Electrical Resistive Components

Lesley Redden

First Edition, 2012

ISBN 978-81-323-4359-2

Published by:
White Word Publications
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter 1

Electrical Resistance and Conductance

The **electrical resistance** of an electrical element measures its opposition to the passage of an electric current; the inverse quantity is **electrical conductance**, measuring how easily electricity flows along a certain path. Electrical resistance shares some conceptual parallels with the mechanical notion of friction. The SI unit of electrical resistance is the ohm (Ω), while electrical conductance is measured in siemens (S).

An object of uniform cross section has a resistance proportional to its resistivity and length and inversely proportional to its cross-sectional area. All materials show some resistance, except for superconductors, which have a resistance of zero.

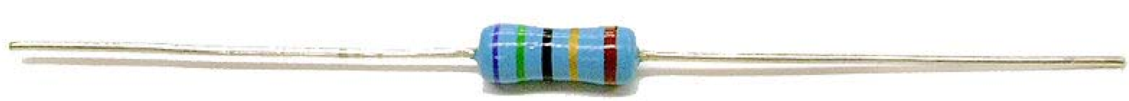
The resistance of an object is defined as the ratio of voltage across it to current through it:

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

For a wide variety of materials and conditions, the electrical resistance R is constant for a given temperature; it does not depend on the amount of current through or the potential difference (voltage) across the object. Such materials are called Ohmic materials. For objects made of ohmic materials the definition of the resistance, with R being a constant for that resistor, is known as Ohm's law.

In the case of a nonlinear conductor (not obeying Ohm's law), this ratio can change as current or voltage changes; the inverse slope of a chord to an I–V curve is sometimes referred to as a "chordal resistance" or "static resistance".

Conductors and resistors



A 65-Ω resistor, as identified by its electronic color code (blue–green–black). An ohmmeter could be used to verify this value.

Objects such as wires that are designed to have low resistance so that they transfer current with the least loss of electrical energy are called conductors. Objects that are designed to have a specific resistance so that they can dissipate electrical energy or otherwise modify how a circuit behaves are called resistors. Conductors are made of highly conductive materials such as metals, in particular copper and aluminium. Resistors, on the other hand, are made of a wide variety of materials depending on factors such as the desired resistance, amount of energy that it needs to dissipate, precision, and cost.

DC resistance

The resistance of a given resistor or conductor grows with the length of conductor and decreases for larger cross-sectional area. The resistance R and conductance G of a conductor of uniform cross section, therefore, can be computed as

$$R = \rho \frac{\ell}{A},$$
$$G = \frac{\sigma A}{\ell}.$$

where ℓ is the length of the conductor, measured in metres [m], *A* is the cross-sectional area of the conductor measured in square metres [m²], and ρ (rho) is the electrical resistivity (also called *specific electrical resistance*) of the material, measured in ohm-metres (Ωm). Resistivity is a measure of the material's ability to oppose electric current. For purely resistive circuits conductance is related to resistance *R* by:

$$G = \frac{1}{R}$$

For practical reasons, any connections to a real conductor will almost certainly mean the current density is not totally uniform. However, this formula still provides a good approximation for long thin conductors such as wires.

AC resistance

A wire carrying alternating current has a reduced effective cross sectional area because of the skin effect. Adjacent conductors carryng alternating current have a higher resistance than they would in isolation or when carrying direct current, due to the proximity effect. At commercial power frequency, these effects are significant for large conductors carrying large currents, such as busbars in an electrical substation, or large power cables carrying more than a few hundred amperes.

When an alternating current flows through the circuit, its flow is not opposed only by the circuit resistance, but also by the opposition of electric and magnetic fields to the current change. That effect is measured by electrical reactance. The combined effects of reactance and resistance are expressed by electrical impedance.

Measuring resistance

An instrument for measuring resistance is called an ohmmeter. Simple ohmmeters cannot measure low resistances accurately because the resistance of their measuring leads causes a voltage drop that interferes with the measurement, so more accurate devices use four-terminal sensing.

Causes of resistance

In metals

A metal consists of a lattice of atoms, each with a shell of electrons. This is also known as a positive ionic lattice. The outer electrons are free to dissociate from their parent atoms and travel through the lattice, creating a 'sea' of electrons, making the metal a conductor. When an electrical potential difference (a voltage) is applied across the metal, the electrons drift from one end of the conductor to the other under the influence of the electric field.

Near room temperatures, the thermal motion of ions is the primary source of scattering of electrons (due to destructive interference of free electron waves on non-correlating potentials of ions), and is thus the prime cause of metal resistance. Imperfections of lattice also contribute into resistance, although their contribution in pure metals is negligible.

The larger the cross-sectional area of the conductor, the more electrons are available to carry the current, so the lower the resistance. The longer the conductor, the more scattering events occur in each electron's path through the material, so the higher the resistance. Different materials also affect the resistance.

In semiconductors and insulators

In metals, the Fermi level lies in the conduction band giving rise to free conduction electrons. However, in semiconductors the position of the Fermi level is within the band gap, approximately half-way between the conduction band minimum and valence band maximum for intrinsic (undoped) semiconductors. This means that at 0 kelvins, there are no free conduction electrons and the resistance is infinite. However, the resistance will continue to decrease as the charge carrier density in the conduction band increases. In extrinsic (doped) semiconductors, dopant atoms increase the majority charge carrier concentration by donating electrons to the conduction band or accepting holes in the valence band. For both types of donor or acceptor atoms, increasing the dopant density leads to a reduction in the resistance. Highly doped semiconductors hence behave metallic. At very high temperatures, the contribution of thermally generated carriers will dominate over the contribution from dopant atoms and the resistance will decrease exponentially with temperature.

Small-signal device conductances

The term conductance applies to electronic devices such as transistors and diodes, where it usually refers to a small-signal model that is a linearization of the underlying device equations about a selected DC operating point or Q-point. This conductance is the reciprocal of the small-signal device resistance.

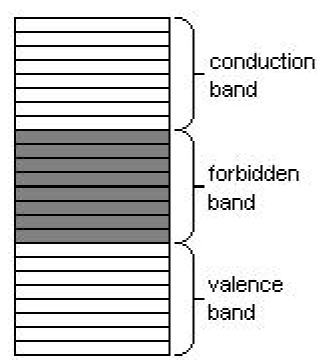
In ionic liquids/electrolytes

In electrolytes, electrical conduction happens not by band electrons or holes, but by full atomic species (ions) traveling, each carrying an electrical charge. The resistivity of ionic liquids varies tremendously by the concentration - while distilled water is almost an insulator, salt water is a very efficient electrical conductor. In biological membranes, currents are carried by ionic salts. Small holes in the membranes, called ion channels, are selective to specific ions and determine the membrane resistance.

Resistivity of various materials

Material	**Resistivity**, ρ ohm-metre
Metals	10^{-8}
Semiconductors	variable
Electrolytes	variable
Insulators	10^{16}
Superconductors	0 (exactly)

Band theory simplified



Electron energy levels in an insulator

Quantum mechanics states that the energy of an electron in an atom cannot be any arbitrary value. Rather, there are fixed energy levels which the electrons can occupy, and values in between these levels are impossible. The energy levels are grouped into two bands: the **valence band** and the **conduction band** (the latter is generally above the former). Electrons in the conduction band may move freely throughout the substance in the presence of an electrical field.

In insulators and semiconductors, the atoms in the substance influence each other so that between the valence band and the conduction band there exists a forbidden band of energy levels, which the electrons cannot occupy. In order for a current to flow, a relatively large amount of energy must be furnished to an electron for it to leap across this forbidden gap and into the conduction band. Thus, even large voltages can yield relatively small currents.

Differential resistance

When the current–voltage dependence is not linear, **differential resistance**, **incremental resistance** or **slope resistance** is defined as the slope of the *V-I* graph at a particular point, thus:

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

This quantity is sometimes called simply *resistance*, although the two definitions are equivalent only for an ohmic component such as an ideal resistor. For example, a diode is a circuit element for which the resistance depends on the applied voltage or current.

If the *V-I* graph is not monotonic (i.e. it has a peak or a trough), the differential resistance will be negative for some values of voltage and current. This property is often known as *negative resistance*, although it is more correctly called *negative differential resistance*, since the absolute resistance V/I is still positive. An example of such an element is the tunnel diode.

Differential resistance is only useful to compare a nonlinear device with a linear source/load in some small interval; for example if it is necessary to evaluate a zener diode's voltage stability under different current values.

Temperature dependence

Near room temperature, the electric resistance of a typical metal increases linearly with rising temperature, while the electrical resistance of a typical semiconductor decreases with rising temperature. The amount of that change in resistance can be calculated using the temperature coefficient of resistivity of the material using the following formula:

$$R(T) = R_0[1 + \alpha(T - T_0)]$$

where T is its temperature, T_0 is a reference temperature (usually room temperature), R_0 is the resistance at T_0, and α is the percentage change in resistivity per unit temperature. The constant α depends only on the material being considered. The relationship stated is actually only an approximate one, the true physics being somewhat non-linear, or looking at it another way, α itself varies with temperature. For this reason it is usual to specify the temperature that α was measured at with a suffix, such as α_{15} and the relationship only holds in a range of temperatures around the reference.

At lower temperatures (less than the Debye temperature), the resistance of a metal decreases as T^5 due to the electrons scattering off of phonons. At even lower temperatures, the dominant scattering mechanism for electrons is other electrons, and the resistance decreases as T^2. At some point, the impurities in the metal will dominate the behavior of the electrical resistance which causes it to saturate to a constant value. Matthiessen's Rule (first formulated by Augustus

Matthiessen in the 1860s; the equation below gives its modern form) says that all of these different behaviors can be summed up to get the total resistance as a function of temperature,

$$R = R_{\text{imp}} + aT^2 + bT^5 + cT$$

where R_{imp} is the temperature independent electrical resistivity due to impurities, and a, b, and c are coefficients which depend upon the metal's properties. This rule can be seen as the motivation to Heike Kamerlingh Onnes's experiments that led in 1911 to discovery of superconductivity.

Intrinsic semiconductors become better conductors as the temperature increases; the electrons are bumped to the conduction energy band by thermal energy, where they flow freely and in doing so leave behind holes in the valence band which also flow freely. The electric resistance of a typical intrinsic (non doped) semiconductor decreases exponentially with the temperature:

$$R = R_0 e^{-aT}$$

Extrinsic (doped) semiconductors have a far more complicated temperature profile. As temperature increases starting from absolute zero they first decrease steeply in resistance as the carriers leave the donors or acceptors. After most of the donors or acceptors have lost their carriers the resistance starts to increase again slightly due to the reducing mobility of carriers (much as in a metal). At higher temperatures it will behave like intrinsic semiconductors as the carriers from the donors/acceptors become insignificant compared to the thermally generated carriers.

The electric resistance of electrolytes and insulators is highly nonlinear, and case by case dependent, therefore no generalized equations are given.

Strain dependence

Just as the resistance of a conductor depends upon temperature, the resistance of a conductor depends upon strain. By placing a conductor under tension (a form of stress that leads to strain in the form of stretching of the conductor), the length of the section of conductor under tension increases and its cross-sectional area decreases. Both these effects contribute to increasing the resistance of the strained section of conductor. Under compression (strain in the opposite direction), the resistance of the strained section of conductor decreases.

Chapter 2

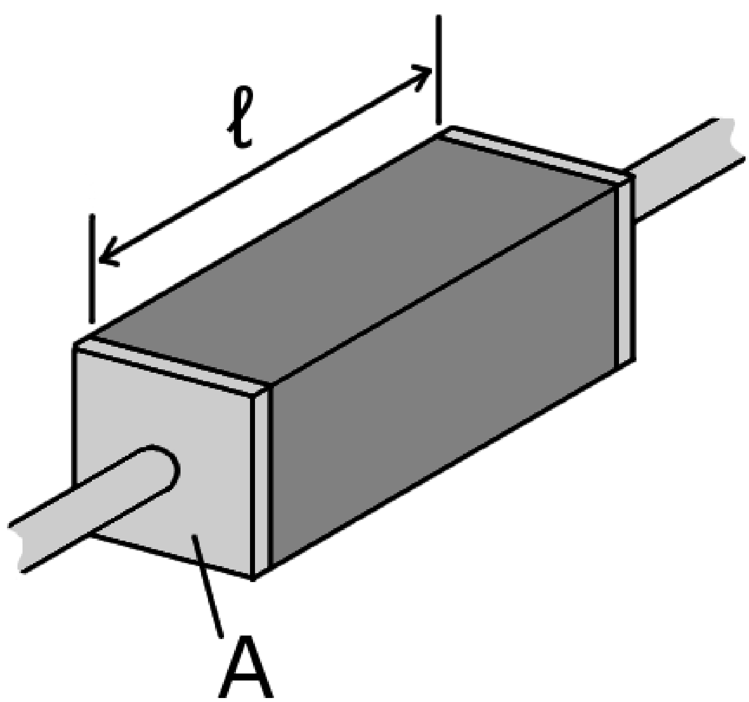
Electrical Resistivity and Conductivity

Electrical resistivity (also known as **resistivity**, **specific electrical resistance**, or **volume resistivity**) is a measure of how strongly a material opposes the flow of electric current. A low resistivity indicates a material that readily allows the movement of electric charge. The SI unit of electrical resistivity is the ohm metre [Ωm]. It is commonly represented by the Greek letter ρ (rho).

