JTAG) standard. The JTAG test architecture provides a means to test interconnects between integrated circuits on a board without using physical test probes. JTAG tool vendors provide various types of stimulus and sophisticated algorithms, not only to detect the failing nets, but also to isolate the faults to specific nets, devices, and pins.

When boards fail the test, technicians may desolder and replace failed components, a task known as *rework*.

## Protection and packaging

PCBs intended for extreme environments often have a conformal coating, which is applied by dipping or spraying after the components have been soldered. The coat prevents corrosion and leakage currents or shorting due to condensation. The earliest conformal coats were wax; modern conformal coats are usually dips of dilute solutions of silicone rubber, polyurethane, acrylic, or epoxy. Another technique for applying a conformal coating is for plastic to be sputtered onto the PCB in a vacuum chamber. The chief disadvantage of conformal coatings is that servicing of the board is rendered extremely difficult.

Many assembled PCBs are static sensitive, and therefore must be placed in antistatic bags during transport. When handling these boards, the user must be grounded (earthed). Improper handling techniques might transmit an accumulated static charge through the board, damaging or destroying components. Even bare boards are sometimes static sensitive. Traces have become so fine that it's quite possible to blow an etch off the board (or change its characteristics) with a static charge. This is especially true on non-traditional PCBs such as MCMs and microwave PCBs.

## Design

- Schematic capture or schematic entry is done through an EDA tool.
- Card dimensions and template are decided based on required circuitry and case of the PCB. Determine the fixed components and heat sinks if required.
- Deciding stack layers of the PCB. 4 to 12 layers or more depending on design complexity. Ground plane and Power plane are decided. Signal planes where signals are routed are in top layer as well as internal layers.
- Line impedance determination using dielectric layer thickness, routing copper thickness and trace-width. Trace separation also taken into account in case of differential signals. Microstrip, stripline or dual stripline can be used to route signals.
- Placement of the components. Thermal considerations and geometry are taken into account. Vias and lands are marked.
- Routing the signal trace. For optimal EMI performance high frequency signals are routed in internal layers between power or ground planes as power plane behaves as ground for AC.
- Gerber file generation for manufacturing.

## Safety certification (US)

Safety Standard UL 796 covers component safety requirements for printed wiring boards for use as components in devices or appliances. Testing analyzes characteristics such as flammability, maximum operating temperature, electrical tracking, heat deflection, and direct support of live electrical parts.

## "Cordwood" construction

A cordwood module.

Cordwood construction can save significant space and was often used with wire-ended components in applications where space was at a premium (such as missile guidance and telemetry systems) and in high-speed computers, where short traces were important. In "cordwood" construction, axial-leaded components were mounted between two parallel planes. The components were either soldered together with jumper wire, or they were connected to other components by thin nickel ribbon welded at right angles onto the component leads. To avoid shorting together different interconnection layers, thin insulating cards were placed between them. Perforations or holes in the cards allowed component leads to project through to the next interconnection layer. One disadvantage of this system was that special nickel leaded components had to be used to allow the interconnecting welds to be made. Some versions of cordwood construction used single sided PCBs as the interconnection method (as pictured). This meant that normal leaded components could be used. Another disadvantage of this system is that components located in the interior are difficult to replace.