Electrical conductivity or **specific conductance** is the reciprocal quantity, and measures a material's ability to conduct an electric current. It is commonly represented by the Greek letter σ, but κ (esp. in electrical engineering) or γ are also occasionally used. Its SI unit is siemens per metre ($S \cdot m^{-1}$) and CGSE unit is reciprocal second (s^{-1}):

$$\sigma = \frac{1}{\rho}.$$

Definitions



A piece of resistive material with electrical contacts on both ends.

Electrical resistivity ρ (Greek: rho) is defined by,

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

where

ρ is the static resistivity (measured in ohm-metres, Ω-m)
E is the magnitude of the electric field (measured in volts per metre, V/m);
J is the magnitude of the current density (measured in amperes per square metre, A/m^2).

Many resistors and conductors have a uniform cross section with a uniform flow of electric current and are made of one material. In this case, the above definition of ρ leads to:

$$\rho = R\frac{A}{\ell},$$

where

R is the electrical resistance of a uniform specimen of the material (measured in ohms, Ω)
ℓ is the length of the piece of material (measured in metres, m)
A is the cross-sectional area of the specimen (measured in square metres, m^2).

The reason resistivity has the dimension units of ohm-metres can be seen by transposing the definition to make resistance the subject:

$$R = \rho\frac{\ell}{A}$$

The resistance of a given sample will increase with the length, but decrease with greater cross-sectional area. Resistance is measured in ohms. Length over area has units of 1/distance. To end up with ohms, resistivity must be in the units of "ohms × distance" (SI ohm-metre, US ohm-inch).

In a hydraulic analogy, increasing the cross-sectional area of a pipe reduces its resistance to flow, and increasing the length increases resistance to flow (and pressure drop for a given flow).

Resistivity of various materials

- A conductor such as a metal has high conductivity and a low resistivity.
- An insulator like glass has low conductivity and a high resistivity.
- The conductivity of a semiconductor is generally intermediate, but varies widely under different conditions, such as exposure of the material to electric fields or specific frequencies of light, and, most important, with temperature and composition of the semiconductor material.

The degree of doping in semiconductors makes a large difference in conductivity. To a point, more doping leads to higher conductivity. The conductivity of a solution of water is highly dependent on its concentration of dissolved salts, and other chemical species that ionize in the solution. Electrical conductivity of water samples is used as an indicator of how salt-free, ion-free, or impurity-free the sample is; the purer the water, the lower the conductivity (the higher the resistivity). Conductivity measurements in water are often reported as *specific conductance*, the conductivity of the water at 25 °C. An EC meter is normally used to measure conductivity in a solution.

This table shows the resistivity, conductivity and temperature coefficient of various materials at 20 °C (68 °F)

Material	ρ [Ω·m] at 20 °C	σ [S/m] at 20 °C	Temperature coefficient [K^{-1}]
Silver	1.59×10^{-8}	6.30×10^{7}	0.0038
Copper	1.68×10^{-8}	5.96×10^{7}	0.0039
Annealed Copper		5.80×10^{7}	
Gold	2.44×10^{-8}	4.52×10^{7}	0.0034
Aluminium	2.82×10^{-8}	3.5×10^{7}	0.0039
Calcium	$3.36x10^{-8}$		0.0041
Tungsten	5.60×10^{-8}		0.0045
Zinc	5.90×10^{-8}		0.0037
Nickel	6.99×10^{-8}		0.006
Iron	1.0×10^{-7}		0.005
Platinum	1.06×10^{-7}		0.00392
Tin	1.09×10^{-7}		0.0045
Lead	2.2×10^{-7}		0.0039
Titanium	$4.20x10^{-7}$		X
Manganin	4.82×10^{-7}		0.000002
Constantan	4.9×10^{-7}		0.000008
Mercury	9.8×10^{-7}		0.0009
Nichrome	1.10×10^{-6}		0.0004
Carbon (amorphous)	$5\text{-}8\times10^{-4}$		−0.0005
Carbon (graphite)	$2.5\text{-}5.0\times10^{-6}$ ⊥ basal plane 3.0×10^{-3} // basal plane		
Carbon (diamond)	$\sim10^{12}$		
Germanium	4.6×10^{-1}		−0.048

Sea water	2×10^{-1}	4.8	
Drinking water		0.0005 to 0.05	
Deionized water		5.5×10^{-6}	
Silicon	6.40×10^{2}		−0.075
Glass	10^{10} to 10^{14}		?
Hard rubber	approx. 10^{13}		?
Sulfur	10^{15}		?
Air		3 to 8×10^{-15}	
Paraffin	10^{17}		?
Quartz (fused)	7.5×10^{17}		?
PET	10^{20}		?
Teflon	10^{22} to 10^{24}		?

The effective temperature coefficient varies with temperature and purity level of the material. The 20 °C value is only an approximation when used at other temperatures. For example, the coefficient becomes lower at higher temperatures for copper, and the value 0.00427 is commonly specified at 0 °C.

The extremely low resistivity (high conductivity) of silver is characteristic of metals. George Gamow tidily summed up the nature of the metals' dealings with electrons 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." More technically, the free electron model gives a basic description of electron flow in metals.

Temperature dependence

In general, electrical resistivity of metals increases with temperature, while the resistivity of semiconductors decreases with increasing temperature. In both cases, electron–phonon interactions can play a key role. At high temperatures, the resistance of a metal increases linearly with temperature. As the temperature of a metal is reduced, the temperature dependence of resistivity follows a power law function of temperature. Mathematically the temperature dependence of the resistivity ρ of a metal is given by the Bloch–Grüneisen formula:

$$\rho(T) = \rho(0) + A\left(\frac{T}{\Theta_R}\right)^n \int_0^{\frac{\Theta_R}{T}} \frac{x^n}{(e^x - 1)(1 - e^{-x})} dx$$

where ρ(0) is the residual resistivity due to defect scattering, A is a constant that depends on the velocity of electrons at the Fermi surface, the Debye radius and the number density of electrons in the metal. Θ_R is the Debye temperature as obtained from resistivity measurements and matches very closely with the values of Debye temperature obtained from specific heat measurements. n is an integer that depends upon the nature of interaction:

1. n=5 implies that the resistance is due to scattering of electrons by phonons (as it is for simple metals)
2. n=3 implies that the resistance is due to s-d electron scattering (as is the case for transition metals)
3. n=2 implies that the resistance is due to electron–electron interaction.

As the temperature of the metal is sufficiently reduced (so as to 'freeze' all the phonons), the resistivity usually reaches a constant value, known as the **residual resistivity**. This value depends not only on the type of metal, but on its purity and thermal history. The value of the residual resistivity of a metal is decided by its impurity concentration. Some materials lose all electrical resistivity at sufficiently low temperatures, due to an effect known as superconductivity.

An even better approximation of the temperature dependence of the resistivity of a semiconductor is given by the Steinhart–Hart equation:

$$1/T = A + B\ln(\rho) + C(\ln(\rho))^3$$

where *A*, *B* and *C* are the so-called **Steinhart–Hart coefficients**.

This equation is used to calibrate thermistors.

In non-crystalline semi-conductors, conduction can occur by charges quantum tunnelling from one localised site to another. This is known as variable range hopping and has the characteristic form of $\rho = Ae^{T^{-1/n}}$, where n=2,3,4 depending on the dimensionality of the system.

Complex resistivity and conductivity

When analyzing the response of materials to alternating electric fields, in applications such as electrical impedance tomography, it is necessary to replace resistivity with a complex quantity called **impeditivity** (in analogy to electrical impedance). Impeditivity is the sum of a real component, the resistivity, and an imaginary component, the **reactivity** (in analogy to reactance). The magnitude of Impeditivity is the square root of sum of squares of magnitudes of resistivity and reactivity.

Conversely, in such cases the conductivity must be expressed as a complex number (or even as a matrix of complex numbers, in the case of anisotropic materials) called the *admittivity*. Admittivity is the sum of a real component called the conductivity and an imaginary component called the susceptivity.

An alternative description of the response to alternating currents uses a real (but frequency-dependent) conductivity, along with a real permittivity. The larger the conductivity is, the more quickly the alternating-current signal is absorbed by the material (i.e., the more opaque the material is).

Resistivity density products

In some applications where the weight of an item is very important resistivity density products are more important than absolute low resistivity- it is often possible to make the conductor thicker to make up for a higher resistivity; and then a low resistivity density product material (or equivalently a high conductance to density ratio) is desirable. For example, for long distance overhead power lines— aluminium is frequently used rather than copper because it is lighter for the same conductance.

Material	Resistivity [nΩ·m]	Density [g/cm³]	Resistivity-density product [nΩ·m·g/cm³]
Sodium	47.7	0.97	46
Lithium	92.8	0.53	49
Calcium	33.6	1.55	52
Potassium	72.0	0.89	64
Aluminium	26.50	2.70	72
Copper	16.78	8.96	150
Silver	15.87	10.49	166

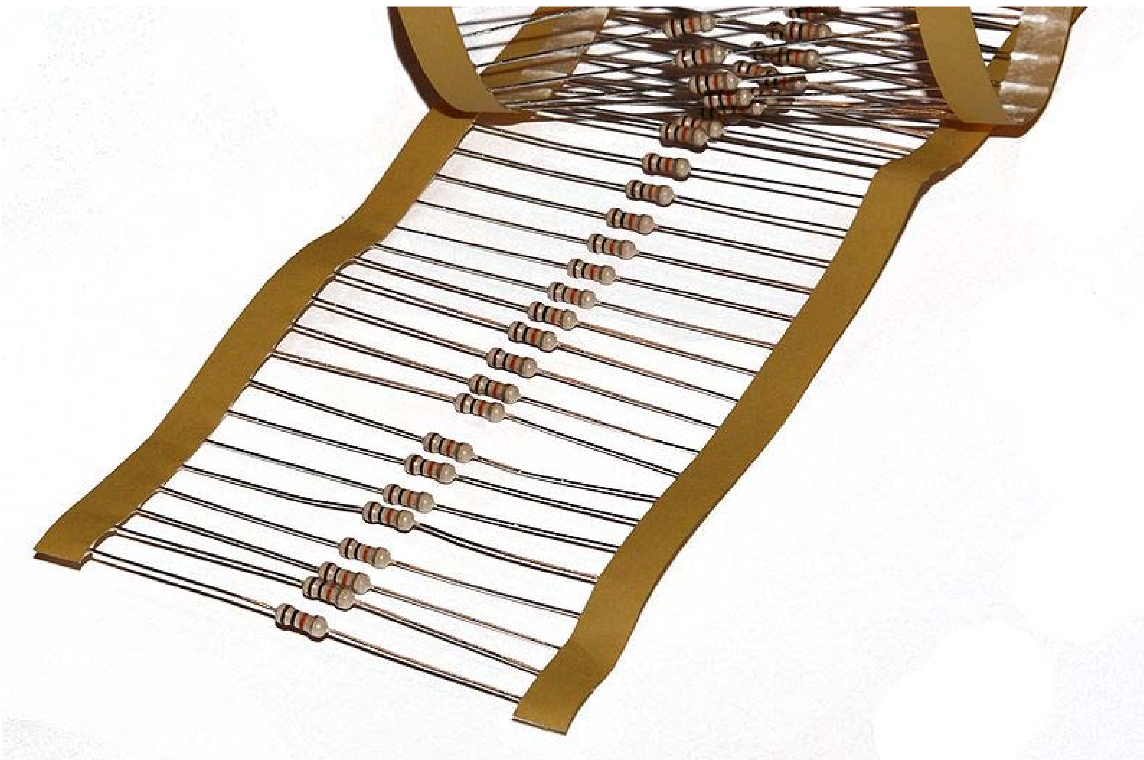
Silver, although it is the least resistive metal known, has a high density and does poorly by this measure. The calcium and the alkali metals have the best products, but are rarely used for conductors due to their high reactivity with water and oxygen. Aluminium is far more stable.

Chapter 3

Resistor

A typical axial-lead resistor

Partially exposed Tesla TR-212 1 kΩ carbon film resistor



Axial-lead resistors on tape. The tape is removed during assembly before the leads are formed and the part is inserted into the board.

Three carbon composition resistors in a 1960s valve (vacuum tube) radio

A **resistor** is a two-terminal passive electronic component which implements electrical resistance as a circuit element. When a voltage V is applied across the terminals of a resistor, a current I will flow through the resistor in direct proportion to that voltage. The reciprocal of the constant of proportionality is known as the resistance R, since, with a given voltage V, a larger value of R further "resists" the flow of current I as given by Ohm's law:

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

Resistors are common elements of electrical networks and electronic circuits and are ubiquitous in most electronic equipment. Practical resistors can be made of various compounds and films, as well as resistance wire (wire made of a high-resistivity alloy, such as nickel-chrome). Resistors are also implemented within integrated circuits, particularly analog devices, and can also be integrated into hybrid and printed circuits.

The electrical functionality of a resistor is specified by its resistance: common commercial resistors are manufactured over a range of more than 9 orders of magnitude. When specifying that resistance in an electronic design, the required precision of the resistance may require attention to the manufacturing tolerance of the chosen resistor, according to its specific application. The temperature coefficient of the resistance may also be of concern in some precision applications. Practical resistors are also specified as having a maximum power rating which must exceed the anticipated power dissipation of that resistor in a particular circuit: this is mainly of concern in power electronics applications. Resistors with higher power ratings are physically larger and may require heat sinking. In a high voltage circuit, attention must sometimes be paid to the rated maximum working voltage of the resistor.

The series inductance of a practical resistor causes its behavior to depart from ohms law; this specification can be important in some high-frequency applications for smaller values of resistance. In a low-noise amplifier or pre-amp the noise characteristics of a resistor may be an issue. The unwanted inductance, excess noise, and temperature coefficient are mainly dependent on the technology used in manufacturing the resistor. They are not normally specified individually for a particular family of resistors manufactured using a particular technology. A family of discrete resistors is also characterized according to its form factor, that is, the size of the device and position of its leads (or terminals) which is relevant in the practical manufacturing of circuits using them.

Units

The ohm (symbol: Ω) is the SI unit of electrical resistance, named after Georg Simon Ohm. An ohm is equivalent to a volt per ampere. Since resistors are specified and manufactured over a very large range of values, the derived units of milliohm (1 mΩ = 10^{-3} Ω), kilohm (1 kΩ = 10^{3} Ω), and megohm (1 MΩ = 10^{6} Ω) are also in common usage.

The reciprocal of resistance R is called conductance G = 1/R and is measured in Siemens (SI unit), sometimes referred to as a mho. Thus a Siemens is the reciprocal of an ohm: $S = \Omega^{-1}$. Although the concept of conductance is often used in circuit analysis, practical resistors are always specified in terms of their resistance (ohms) rather than conductance.

Theory of operation

Ohm's law

The behavior of an ideal resistor is dictated by the relationship specified in Ohm's law:

$$V = I \cdot R$$

Ohm's law states that the voltage (V) across a resistor is proportional to the current (I) passing through it, where the constant of proportionality is the resistance (R).

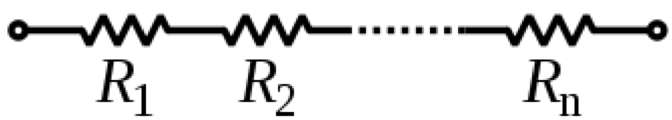
Equivalently, Ohm's law can be stated:

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

This formulation of Ohm's law states that, when a voltage (V) is present across a resistance (R), a current (I) will flow through the resistance. This is directly used in practical computations. For example, if a 300 ohm resistor is attached across the terminals of a 12 volt battery, then a current of 12 / 300 = 0.04 amperes (or 40 milliamperes) will flow through that resistor.

Series and parallel resistors

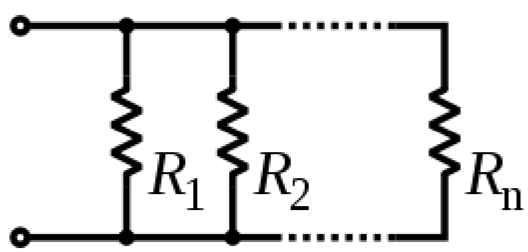
In a series configuration, the current through all of the resistors is the same, but the voltage across each resistor will be in proportion to its resistance. The potential difference (voltage) seen across the network is the sum of those voltages, thus the total resistance can be found as the sum of those resistances:



$$R_{\mathrm{eq}} = R_1 + R_2 + \cdots + R_n$$

As a special case, the resistance of N resistors connected in series, each of the same resistance R, is given by NR.

Resistors in a parallel configuration are each subject to the same potential difference (voltage), however the currents through them add. The conductances of the resistors then add to determine the conductance of the network. Thus the equivalent resistance (R_{eq}) of the network can be computed:



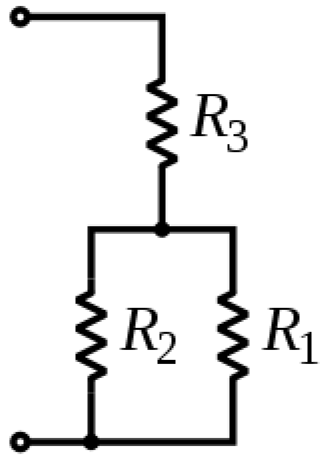
$$\frac{1}{R_{\mathrm{eq}}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}$$

The parallel equivalent resistance can be represented in equations by two vertical lines "||" (as in geometry) as a simplified notation. For the case of two resistors in parallel, this can be calculated using:

$$R_{\mathrm{eq}} = R_1 \| R_2 = \frac{R_1 R_2}{R_1 + R_2}$$

As a special case, the resistance of N resistors connected in parallel, each of the same resistance R, is given by R/N.

A resistor network that is a combination of parallel and series connections can be broken up into smaller parts that are either one or the other. For instance,



$$R_{\mathrm{eq}} = (R_1 \| R_2) + R_3 = \frac{R_1 R_2}{R_1 + R_2} + R_3$$

However, some complex networks of resistors cannot be resolved in this manner, requiring more sophisticated circuit analysis. For instance, consider a cube, each edge of which has been replaced by a resistor. What then is the resistance that would be measured between two opposite vertices? In the case of 12 equivalent resistors, it can be shown that the corner-to-corner resistance is $\frac{5}{6}$ of the individual resistance. More generally, the Y-Δ transform, or matrix methods can be used to solve such a problem.

One practical application of these relationships is that a non-standard value of resistance can generally be synthesized by connecting a number of standard values in series and/or parallel. This can also be used to obtain a resistance with a higher power rating than that of the individual resistors used. In the special case of N identical resistors all connected in series or all connected in parallel, the power rating of the individual resistors is thereby multiplied by N.

Power dissipation

The power P dissipated by a resistor (or the equivalent resistance of a resistor network) is calculated as: $P = I^2 R = IV = \frac{V^2}{R}$

The first form is a restatement of Joule's first law. Using Ohm's law, the two other forms can be derived.

The total amount of heat energy released over a period of time can be determined from the integral of the power over that period of time:

$$W = \int_{t_1}^{t_2} v(t)i(t)\,dt.$$

Practical resistors are rated according to their maximum power dissipation. The vast majority of resistors used in electronic circuits absorb much less than a watt of electrical power and require no attention to their power rating. Such resistors in their discrete form, including most of the packages detailed below, are typically rated as 1/10, 1/8, or 1/4 watt.

Resistors required to dissipate substantial amounts of power, particularly used in power supplies, power conversion circuits, and power amplifiers, are generally referred to as *power resistors*; this designation is loosely applied to resistors with power ratings of 1 watt or greater. Power resistors are physically larger and tend not to use the preferred values, color codes, and external packages described below.

If the average power dissipated by a resistor is more than its power rating, damage to the resistor may occur, permanently altering its resistance; this is distinct from the reversible change in resistance due to its temperature coefficient when it warms. Excessive power dissipation may raise the temperature of the resistor to a point where it can burn the circuit board or adjacent components, or even cause a fire. There are flameproof resistors that fail (open circuit) before they overheat dangerously.

Note that the nominal power rating of a resistor is not the same as the power that it can safely dissipate in practical use. Air circulation and proximity to a circuit board, ambient temperature, and other factors can reduce acceptable dissipation significantly. Rated power dissipation may be given for an ambient temperature of 25 °C in free air. Inside an equipment case at 60 °C, rated dissipation will be significantly less; a resistor dissipating a bit less than the maximum figure given by the manufacturer may still be outside the safe operating area and may prematurely fail.

Construction

A single in line (SIL) resistor package with 8 individual, 47 ohm resistors. One end of each resistor is connected to a separate pin and the other ends are all connected together to the remaining (common) pin - pin 1, at the end identified by the white dot.

Lead arrangements

Resistors with wire leads for through-hole mounting

Through-hole components typically have leads leaving the body axially. Others have leads coming off their body radially instead of parallel to the resistor axis. Other components may be SMT (surface mount technology) while high power resistors may have one of their leads designed into the heat sink.

Carbon composition

Carbon composition resistors consist of a solid cylindrical resistive element with embedded wire leads or metal end caps to which the lead wires are attached. The body of the resistor is protected with paint or plastic. Early 20th-century carbon composition resistors had uninsulated bodies; the lead wires were wrapped around the ends of the resistance element rod and soldered. The completed resistor was painted for color coding of its value.

The resistive element is made from a mixture of finely ground (powdered) carbon and an insulating material (usually ceramic). A resin holds the mixture together. The resistance is determined by the ratio of the fill material (the powdered ceramic) to the carbon. Higher concentrations of carbon, a weak conductor, result in lower resistance. Carbon composition resistors were commonly used in the 1960s and earlier, but are not so popular for general use now as other types have better specifications, such as tolerance, voltage dependence, and stress (carbon composition resistors will change value when stressed with over-voltages). Moreover, if internal moisture content (from exposure for some length of time to a humid environment) is significant, soldering heat will create a non-reversible change in resistance value. Carbon composition resistors have poor stability with time and were consequently factory sorted to, at best, only 5% tolerance. These resistors, however, if never subjected to overvoltage nor overheating were remarkably reliable considering the component's size

They are still available, but comparatively quite costly. Values ranged from fractions of an ohm to 22 megohms. Because of the high price, these resistors are no longer used in most applications. However, carbon resistors are used in power supplies and welding controls.

Carbon film

A carbon film is deposited on an insulating substrate, and a helix cut in it to create a long, narrow resistive path. Varying shapes, coupled with the resistivity of carbon, (ranging from 90 to 400 nΩ m) can provide a variety of resistances. Carbon film resistors feature a power rating range of 0.125 W to 5 W at 70 °C. Resistances available range from 1 ohm to 10 megohm. The carbon film resistor has an operating temperature range of -55 °C to 155 °C. It has 200 to 600 volts maximum working voltage range. Special carbon film resistors are used in applications requiring high pulse stability.

Thick and thin film

Thick film resistors became popular during the 1970s, and most SMD (surface mount device) resistors today are of this type. The principal difference between thin film and thick film resistors is not the actual thickness of the film, but rather how the film is applied to the cylinder (axial resistors) or the surface (SMD resistors).

Thin film resistors are made by sputtering (a method of vacuum deposition) the resistive material onto an insulating substrate. The film is then etched in a similar manner to the old (subtractive) process for making printed circuit boards; that is, the surface is coated with a photo-sensitive material, then covered by a pattern film, irradiated with ultraviolet light, and then the exposed photo-sensitive coating is developed, and underlying thin film is etched away.

Thick film resistors are manufactured using screen and stencil printing processes.

Because the time during which the sputtering is performed can be controlled, the thickness of the thin film can be accurately controlled. The type of material is also usually different consisting of one or more ceramic (cermet) conductors such as tantalum nitride (TaN), ruthenium dioxide (RuO_2), lead oxide (PbO), bismuth ruthenate ($Bi_2Ru_2O_7$), nickel chromium (NiCr), and/or bismuth iridate ($Bi_2Ir_2O_7$).

The resistance of both thin and thick film resistors after manufacture is not highly accurate; they are usually trimmed to an accurate value by abrasive or laser trimming. Thin film resistors are usually specified with tolerances of 0.1, 0.2, 0.5, or 1%, and with temperature coefficients of 5 to 25 ppm/K.

Thick film resistors may use the same conductive ceramics, but they are mixed with sintered (powdered) glass and some kind of liquid so that the composite can be screen-printed. This composite of glass and conductive ceramic (cermet) material is then fused (baked) in an oven at about 850 °C.

Thick film resistors, when first manufactured, had tolerances of 5%, but standard tolerances have improved to 2% or 1% in the last few decades. Temperature coefficients of thick film resistors are high, typically ±200 or ±250 ppm/K; a 40 kelvin (70 °F) temperature change can change the resistance by 1%.