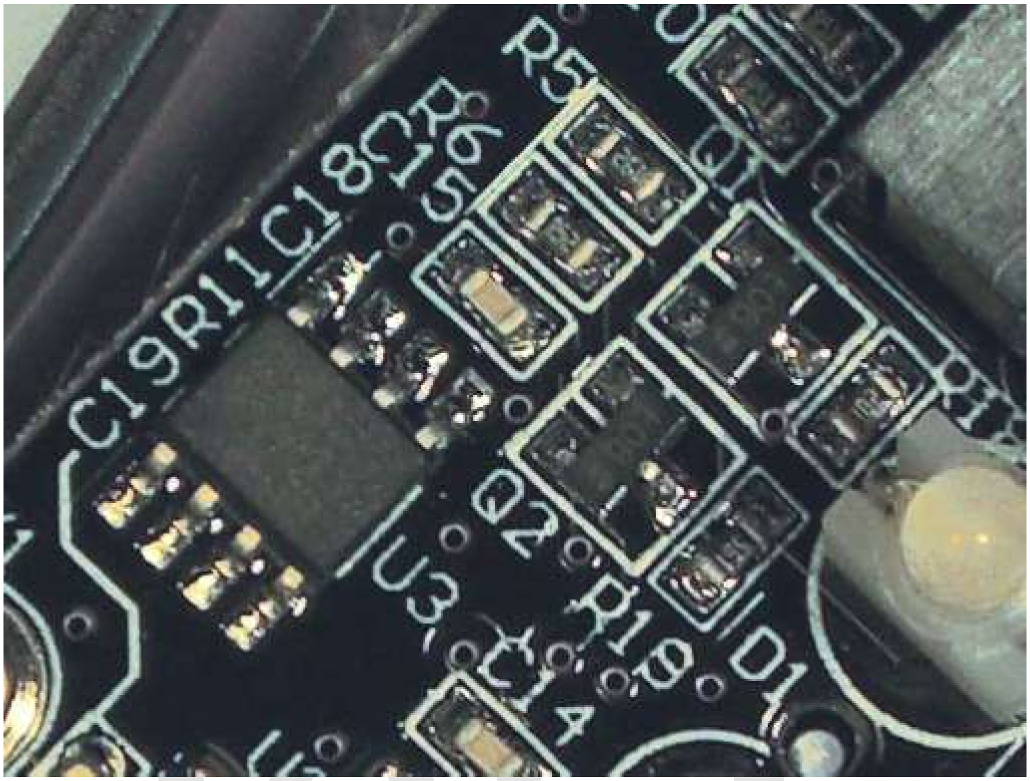
Before the advent of integrated circuits, this method allowed the highest possible component packing density; because of this, it was used by a number of computer vendors including Control Data Corporation. The cordwood method of construction now appears to have fallen into disuse, probably because high packing densities can be more easily achieved using surface mount techniques and integrated circuits.

### *Multiwire boards*

Multiwire is a patented technique of interconnection which uses machine-routed insulated wires embedded in a non-conducting matrix (often plastic resin). It was used during the 1980s and 1990s. (Kollmorgen Technologies Corp., U.S. Patent 4,175,816) Multiwire is still available in 2010 through Hitachi. There are other competitive discrete wiring technologies that have been developed (Jumatech ).

Since it was quite easy to stack interconnections (wires) inside the embedding matrix, the approach allowed designers to forget completely about the routing of wires (usually a time-consuming operation of PCB design): Anywhere the designer needs a connection, the machine will draw a wire in straight line from one location/pin to another. This led to very short design times (no complex algorithms to use even for high density designs) as well as reduced crosstalk (which is worse when wires run parallel to each other—which almost never happens in Multiwire), though the cost is too high to compete with cheaper PCB technologies when large quantities are needed.

### *Surface-mount technology*



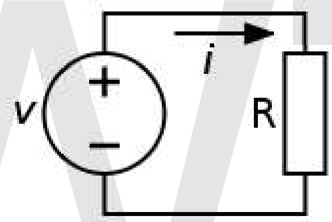

Surface mount components, including resistors, transistors and an integrated circuit

Surface-mount technology emerged in the 1960s, gained momentum in the early 1980s and became widely used by the mid 1990s. Components were mechanically redesigned to have small metal tabs or end caps that could be soldered directly on to the PCB surface. Components became much smaller and component placement on both sides of the board became more common than with through-hole mounting, allowing much higher circuit densities. Surface mounting lends itself well to a high degree of automation, reducing labour costs and greatly increasing production and quality rates. Carrier Tapes provide a stable and protective environment for Surface mount devices (SMDs) which can be one-quarter to one-tenth of the size and weight, and passive components can be one-half to one-quarter of the cost of corresponding through-hole parts. However, integrated circuits are often priced the same regardless of the package type, because the chip itself is the most expensive part. As of 2006, some wire-ended components, such as small-signal switch diodes, e.g. 1N4148, are actually significantly cheaper than corresponding SMD versions.