Thin film resistors are usually far more expensive than thick film resistors. For example, SMD thin film resistors, with 0.5% tolerances, and with 25 ppm/K temperature coefficients, when bought in full size reel quantities, are about twice the cost of 1%, 250 ppm/K thick film resistors.

Metal film

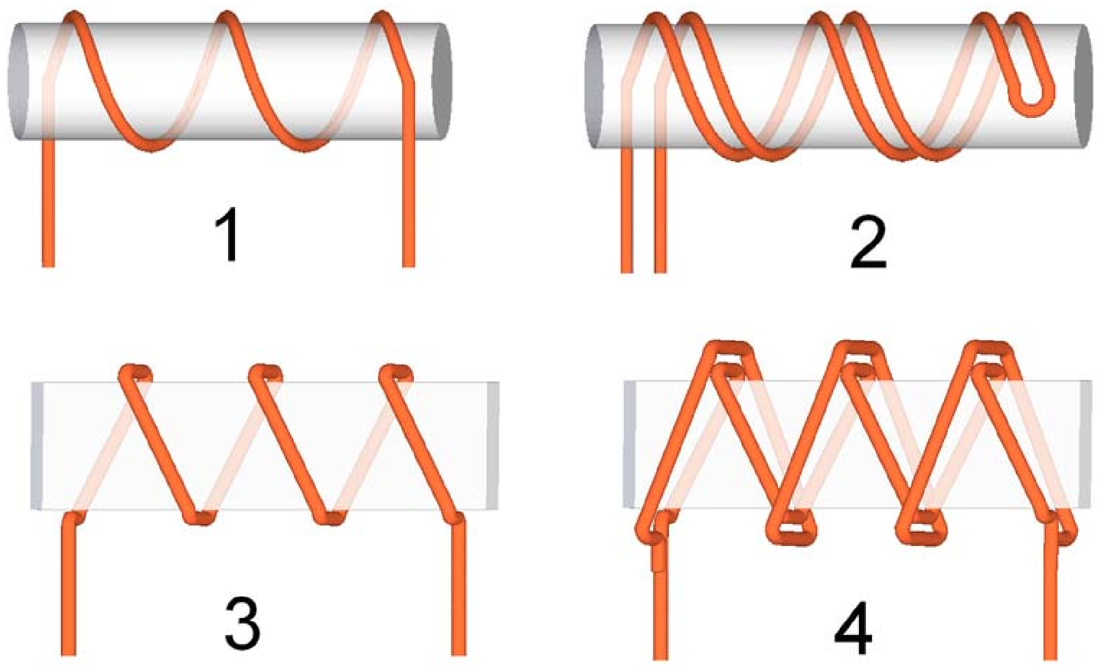
A common type of axial resistor today is referred to as a metal-film resistor. Metal electrode leadless face (MELF) resistors often use the same technology, but are a cylindrically shaped resistor designed for surface mounting. Note that other types of resistors (e.g., carbon composition) are also available in MELF packages.

Metal film resistors are usually coated with nickel chromium (NiCr), but might be coated with any of the cermet materials listed above for thin film resistors. Unlike thin film resistors, the material may be applied using different techniques than sputtering (though that is one such technique). Also, unlike thin-film resistors, the resistance value is determined by cutting a helix through the coating rather than by etching. (This is similar to the way carbon resistors are made.) The result is a reasonable tolerance (0.5, 1, or 2%) and a temperature coefficient that is generally between 50 and 100 ppm/K. Metal film resistors possess good noise characteristics and low non-linearity due to a low voltage coefficient. Also beneficial are the components efficient tolerance, temperature coefficient and stability.

Metal Oxide film

Metal-Oxide film resistors resemble Metal film types, but are made of metal oxides such as tin oxide. This results in a higher operating temperature and greater stability/reliability than Metal film. They are used in applications with high endurance demands.

Wirewound



Types of windings in wire resistors:
1 - common
2 - bifilar
3 - common on a thin former
4 - Ayrton-Perry

Wirewound resistors are commonly made by winding a metal wire, usually nichrome, around a ceramic, plastic, or fiberglass core. The ends of the wire are soldered or welded to two caps or rings, attached to the ends of the core. The assembly is protected with a layer of paint, molded plastic, or an enamel coating baked at high temperature. Because of the very high surface temperature these resistors can withstand temperatures of up to +450 °C. Wire leads in low power wirewound resistors are usually between 0.6 and 0.8 mm in diameter and tinned for ease of soldering. For higher power wirewound resistors, either a ceramic outer case or an aluminum outer case on top of an insulating layer is used. The aluminum-cased types are designed to be attached to a heat sink to dissipate the heat; the rated power is dependent on being used with a suitable heat sink, e.g., a 50 W power rated resistor will overheat at a fraction of the power dissipation if not used with a heat sink. Large wirewound resistors may be rated for 1,000 watts or more.

Because wirewound resistors are coils they have more undesirable inductance than other types of resistor, although winding the wire in sections with alternately reversed direction can minimize inductance. Other techniques employ bifilar winding, or a flat thin former (to reduce cross-section area of the coil). For most demanding circuits resistors with Ayrton-Perry winding are used.

Applications of wirewound resistors are similar to those of composition resistors with the exception of the high frequency. The high frequency of wirewound resistors is substantially worse than that of a composition resistor

Foil resistor

The primary resistance element of a foil resistor is a special alloy foil several micrometres thick. Since their introduction in the 1960s, foil resistors have had the best precision and stability of any resistor available. One of the important parameters influencing stability is the temperature coefficient of resistance (TCR). The TCR of foil resistors is extremely low, and has been further improved over the years. One range of ultra-precision foil resistors offers a TCR of 0.14 ppm/°C, tolerance ±0.005%, long-term stability (1 year) 25 ppm, (3 year) 50 ppm (further improved 5-fold by hermetic sealing), stability under load (2000 hours) 0.03%, thermal EMF 0.1 μV/°C, noise -42 dB, voltage coefficient 0.1 ppm/V, inductance 0.08 μH, capacitance 0.5 pF.

Ammeter shunts

An ammeter shunt is a special type of current-sensing resistor, having four terminals and a value in milliohms or even micro-ohms. Current-measuring instruments, by themselves, can usually accept only limited currents. To measure high currents, the current passes through the shunt, where the voltage drop is measured and interpreted as current. A typical shunt consists of two solid metal blocks, sometimes brass, mounted on to an insulating base. Between the blocks, and soldered or brazed to them, are one or more strips of low temperature coefficient of resistance (TCR) manganin alloy. Large bolts threaded into the blocks make the current connections, while much-smaller screws provide voltage connections. Shunts are rated by full-scale current, and often have a voltage drop of 50 mV at rated current. Such meters are adapted to the shunt full current rating by using an appropriately marked dial face; no change need be made to the other parts of the meter.

Grid resistor

In heavy-duty industrial high-current applications, a grid resistor is a large convection-cooled lattice of stamped metal alloy strips connected in rows between two electrodes. Such industrial grade resistors can be as large as a refrigerator; some designs can handle over 500 amperes of current, with a range of resistances extending lower than 0.04 ohms. They are used in applications such as dynamic braking and load banking for locomotives and trams, neutral grounding for industrial AC distribution, control loads for cranes and heavy equipment, load testing of generators and harmonic filtering for electric substations.

The term *grid resistor* is sometimes used to describe a resistor of any type connected to the control grid of a vacuum tube. This is not a resistor technology; it is an electronic circuit topology.

Special varieties

- Metal oxide varistor

- Cermet
- Phenolic
- Tantalum
- Water resistor

Variable resistors

Adjustable resistors

A resistor may have one or more fixed tapping points so that the resistance can be changed by moving the connecting wires to different terminals. Some wirewound power resistors have a tapping point that can slide along the resistance element, allowing a larger or smaller part of the resistance to be used.

Where continuous adjustment of the resistance value during operation of equipment is required, the sliding resistance tap can be connected to a knob accessible to an operator. Such a device is called a rheostat and has two terminals.

Potentiometers

A common element in electronic devices is a three-terminal resistor with a continuously adjustable tapping point controlled by rotation of a shaft or knob. These variable resistors are known as potentiometers when all three terminals are present, since they act as a continuously adjustable voltage divider. A common example is a volume control for a radio receiver.

Accurate, high-resolution panel-mounted potentiometers (or "pots") have resistance elements typically wirewound on a helical mandrel, although some include a conductive-plastic resistance coating over the wire to improve resolution. These typically offer ten turns of their shafts to cover their full range. They are usually set with dials that include a simple turns counter and a graduated dial. Electronic analog computers used them in quantity for setting coefficients, and delayed-sweep oscilloscopes of recent decades included one on their panels.

Resistance decade boxes

A resistance decade box or resistor substitution box is a unit containing resistors of many values, with one or more mechanical switches which allow any one of various discrete resistances offered by the box to be dialed in. Usually the resistance is accurate to high precision, ranging from laboratory/calibration grade accuracy of 20 parts per million, to field grade at 1%. Inexpensive boxes with lesser accuracy are also available. All types offer a convenient way of selecting and quickly changing a resistance in laboratory, experimental and development work without needing to attach resistors one by one, or even stock each value. The range of resistance provided, the maximum resolution, and the accuracy characterize the box. For example, one box offers resistances from 0 to 24 megohms, maximum resolution 0.1 ohm, accuracy 0.1%.

Special devices

There are various devices whose resistance changes with various quantities. The resistance of thermistors exhibit a strong negative temperature coefficient, making them useful for measuring temperatures. Since their resistance can be large until they are allowed to heat up due to the passage of current, they are also commonly used to prevent excessive current surges when equipment is powered on. Similarly, the resistance of a humistor varies with humidity. Metal oxide varistors drop to a very low resistance when a high voltage is applied, making them useful for protecting electronic equipment by absorbing dangerous voltage surges. One sort of photodetector, the photoresistor, has a resistance which varies with illumination.

The strain gauge, invented by Edward E. Simmons and Arthur C. Ruge in 1938, is a type of resistor that changes value with applied strain. A single resistor may be used, or a pair (half bridge), or four resistors connected in a Wheatstone bridge configuration. The strain resistor is bonded with adhesive to an object that will be subjected to mechanical strain. With the strain gauge and a filter, amplifier, and analog/digital converter, the strain on an object can be measured.

A related but more recent invention uses a Quantum Tunnelling Composite to sense mechanical stress. It passes a current whose magnitude can vary by a factor of 10^{12} in response to changes in applied pressure.

Measurement

The value of a resistor can be measured with an ohmmeter, which may be one function of a multimeter. Usually, probes on the ends of test leads connect to the resistor. A simple ohmmeter may apply a voltage from a battery across the unknown resistor (with an internal resistor of a known value in series) producing a current which drives a meter movement. The current flow, in accordance with Ohm's Law, is inversely proportional to the sum of the internal resistance and the resistor being tested, resulting in an analog meter scale which is very non-linear, calibrated from infinity to 0 ohms. A digital multimeter, using active electronics, may instead pass a specified current through the test resistance. The voltage generated across the test resistance in that case is linearly proportional to its resistance, which is measured and displayed. In either case the low-resistance ranges of the meter pass much more current through the test leads than do high-resistance ranges, in order for the voltages present to be at reasonable levels (generally below 10 volts) but still measurable.

Measuring low-value resistors, such as fractional-ohm resistors, with acceptable accuracy requires four-terminal connections. One pair of terminals applies a known, calibrated current to the resistor, while the other pair senses the voltage drop across the resistor. Some laboratory quality ohmmeters, especially milliohmmeters, and even some of the better digital multimeters sense using four input terminals for this purpose, which may be used with special test leads. Each of the two so-called Kelvin clips has a pair of jaws insulated from each other. One side of each clip applies the measuring current, while the other connections are only to sense the voltage drop. The resistance is again calculated using Ohm's Law as the measured voltage divided by the applied current.

Standards

Production resistors

Resistor characteristics are quantified and reported using various national standards. In the US, MIL-STD-202 contains the relevant test methods to which other standards refer.

There are various standards specifying properties of resistors for use in equipment:

- BS 1852
- EIA-RS-279
- MIL-PRF-26
- MIL-PRF-39007 (Fixed Power, established reliability)
- MIL-PRF-55342 (Surface-mount thick and thin film)
- MIL-PRF-914
- MIL-R-11
- MIL-R-39017 (Fixed, General Purpose, Established Reliability)
- MIL-PRF-32159 (zero ohm jumpers)

There are other United States military procurement MIL-R- standards.

Resistance standards

The primary standard for resistance, the "mercury ohm" was initially defined in 1884 in as a column of mercury 106 mm long and 1 square millimeter in cross-section, at 0 degrees Celsius. Difficulties in precisely measuring the physical constants to replicate this standard result in variations of as much as 30 ppm. From 1900 the mercury ohm was replaced with a precision machined plate of manganin. Since 1990 the international resistance standard has been based on the quantized Hall effect discovered by Klaus von Klitzing, for which he won the Nobel Prize in Physics in 1985.

Resistors of extremely high precision are manufactured for calibration and laboratory use. They may have four terminals, using one pair to carry an operating current and the other pair to measure the voltage drop; this eliminates errors caused by voltage drops across the lead resistances, because no current flows through voltage sensing leads. It is important in small value resistors (100–0.0001 ohm) where lead resistance is significant or even comparable with respect to resistance standard value.

Resistor marking

Most axial resistors use a pattern of colored stripes to indicate resistance. Surface-mount resistors are marked numerically, if they are big enough to permit marking; more-recent small sizes are impractical to mark. Cases are usually tan, brown, blue, or green, though other colors are occasionally found such as dark red or dark gray.

Early 20th century resistors, essentially uninsulated, were dipped in paint to cover their entire body for color coding. A second color of paint was applied to one end of the element, and a color dot (or band) in the middle provided the third digit. The rule was "body, tip, dot", providing two significant digits for value and the decimal multiplier, in that sequence. Default tolerance was ±20%. Closer-tolerance resistors had silver (±10%) or gold-colored (±5%) paint on the other end.

Four-band resistors

Four-band identification is the most commonly used color-coding scheme on resistors. It consists of four colored bands that are painted around the body of the resistor. The first two bands encode the first two significant digits of the resistance value, the third is a power-of-ten multiplier or number-of-zeroes, and the fourth is the tolerance accuracy, or acceptable error, of the value. The first three bands are equally spaced along the resistor; the spacing to the fourth band is wider. Sometimes a fifth band identifies the thermal coefficient, but this must be distinguished from the true 5-color system, with 3 significant digits.

For example, green-blue-yellow-red is 56×10^4 Ω = 560 kΩ ± 2%. An easier description can be as followed: the first band, green, has a value of 5 and the second band, blue, has a value of 6, and is counted as 56. The third band, yellow, has a value of 10^4, which adds four 0's to the end, creating 560,000 Ω at ±2% tolerance accuracy. 560,000 Ω changes to 560 kΩ ±2% (as a kilo- is 10^3).

Each color corresponds to a certain digit, progressing from darker to lighter colors, as shown in the chart below.

Color	**1st band**	**2nd band**	**3rd band (multiplier)**	**4th band (tolerance)**	**Temp. Coefficient**
Black	0	0	$\times10^0$		
Brown	1	1	$\times10^1$	±1% (F)	100 ppm
Red	2	2	$\times10^2$	±2% (G)	50 ppm
Orange	3	3	$\times10^3$		15 ppm
Yellow	4	4	$\times10^4$		25 ppm
Green	5	5	$\times10^5$	±0.5% (D)	
Blue	6	6	$\times10^6$	±0.25% (C)	
Violet	7	7	$\times10^7$	±0.1% (B)	
Gray	8	8	$\times10^8$	±0.05% (A)	
White	9	9	$\times10^9$		
Gold			$\times10^{-1}$	±5% (J)	
Silver			$\times10^{-2}$	±10% (K)	
None				±20% (M)	

There are many mnemonics for remembering these colors.

Preferred values

Early resistors were made in more or less arbitrary round numbers; a series might have 100, 125, 150, 200, 300, etc. Resistors as manufactured are subject to a certain percentage tolerance, and it makes sense to manufacture values that correlate with the tolerance, so that the actual value of a resistor overlaps slightly with its neighbors. Wider spacing leaves gaps; narrower spacing increases manufacturing and inventory costs to provide resistors that are more or less interchangeable.

A logical scheme is to produce resistors in a range of values which increase in a geometrical progression, so that each value is greater than its predecessor by a fixed multiplier or percentage, chosen to match the tolerance of the range. For example, for a tolerance of ±20% it makes sense to have each resistor about 1.5 times its predecessor, covering a decade in 6 values. In practice the factor used is 1.4678, giving values of 1.47, 2.15, 3.16, 4.64, 6.81, 10 for the 1-10 decade (a decade is a range increasing by a factor of 10; 0.1-1 and 10-100 are other examples); these are rounded in practice to 1.5, 2.2, 3.3, 4.7, 6.8, 10; followed, of course by 15, 22, 33, … and preceded by … 0.47, 0.68, 1. This scheme has been adopted as the **E6** range of the IEC 60063 preferred number series. There are also **E12**, **E24**, **E48**, **E96** and **E192** ranges for components of ever tighter tolerance, with 12, 24, 96, and 192 different values within each decade. The actual values used are in the IEC 60063 lists of preferred numbers.

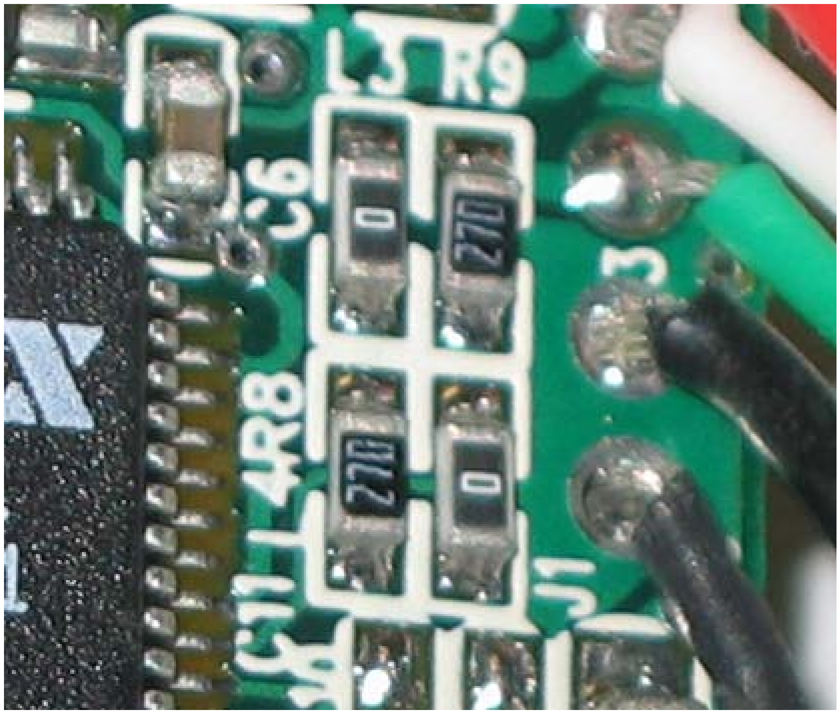
A resistor of 100 ohms ±20% would be expected to have a value between 80 and 120 ohms; its E6 neighbors are 68 (54-82) and 150 (120-180) ohms. A sensible spacing, E6 is used for ±20% components; E12 for ±10%; E24 for ±5%; E48 for ±2%, E96 for ±1%; E192 for ±0.5% or better. Resistors are manufactured in values from a few milliohms to about a gigaohm in IEC60063 ranges appropriate for their tolerance.

Earlier power wirewound resistors, such as brown vitreous-enameled types, however, were made with a different system of preferred values, such as some of those mentioned in the first sentence of this section.

5-band axial resistors

5-band identification is used for higher precision (lower tolerance) resistors (1%, 0.5%, 0.25%, 0.1%), to specify a third significant digit. The first three bands represent the significant digits, the fourth is the multiplier, and the fifth is the tolerance. Five-band resistors with a gold or silver 4th band are sometimes encountered, generally on older or specialized resistors. The 4th band is the tolerance and the 5th the temperature coefficient.

SMD resistors



This image shows four surface-mount resistors (the component at the upper left is a capacitor) including two zero-ohm resistors. Zero-ohm links are often used instead of wire links, so that they can be inserted by a resistor-inserting machine. Of course, their resistance is non-zero, although quite low. *Zero* is simply a brief description of their function.

Surface mounted resistors are printed with numerical values in a code related to that used on axial resistors. Standard-tolerance surface-mount technology (SMT) resistors are marked with a three-digit code, in which the first two digits are the first two significant digits of the value and the third digit is the power of ten (the number of zeroes). For example:

334 = 33 × 10^4 ohms = 330 kilohms

222 = 22 × 10^2 ohms = 2.2 kilohms

473 = 47 × 10^3 ohms = 47 kilohms

105 = 10 × 10^5 ohms = 1.0 megohm

Resistances less than 100 ohms are written: 100, 220, 470. The final zero represents ten to the power zero, which is 1. For example:

100 = 10 × 10^0 ohm = 10 ohms

220 = 22 × 10^0 ohm = 22 ohms

Sometimes these values are marked as *10* or *22* to prevent a mistake.

Resistances less than 10 ohms have 'R' to indicate the position of the decimal point (radix point). For example:

4R7 = 4.7 ohms
R300 = 0.30 ohms
0R22 = 0.22 ohms
0R01 = 0.01 ohms

Precision resistors are marked with a four-digit code, in which the first three digits are the significant figures and the fourth is the power of ten. For example:

1001 = 100 × 10^1 ohms = 1.00 kilohm
4992 = 499 × 10^2 ohms = 49.9 kilohm
1000 = 100 × 10^0 ohm = 100 ohms

000 and *0000* sometimes appear as values on surface-mount zero-ohm links, since these have (approximately) zero resistance.

More recent surface-mount resistors are too small, physically, to permit practical markings to be applied.

Industrial type designation

Format: *[two letters]<space>[resistance value (three digit)]<nospace>[tolerance code(numerical - one digit)]*

Power Rating at 70 °C

Type No.	Power rating (watts)	MIL-R-11 Style	MIL-R-39008 Style
BB	⅛	RC05	RCR05
CB	¼	RC07	RCR07
EB	½	RC20	RCR20
GB	1	RC32	RCR32
HB	2	RC42	RCR42
GM	3	-	-
HM	4	-	-

Tolerance Code

Industrial type designation	Tolerance	MIL Designation
5	±5%	J
2	±20%	M
1	±10%	K
-	±2%	G
-	±1%	F
-	±0.5%	D
-	±0.25%	C
-	±0.1%	B

The operational temperature range distinguishes commercial grade, industrial grade and military grade components.

- Commercial grade: 0 °C to 70 °C
- Industrial grade: −40 °C to 85 °C (sometimes −25 °C to 85 °C)
- Military grade: −55 °C to 125 °C (sometimes -65 °C to 275 °C)
- Standard Grade -5 °C to 60 °C

Electrical and thermal noise

In amplifying faint signals, it is often necessary to minimize electronic noise, particularly in the first stage of amplification. As dissipative elements, even an ideal resistor will naturally produce a randomly fluctuating voltage or "noise" across its terminals. This Johnson–Nyquist noise is a fundamental noise source which depends only upon the temperature and resistance of the resistor, and is predicted by the fluctuation–dissipation theorem. Using a larger resistor produces a larger voltage noise, whereas with a smaller value of resistance there will be more current noise, assuming a given temperature. The thermal noise of a practical resistor may also be somewhat larger than the theoretical prediction and that increase is typically frequency-dependent.

However the "excess noise" of a practical resistor is an additional source of noise observed only when a current flows through it. This is specified in unit of μV/V/decade - μV of noise per volt applied across the resistor per decade of frequency. The μV/V/decade value is frequently given in dB so that a resistor with a noise index of 0 dB will exhibit 1 μV (rms) of excess noise for each volt across the resistor in each frequency decade. Excess noise is thus an example of $1/f$ noise. Thick-film and carbon composition resistors generate more excess noise than other types at low frequencies; wire-wound and thin-film resistors, though much more expensive, are often utilized for their better noise characteristics. Carbon composition resistors can exhibit a noise index of 0 dB while bulk metal foil resistors may have a noise index of -40 dB, usually making the excess noise of metal foil resistors insignificant. Thin film surface mount resistors typically have lower noise and better thermal stability than thick film surface mount resistors. However, the design engineer must read the data sheets for the family of devices to weigh the various device tradeoffs.

While not an example of "noise" per se, a resistor may act as a thermocouple, producing a small DC voltage differential across it due to the thermoelectric effect if its ends are at somewhat different temperatures. This induced DC voltage can degrade the precision of instrumentation amplifiers in particular. Such voltages appear in the junctions of the resistor leads with the circuit board and with the resistor body. Common metal film resistors show such an effect at a magnitude of about 20 μV/°C. Some carbon composition resistors can exhibit thermoelectric offsets as high as 400 μV/°C, whereas specially constructed resistors can reduce this number to 0.05 μV/°C. In applications where the thermoelectric effect may become important, care has to be taken (for example) to mount the resistors horizontally to avoid temperature gradients and to mind the air flow over the board.