**Chapter-16**

# Electrical Network and Pre-charge

## Electrical network




A simple electric circuit made up of a voltage source and a resistor. Here, $V = iR$, according to Ohm's Law.

An **electrical network** is an interconnection of electrical elements such as resistors, inductors, capacitors, transmission lines, voltage sources, current sources and switches. An **electrical circuit** is a special type of network, one that has a closed loop giving a return path for the current. Electrical networks that consist only of sources (voltage or current), linear lumped elements (resistors, capacitors, inductors), and linear distributed elements (transmission lines) can be analyzed by algebraic and transform methods to determine DC response, AC response, and transient response.

A network that contains active electronic components is known as an **electronic circuit**. Such networks are generally nonlinear and require more complex design and analysis tools.

## *Design methods*

To design any electrical circuit, either analog or digital, electrical engineers need to be able to predict the voltages and currents at all places within the circuit. Linear circuits, that is, circuits with the same input and output frequency, can be analyzed by hand using complex number theory. Other circuits can only be analyzed with specialized software programs or estimation techniques such as the piecewise-linear model.

Circuit simulation software, such as HSPICE, and languages such as VHDL-AMS and verilog-AMS allow engineers to design circuits without the time, cost and risk of error involved in building circuit prototypes.

## *Electrical laws*

A number of electrical laws apply to all electrical networks. These include:

- Kirchhoff's current law: The sum of all currents entering a node is equal to the sum of all currents leaving the node.
- Kirchhoff's voltage law: The directed sum of the electrical potential differences around a loop must be zero.
- Ohm's law: The voltage across a resistor is equal to the product of the resistance and the current flowing through it (at constant temperature).
- Norton's theorem: Any network of voltage and/or current sources and resistors is electrically equivalent to an ideal current source in parallel with a single resistor.
- Thévenin's theorem: Any network of voltage and/or current sources and resistors is electrically equivalent to a single voltage source in series with a single resistor.

Other more complex laws may be needed if the network contains nonlinear or reactive components. Non-linear self-regenerative heterodyning systems can be approximated. Applying these laws results in a set of simultaneous equations that can be solved either algebraically or numerically.

## *Network simulation software*

More complex circuits can be analyzed numerically with software such as SPICE or GNUCAP, or symbolically using software such as SapWin.

### Linearization around operating point

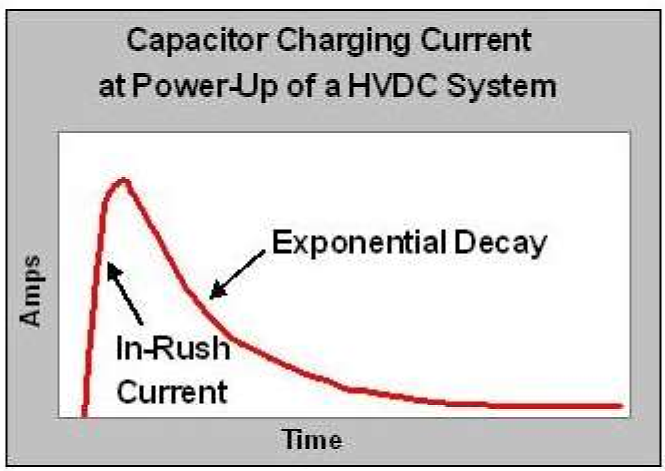
When faced with a new circuit, the software first tries to find a steady state solution, that is, one where all nodes conform to Kirchhoff's Current Law *and* the voltages across and through each element of the circuit conform to the voltage/current equations governing that element.

Once the steady state solution is found, the **operating points** of each element in the circuit are known. For a small signal analysis, every non-linear element can be linearized around its operation point to obtain the small-signal estimate of the voltages and currents. This is an application of Ohm's Law. The resulting linear circuit matrix can be solved with Gaussian elimination.

### Piecewise-linear approximation

Software such as the PLECS interface to Simulink uses piecewise-linear approximation of the equations governing the elements of a circuit. The circuit is treated as a completely linear network of ideal diodes. Every time a diode switches from on to off or vice versa, the configuration of the linear network changes. Adding more detail to the approximation of equations increases the accuracy of the simulation, but also increases its running time.