Failure modes

The failure rate of resistors in a properly designed circuit is low compared to other electronic components such as semiconductors and electrolytic capacitors. Damage to resistors most often occurs due to overheating when the average power delivered to it (as computed above) greatly exceeds its ability to dissipate heat (specified by the resistor's *power rating*). This may be due to a fault external to the circuit, but is frequently caused by the failure of another component (such as a transistor that shorts out) in the circuit connected to the resistor. Operating a resistor too close to its power rating can limit the resistor's lifespan or cause a change in its resistance over time which may or may not be noticeable. A safe design generally uses overrated resistors in power applications to avoid this danger.

When overheated, carbon-film resistors may decrease or increase in resistance. Carbon film and composition resistors can fail (open circuit) if running close to their maximum dissipation. This is also possible but less likely with metal film and wirewound resistors.

There can also be failure of resistors due to mechanical stress and adverse environmental factors including humidity. If not enclosed, wirewound resistors can corrode.

Variable resistors degrade in a different manner, typically involving poor contact between the wiper and the body of the resistance. This may be due to dirt or corrosion and is typically perceived as "crackling" as the contact resistance fluctuates; this is especially noticed as the device is adjusted. This is similar to crackling caused by poor contact in switches, and like switches, potentiometers are to some extent self-cleaning: running the wiper across the resistance may improve the contact. Potentiometers which are seldom adjusted, especially in dirty or harsh environments, are most likely to develop this problem. When self-cleaning of the contact is insufficient, improvement can usually be obtained through the use of contact cleaner (also known as "tuner cleaner") spray. The crackling noise associated with turning the shaft of a dirty potentiometer in an audio circuit (such as the volume control) is greatly accentuated when an undesired DC voltage is present, often implicating the failure of a DC blocking capacitor in the circuit.

Chapter 4

Electrical Ballast

"Choke ballast" (inductor) used in older lighting. This example is from a tanning bed. Requires a lamp starter (below) and capacitor.

Lamp starter, required with some inductor type ballasts. Connects both ends of the lamp together to "preheat" the lamp ends for 1 second before lighting.

An **electrical ballast** (sometimes called **control gear**) is a device intended to limit the amount of current in an electric circuit.

Ballasts vary greatly in complexity. They can be as simple as a series resistor as commonly used with small neon lamps or light-emitting diodes (LEDs). For higher-power installations, too much energy would be wasted in a resistive ballast, so alternatives are used that depend upon the reactance of inductors, capacitors, or both. Finally, ballasts can be as complex as the computerized, remote-controlled electronic ballasts now often used with fluorescent lamps.

Current limiting

Ballasts stabilize the current through an electrical load. These are most often used when an electrical circuit or device presents a negative (differential) resistance to the supply. If such a device were connected to a constant-voltage power supply, it would draw an increasing amount of current until it was destroyed or caused the power supply to fail. To prevent this, a ballast provides a positive resistance or reactance that limits the ultimate current to an appropriate level. In this way, the ballast provides for the proper operation of the negative-resistance device by appearing to be a legitimate, stable resistance in the circuit.

An example of a negative-resistance device is a gas-discharge lamp, where after lamp ignition, increasing arc current reduces the voltage drop.

Ballasts can also be used simply to deliberately reduce the current in an ordinary, positive-resistance circuit.

Prior to the advent of solid-state ignition, automobile ignition systems commonly included a ballast resistor to regulate the voltage applied to the ignition system.

Although LEDs are positive resistance devices, they have insufficient resistance to regulate their current consumption when operated from a voltage controlled source, so ballasts are used to control the current through the LED. Because the power dissipation is minuscule, simple resistor ballasts are normally used.

Resistors

A **ballast resistor** compensates for normal or incidental changes in the physical state of a system. It may be a fixed or variable resistor.

Fixed resistors

For simple, low-powered loads such as a neon lamp or LED, a fixed resistor is commonly used. Because the resistance of the ballast resistor is large it dominates the current in the circuit, even in the face of negative resistance introduced by the neon lamp.

The term also refers to an automobile engine component that lowers the supply voltage to the ignition system after the engine has been started. Because cranking the engine causes a very heavy load on the battery, the system voltage can drop quite low during cranking. To allow the engine to start, the ignition system must be designed to operate on this lower voltage. But once cranking is completed, the normal operating voltage is regained; this voltage would overload the ignition system. To avoid this problem, a ballast resistor is inserted in series with the supply voltage feeding the ignition system. Occasionally, this ballast resistor will fail and the classic symptom of this failure is that the engine runs while being cranked (while the resistor is bypassed) but stalls immediately when cranking ceases (and the resistor is re-connected in the circuit).

Modern electronic ignition systems do not require a ballast resistor as they are flexible enough to operate on the low cranking voltage or the ordinary operating voltage.

In some old AC/DC receivers (universal sets), the vacuum tube heaters are connected in series. Since the voltage drop across all the filaments in series is sometimes less than the full mains voltage, it was often necessary to get rid of the excess voltage. A ballast resistor was often used for this purpose, as it was cheap and worked with both AC and DC.

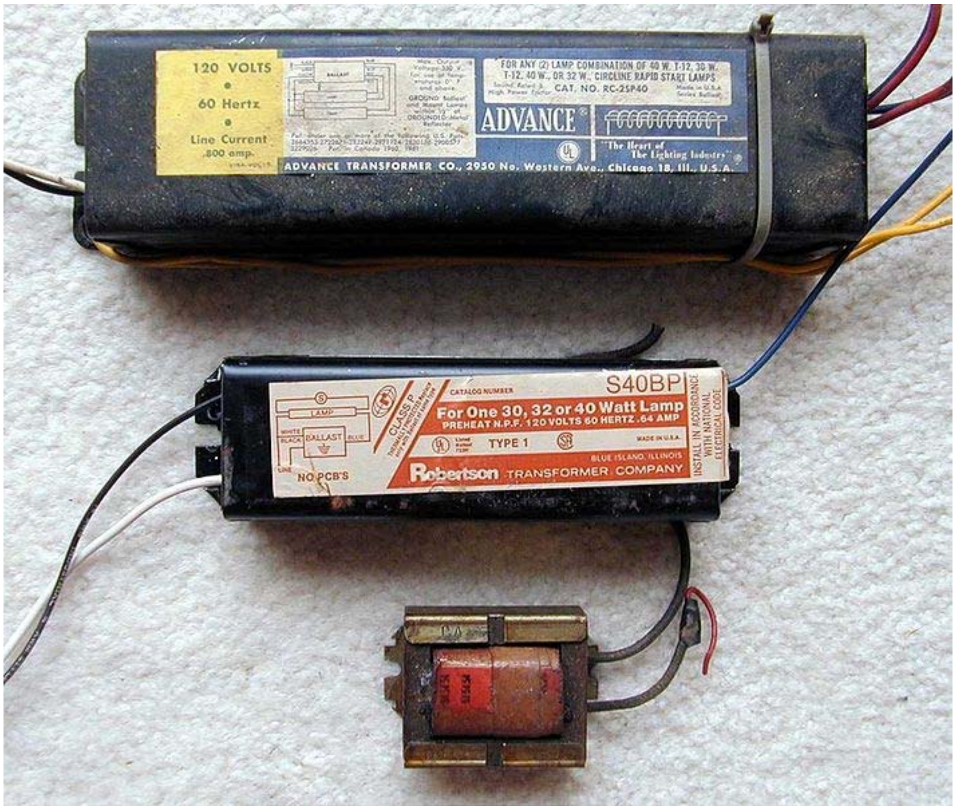
Self-variable resistors

Some ballast resistors have the property of increasing in resistance as current through them increases, and decreasing in resistance as current decreases. Physically, some such devices are often built quite like incandescent lamps. Like the tungsten filament of an ordinary incandescent lamp, if current increases, the ballast resistor gets hotter, its resistance goes up, and its voltage drop increases. If current decreases, the ballast resistor gets colder, its resistance drops, and the voltage drop decreases. Therefore the ballast resistor reduces variations in current, despite variations in applied voltage or changes in the rest of an electric circuit. These devices are sometimes termed barretters.

This property can lead to more precise current control than merely choosing an appropriate fixed resistor. The power lost in the resistive ballast is also reduced because a smaller portion of the overall power is dropped in the ballast compared to what might be required with a fixed resistor.

In times past, household clothes dryers sometimes incorporated a germicidal lamp in series with an ordinary incandescent lamp; the incandescent lamp operated as the ballast for the germicidal lamp. A commonly used light in the home in the 1960s in 220-240V countries was a circleline tube ballasted by an under-run regular mains filament lamp. Self ballasted mercury-vapor lamps incorporate ordinary tungsten filaments within the overall envelope of the lamp to act as the ballast, and it supplements the otherwise lacking red area of the light spectrum produced.

Reactive ballasts



Several typical magnetic ballasts for fluorescent lamps. The top is a high-power factor rapid start series ballast for two 30-40W lamps. The middle is a low power factor preheat ballast for a single 30-40W lamp while the bottom ballast is a simple inductor used with a 15W preheat lamp.

Because of the power that would be lost, resistors are not used as ballasts for lamps of more than about two watts. Instead, a reactance is used. Losses in the ballast due to its resistance and losses in its magnetic core may be significant, on the order of 5 to 25% of the lamp input wattage. Practical lighting design calculations must allow for ballast loss in estimating the running cost of a lighting installation.

An inductor is very common in line-frequency ballasts to provide the proper starting and operating electrical condition to power a fluorescent lamp, neon lamp, or high intensity discharge (HID) lamp. (Because of the use of the inductor, such ballasts are usually called *magnetic ballasts*.) The inductor has two benefits:

1. Its reactance limits the power available to the lamp with only minimal power losses in the inductor
2. The voltage spike produced when current through the inductor is rapidly interrupted is used in some circuits to first strike the arc in the lamp.

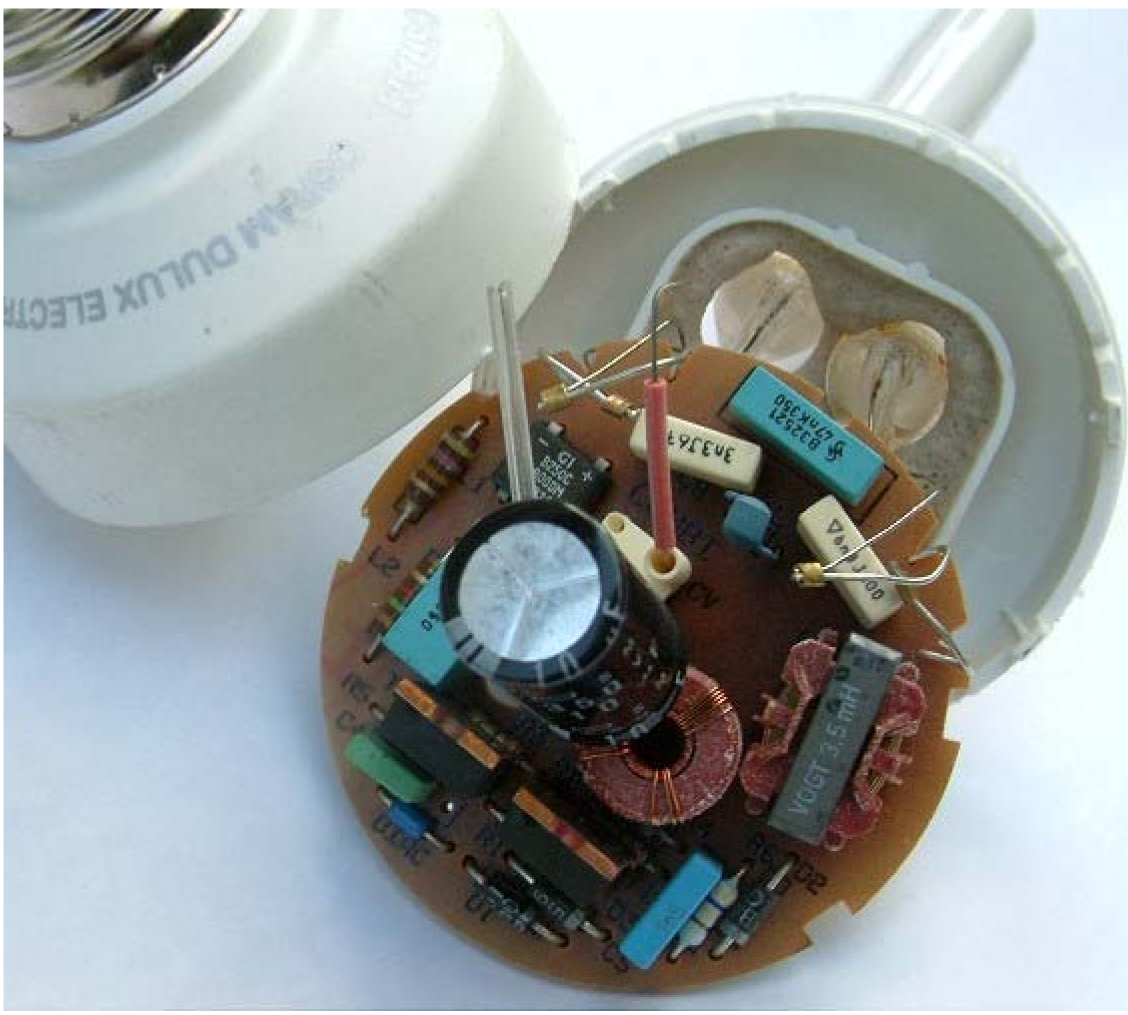
A disadvantage of the inductor is that current is shifted out of phase with the voltage, producing a poor power factor. In more expensive ballasts, a capacitor is often paired with the inductor to correct the power factor. In ballasts that control two or more lamps, line-frequency ballasts commonly use different phase relationships between the multiple lamps. This not only mitigates the flicker of the individual lamps, it also helps maintain a high power factor. These ballasts are often called *lead-lag* ballasts because the current in one lamp leads the mains phase and the current in the other lamp lags the mains phase.

For large lamps, line voltage may not be sufficient to start the lamp, so an autotransformer winding is included in the ballast to step up the voltage. The autotransformer is designed with enough leakage inductance so that the current is appropriately limited.

Because of the large inductors and capacitors that must be used, reactive ballasts operated at line frequency tend to be large and heavy. They commonly also produce acoustic noise (line-frequency hum).

Prior to 1980 in the United States, PCB-based oils were used as an insulating oil in many ballasts to provide cooling and electrical isolation.

Electronic ballasts



Electronic ballast of a compact fluorescent lamp

An **electronic lamp ballast** uses solid state electronic circuitry to provide the proper starting and operating electrical condition to power one or more fluorescent lamps and more recently HID lamps. Electronic ballasts usually change the frequency of the power from the standard mains (e.g., 60 Hz in U.S.) frequency to 20,000 Hz or higher, substantially eliminating the stroboscopic effect of flicker (a product of the line frequency) associated with fluorescent lighting. In addition, because more gas remains ionized in the arc stream, the lamps actually operate at about 9% higher efficacy above approximately 10 kHz. Lamp efficacy increases sharply at about 10 kHz and continues to improve until approximately 20 kHz. Because of the higher efficiency of the ballast itself and the improvement of lamp efficacy by operating at a higher frequency, electronic ballasts offer higher system efficacy for low pressure lamps like the fluorescent lamp. For HID lamps there is no improvement of the lamp efficacy in using higher frequency, but for these lamps the ballast losses are lower at higher frequencies and also the light depreciation is lower meaning more light after a given operating time of say 10 000 hours. Some HID lamp types like the Ceramic discharge metal halide lamp have reduced reliability when operated at high frequencies in the range of 20kHz to 200 kHz and for these lamps a square wave low frequency current drive is mostly used with frequency in the range of 100 to 400 Hz, with the same advantage of lower light depreciation. Electronic ballasts are often based on the SMPS topology, first rectifying the input power and then chopping it at a high frequency. Advanced electronic ballasts may allow dimming via pulse-width modulation or via changing the frequency to a higher value and remote control and monitoring via networks such as LonWorks, DALI, DMX-512, DSI or simple analog control using a 0-10V DC brightness control signal. Recently also systems remotely controlling the dim level via a wireless mesh network have been introduced.

Fluorescent lamp ballasts

Instant start

An instant start ballast starts lamps without heating the cathodes at all by using high voltage (around 600 V). It is the most energy efficient type, but gives the least number of starts from a lamp as emissive oxides are blasted from the cold cathode surfaces each time the lamp is started. This is the best type for installations where lamps are not turned on and off very often.

Rapid start

A rapid start ballast applies voltage and heats the cathodes simultaneously. Provides superior lamp life and more cycle life, but uses slightly more energy as the cathodes in each end of the lamp continue to consume heating power as the lamp operates. A dimming circuit can be used with a dimming ballast, which maintains the heating current while allowing lamp current to be controlled.

Programmed start

A programmed-start ballast is a more advanced version of rapid start. This ballast applies power to the filaments first, then after a short delay to allow the cathodes to preheat, applies voltage to the lamps to strike an arc. This ballast gives the best life and most starts from lamps, and so is preferred for applications with very frequent power cycling such as vision examination rooms and restrooms with a motion detector switch.

Ballast factor

For a lighting ballast, the *ballast factor* is defined as the light output (in lumens) with a test ballast, compared to the light output with a laboratory reference ballast that operates the lamp at its specified nominal power rating. The ballast factor of practical ballasts must be considered in lighting design; a low ballast factor may save energy, but will produce less light. With fluorescent lamps, an electronic ballast may produce more light than the reference test ballast, which operates the lamp with line frequency current; such electronic ballasts have a ballast factor greater than one.

Chapter 5

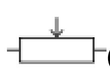
Potentiometer

Potentiometer

A typical single-turn potentiometer

Type Passive

Electronic symbol

(Europe)

(US)

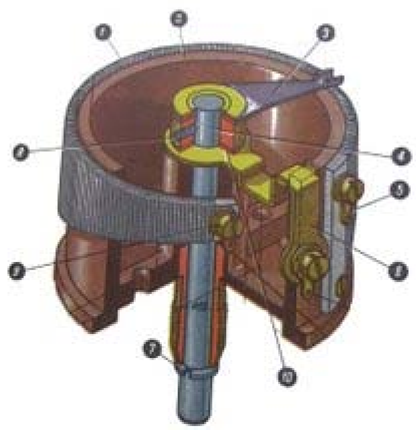
A **potentiometer** (colloquially known as a "**pot**") is a three-terminal resistor with a sliding contact that forms an adjustable voltage divider. If only two terminals are used (one side and the wiper), it acts as a ***variable resistor*** or ***rheostat***. Potentiometers are commonly used to control electrical devices such as volume controls on audio equipment. Potentiometers operated by a mechanism can be used as position transducers, for example, in a joystick.

Potentiometers are rarely used to directly control significant power (more than a watt), since the power dissipated in the potentiometer would be comparable to the power in the controlled load. Instead they are used to adjust the level of analog signals (e.g. volume controls on audio equipment), and as control inputs for electronic circuits. For example, a light dimmer uses a potentiometer to control the switching of a TRIAC and so indirectly control the brightness of lamps.

History

The slide-wire potentiometer was invented by Johann Christian Poggendorff (1796–1877) in 1841. Leeds and Northrup Type K model was a standard piece of apparatus in most college and university electrical measurements laboratories for the first half of the 20th century.

Potentiometer construction



Construction of a wire-wound circular potentiometer. The resistive element (1) of the shown device is trapezoidal, giving a non-linear relationship between resistance and turn angle. The wiper (3) rotates with the axis (4), providing the changeable resistance between the wiper contact (6) and the fixed contacts (5) and (9). The vertical position of the axis is fixed in the body (2) with the ring (7) (below) and the bolt (8) (above).

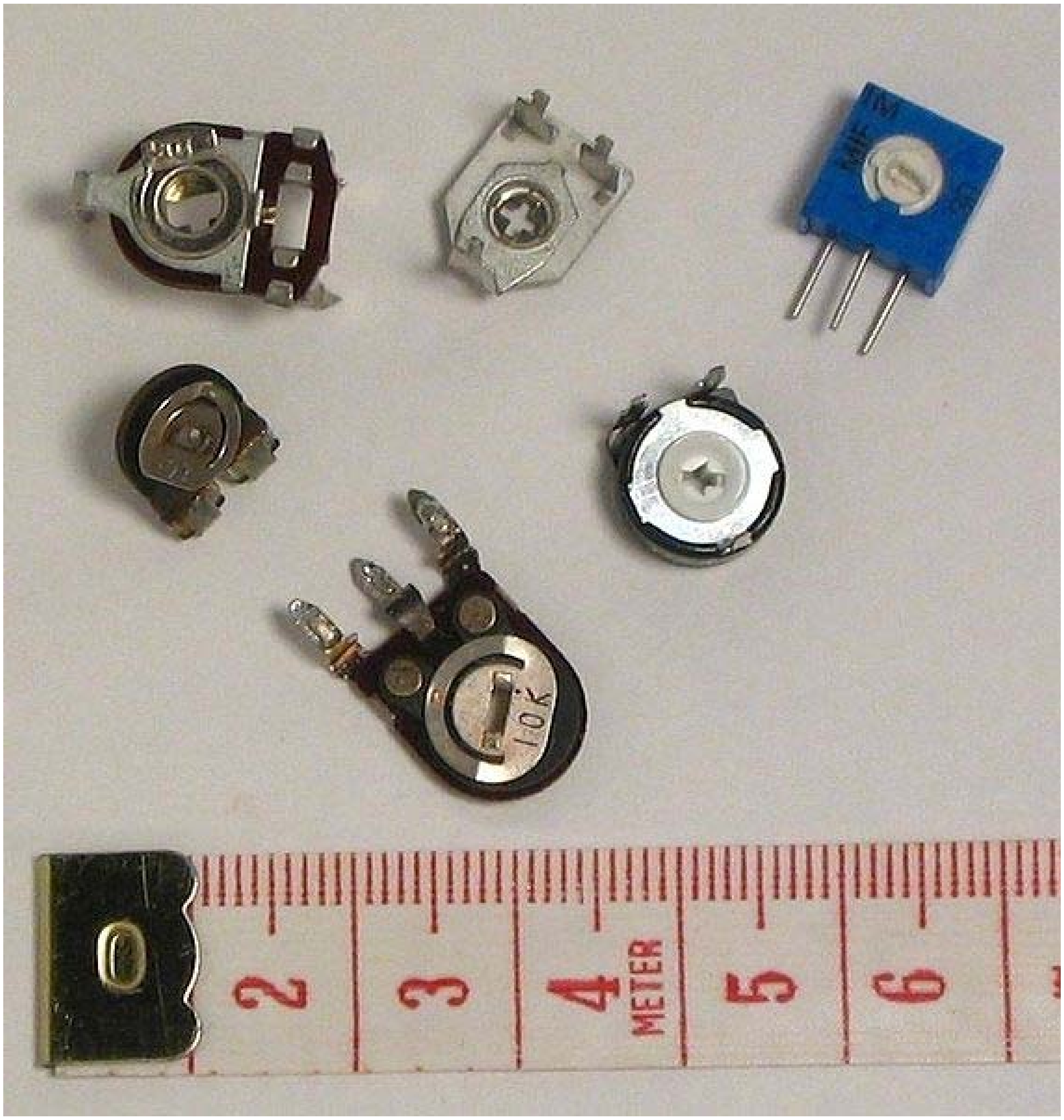
A potentiometer is constructed with a resistive element formed into an arc of a circle, and a sliding contact (wiper) travelling over that arc. The resistive element, with a terminal at one or both ends, is flat or angled, and is commonly made of graphite, although other materials may be used. The wiper is connected through another sliding contact to another terminal. On panel potentiometers, the wiper is usually the center terminal of three. For single-turn potentiometers, this wiper typically travels just under one revolution around the contact. "Multiturn" potentiometers also exist, where the resistor element may be helical and the wiper may move 10, 20, or more complete revolutions, though multiturn potentimeters are usually constructed of a conventional resistive element wiped via a worm gear. Besides graphite, materials used to make the resistive element include resistance wire, carbon particles in plastic, and a ceramic/metal mixture called cermet.

One form of rotary potentiometer is called a String potentiometer. It is a multi-turn potentiometer operated by an attached reel of wire turning against a spring. It is used as a position transducer.

In a linear slider potentiometer, a sliding control is provided instead of a dial control. The resistive element is a rectangular strip, not semi-circular as in a rotary potentiometer. Due to the large opening slot or the wiper, this type of potentiometer has a greater potential for getting contaminated.

Potentiometers can be obtained with either linear or logarithmic relations between the slider position and the resistance (potentiometer laws or "tapers"). A letter code ("A" taper, "B" taper, etc.) may be used to identify which taper is intended, but the letter code definitions are variable over time and between manufacturers.

Manufacturers of conductive track potentiometers use conductive polymer resistor pastes that contain hard wearing resins and polymers, solvents, lubricant and carbon – the constituent that provides the conductive/resistive properties. The tracks are made by screen printing the paste onto a paper based phenolic substrate and then curing it in an oven. The curing process removes all solvents and allows the conductive polymer to polymerize and cross link. This produces a durable track with stable electrical resistance throughout its working life.



PCB mount trimmer potentiometers, or "trimpots", intended for infrequent adjustment.