# Pre-charge




Peak inrush current into a high voltage capacitor upon power up can stress the component, reducing its reliability.

**Pre-charge** of the powerline voltages in a high voltage DC application is a preliminary mode which current-limits the power source such that a controlled rise time of the system voltage during power up is achieved.

When a high-voltage system is designed appropriately to handle the flow of maximum rated power through its distribution system, the components within the system can still undergo considerable stress upon the system "power up". In some applications, the occasion to activate the system is a rare occurrence, such as in commercial utility power distribution which is typically on almost all of the time. Yet in other systems such as in vehicle applications, activation will occur with every individual use of the system. When a long life of the components and a high reliability of the high voltage system is needed, then a power-up method which reduces and limits the power-up stress is required.

## *Background: in-rush currents into capacitors*

In-rush currents into capacitive components are a key concern in power-up stress to components. When DC input power is applied to a capacitive load, the step response of the voltage input will cause the input capacitor to charge. The capacitor charging starts with an inrush current and ends with an exponential decay down to the steady state condition. When the magnitude of the inrush peak is very large compared to the maximum rating of the components, then component stress is to be expected. The current into a capacitor is known to be $I = C(dV / dT)$: the peak inrush current will depend upon the capacitance C and the rate of change of the voltage (dV/dT). The inrush current will increase as the capacitance value increases, and the inrush current will increase as the voltage of the power source increases. This second parameter is of primary concern in high voltage power distribution systems. By their nature, high voltage power sources will deliver high voltage into the distribution system. Capacitive loads will then be subject to high inrush currents upon power-up. The stress to the components must be understood and minimized.

The objective of a pre-charge function is to limit the magnitude of the inrush current into capacitive loads during power-up. This may take several seconds depending on the system. In general, higher voltage systems benefit from longer pre-charge times during power-up.

Peak Inrush Current Into Powerline Capacitors Increases with Power-up dV/dT

| **11,000 µF Powerline Capacitor** | **Peak In-rush Current at Power-Up of a 15 A Feed** | | | |
|---|---|---|---|---|
| | 1 ms | 10 ms | 100 ms | 1 s |
| ***v* = 28 V** | 310 A | 31 A | 3.1 A | 0.31 A |
| ***v* = 610 V** | 6710 A | 671A | 67A | 7A |

Color Key:

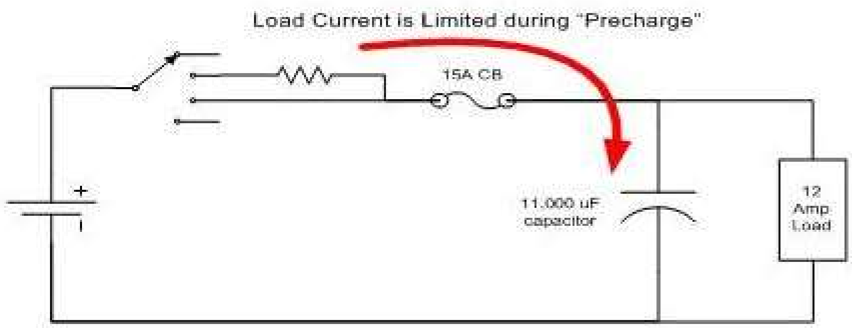
= High Risk of Tripping the Breaker

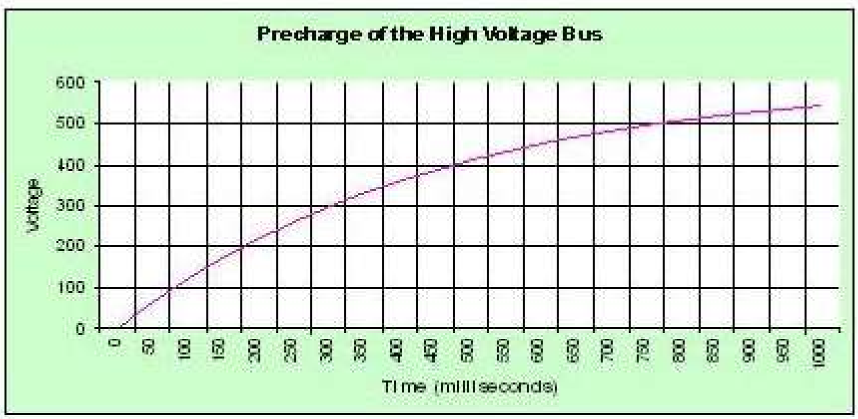
= Careful Selecting the Breaker Rating

= Good

Consider an example where a high voltage source powers up a typical electronics control unit which has an internal power supply with 11000 µF input capacitance. When powered from a 28 V source, the inrush current into the electronics unit would approach 31 amperes in 10 milliseconds. If that same circuit is activated by a 610 V source, then the inrush current would approach 670 A in 10 milliseconds. It is wise not to allow unlimited inrush currents from high voltage power distribution system activation into capacitive loads: instead the inrush current should be controlled to avoid power-up stress to components.

### *Definition of a pre-charge function*




Precharging a high voltage DC power distribution line can control the inrush current into capacitive components, reducing stress and supporting a long component life.

The functional requirement of the high voltage pre-charge circuit is to minimize the peak current out from the power source by slowing down the $dV/dT$ of the input power voltage such that a new "pre-charge mode" is created. Of course the inductive loads on the distribution system must be switched off during the precharge mode. While pre-charging, the system voltage will rise slowly and controllably with power-up current never exceeding the maximum allowed. As the circuit voltage approaches near steady state, then the pre-charge function is complete. Normal operation of a pre-charge circuit is to terminate pre-charge mode when the circuit voltage is 90% or 95% of the operating voltage. Upon completion of pre-charging, the pre-charge resistance is switched out of the power supply circuit and returns to a low impedance power source for normal mode. The high voltage loads are then powered up sequentially.

The simplest inrush-current limiting system, used in many consumer electronics devices, is a NTC resistor. When cold, its high resistance allows a small current to pre-charge the reservoir capacitor. After it warms up, its low resistance more efficiently passes the working current.

Many active power factor correction systems also include soft start.

If the example circuit from before is used with a pre-charge circuit which limits the *dV/dT* to less than 600 volts per second, then the inrush current will be reduced from 670 amperes to 7 amperes. This is a "kinder and gentler" way to activate a high voltage DC power distribution system.

### Benefits of pre-charging

The primary benefit of avoiding component stress during power-up is to realize a long system operating life due to reliable and long lasting components.

There are additional benefits: pre-charging reduces the electrical hazards which may occur when the system integrity is compromised due to hardware damage or failure. Activating the high voltage DC system into a short circuit or a ground fault or into unsuspecting personnel and their equipment can have undesired effects. Arc flash will be minimized if a pre-charge function slows down the activation time of a high voltage power-up. A slow pre-charge will also reduce the voltage into a faulty circuit which builds up while the system diagnostics come on-line. This allows a diagnostic shut down before the fault is fully realized in worst case proportions.

In cases where unlimited inrush current is large enough to trip the source circuit breaker, a slow precharge may even be required to avoid the nuisance trip.

Pre-charging is commonly used in battery electric vehicle applications. The current to the motor is regulated by a *controller* that employs large capacitors in its input circuit. Such systems typically have *contactors* (a high-current relay) to disable the system during inactive periods and to act as an emergency disconnect should the motor current regulator fail in an active state. Without pre-charge the high voltage across the contactors and inrush current can cause a brief arc which will cause pitting of the contacts. Pre-charging the controller input capacitors (typically to 90 to 95 percent of applied battery voltage) eliminates the pitting problem. The current to maintain the charge is so low that some systems apply the pre-charge at all times other than when charging batteries, while more complex systems apply pre-charge as part of the starting sequence and will defer main contactor closure until the pre-charge voltage level is detected as sufficiently high.

### Applications in high voltage power systems

- High-voltage direct current
- Battery Electric Vehicles
- Hybrid Vehicle

- Future Combat System
- Motorized bicycle
- Electric power-assist system