Linear taper potentiometer

A *linear taper potentiometer* has a resistive element of constant cross-section, resulting in a device where the resistance between the contact (wiper) and one end terminal is proportional to the distance between them. *Linear taper* describes the electrical characteristic of the device, not the geometry of the resistive element. Linear taper potentiometers are used when an approximately proportional relation is desired between shaft rotation and the division ratio of the potentiometer; for example, controls used for adjusting the centering of (an analog) cathode-ray oscilloscope.

Logarithmic potentiometer

A *logarithmic taper potentiometer* has a resistive element that either 'tapers' in from one end to the other, or is made from a material whose resistivity varies from one end to the other. This results in a device where output voltage is a logarithmic function of the mechanical angle of the potentiometer.

Most (cheaper) "log" potentiometers are actually not logarithmic, but use two regions of different resistance (but constant resistivity) to approximate a logarithmic law. A logarithmic potentiometer can also be simulated with a linear one and an external resistor. True logarithmic potentiometers are significantly more expensive.

Logarithmic taper potentiometers are often used in connection with audio amplifiers as human perception of audio volume is logarithmic.

A high power wirewound potentiometer. Any potentiometer may be connected as a rheostat.

Rheostat

The most common way to vary the resistance in a circuit is to use a variable resistor or a **rheostat** A rheostat is a two-terminal variable resistor. Often these are designed to handle much higher voltage and current. Typically these are constructed as a resistive wire wrapped to form a toroid coil with the wiper moving over the upper surface of the toroid, sliding from one turn of the wire to the next. Sometimes a rheostat is made from resistance wire wound on a heat-resisting cylinder with the slider made from a number of metal fingers that grip lightly onto a small portion of the turns of resistance wire. The "fingers" can be moved along the coil of resistance wire by a sliding knob thus changing the "tapping" point. They are usually used as variable resistors rather than variable potential dividers.

Any three-terminal potentiometer can be used as a two-terminal variable resistor by not connecting to the third terminal. It is common practice to connect the wiper terminal to the unused end of the resistance track to reduce the amount of resistance variation caused by dirt on the track.

Digital potentiometer

A digital potentiometer is an electronic component that mimics the functions of analog potentiometers. Through digital input signals, the resistance between two terminals can be adjusted, just as in an analog potentiometer.

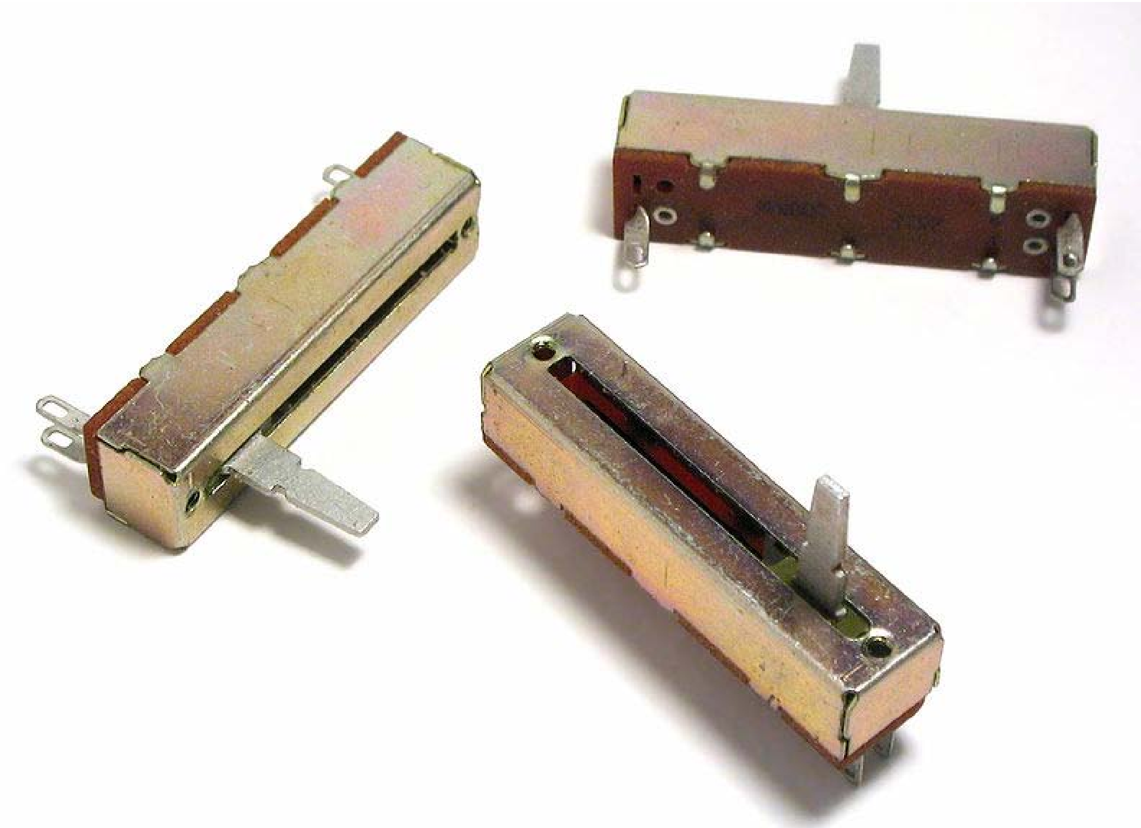
Membrane Potentiometer

A membrane potentiometer uses a conductive membrane that is deformed by a sliding element to contact a resistor voltage divider. Linearity can range from 0.5% to 5% depending on the material, design and manufacturing process. The repeat accuracy is typically between 0.1mm and 1.0mm with a theoretically infinite resolution. The service life of these types of potentiometers is typically 1 million to 20 million cycles depending on the materials used during manufacturing and the actuation method; contact and contactless (magnetic) methods are available. Many different material variations are available such as PET(foil), FR4, and Kapton. Membrane potentiometer manuafacturers offer linear, rotary, and application-specific variations. The linear versions can range from 9mm to 1000mm in length and the rotary versions range from 0° to 360°(multi-turn), with each having a height of 0.5mm. Membrane potentiometers can be used for position sensing.

Potentiometer applications

Potentiometers are widely used as user controls, and may control a very wide variety of equipment functions. The widespread use of potentiometers in consumer electronics has declined in the 1990s, with digital controls now more common. However they remain in many applications, such as volume controls and as position sensors.

Audio control



Linear potentiometers ("faders")

One of the most common uses for modern low-power potentiometers is as audio control devices. Both linear potentiometers and rotary potentiometers are regularly used to adjust loudness, frequency attenuation and other characteristics of audio signals.

The 'log pot' is used as the volume control in audio amplifiers, where it is also called an "audio taper pot", because the amplitude response of the human ear is also logarithmic. It ensures that, on a volume control marked 0 to 10, for example, a setting of 5 sounds half as loud as a setting of 10. There is also an *anti-log pot* or *reverse audio taper* which is simply the reverse of a logarithmic potentiometer. It is almost always used in a ganged configuration with a logarithmic potentiometer, for instance, in an audio balance control.

Potentiometers used in combination with filter networks act as tone controls or equalizers.

Television

Potentiometers were formerly used to control picture brightness, contrast, and color response. A potentiometer was often used to adjust "vertical hold", which affected the synchronization between the receiver's internal sweep circuit (sometimes a multivibrator) and the received picture signal.

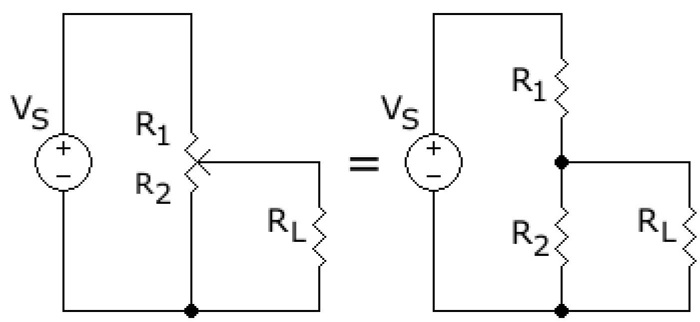
Transducers

Potentiometers are also very widely used as a part of displacement transducers because of the simplicity of construction and because they can give a large output signal.

Computation

In analog computers, high precision potentiometers are used to scale intermediate results by desired constant factors, or to set initial conditions for a calculation. A motor-driven potentiometer may be used as a function generator, using a non-linear resistance card to supply approximations to trigonometric functions. For example, the shaft rotation might represent an angle, and the voltage division ratio can be made proportional to the cosine of the angle.

Theory of operation



A potentiometer with a resistive load, showing equivalent fixed resistors for clarity.

The potentiometer can be used as a voltage divider to obtain a manually adjustable output voltage at the slider (wiper) from a fixed input voltage applied across the two ends of the potentiometer. This is the most common use of them.

The voltage across R_L can be calculated by:

$$V_\mathrm{L} = \frac{R_2 R_\mathrm{L}}{R_1 R_\mathrm{L} + R_2 R_\mathrm{L} + R_1 R_2} \cdot V_s.$$

If R_L is large compared to the other resistances (like the input to an operational amplifier), the output voltage can be approximated by the simpler equation:

$$V_\mathrm{L} = \frac{R_2}{R_1 + R_2} \cdot V_s.$$

As an example, assume

$V_\mathrm{S} = 10\ \mathrm{V}$, $R_1 = 1\ \mathrm{k\Omega}$, $R_2 = 2\ \mathrm{k\Omega}$, and $R_\mathrm{L} = 100\ \mathrm{k\Omega}.$

Since the load resistance is large compared to the other resistances, the output voltage V_L will be approximately:

$$\frac{2\ \mathrm{k\Omega}}{1\ \mathrm{k\Omega} + 2\ \mathrm{k\Omega}} \cdot 10\ \mathrm{V} = \frac{2}{3} \cdot 10\ \mathrm{V} \approx 6.667\ \mathrm{V}.$$

Due to the load resistance, however, it will actually be slightly lower: ≈ 6.623 V.

One of the advantages of the potential divider compared to a variable resistor in series with the source is that, while variable resistors have a maximum resistance where some current will always flow, dividers are able to vary the output voltage from maximum (V_S) to ground (zero volts) as the wiper moves from one end of the potentiometer to the other. There is, however, always a small amount of contact resistance.

In addition, the load resistance is often not known and therefore simply placing a variable resistor in series with the load could have a negligible effect or an excessive effect, depending on the load.

Early patents

- US patent 131,334, Thomas Edison, "Coiled resistance wire rheostat", issued 1872-9-17
- Mary Hallock-Greenewalt invented a type of nonlinear rheostat for use in her visual-music instrument, the Sarabet (US patent 1,357,773)

Chapter 6

Digital Potentiometer & Bleeder Resistor

Digital Potentiometer

A **digital potentiometer** is a digitally-controlled electronic component that mimics the analog functions of a potentiometer. It is often used for trimming and scaling analog signals by microcontrollers. It is either built using an R-2R integrated circuit or a Digital-to-analog converter. A digital potentiometer is an electronic component that is often controlled by digital protocols like I^2C and SPI, as well as more basic Up/Down protocols. Some typical uses of digital potentiometers are in circuits requiring gain control of amplifiers (frequently instrumentation amplifiers), small-signal audio-balancing, and offset adjustment.

Sometimes this device is also referred to as an RDAC, Resistive Digital-to-Analog Converter.

Some Digipots come with non-volatile memory, so that they retain their last programmed position after they have been power cycled. Most, though, are volatile, i.e. after they are power cycled they will default to a standard value, which is usually the mid-point.

The former can be useful, but when they are controlled by a microprocessor, or even via a Field Programmable Gate Array (FPGA), these devices can retain, in other non-volatile memory, the value to initialise the Digipot with. In these circumstances, the need for non-volatile Digipots is less obvious.

Limitations

These devices are extremely useful in the modern, digitally controlled world, but have some limitations. While quite similar to a normal potentiometer, digital potentiometers are somewhat constrained by current limits in the tens of milliamperes. Also, most, if not all digital potentiometers limit the input voltage range to the digital supply range (often 0–5 VDC), so some ingenuity is often required when attempting to replace standard resistive potentiometers with digital potentiometers. Further, instead of the seemingly continuous control that can be obtained from a multiturn resistive potentiometer, digital potentiometers have discrete steps in resistance. Eight-bit pots (256-steps) are most common, but potentiometers between 5 and 10 bits (32 to 1024 steps) are available. A fourth constraint is that special logic is often required to check for zero crossing of an analog AC signal to allow the resistance value to be changed without causing an audible click in the output for audio amplifiers.

The non-volatile Digipots also differ from their electro-mechanical cousins in that on power up, the resistance will default to (possibly) a different value after a power cycle.

Similarly, the Digipot resistance is only valid when the correct DC supply voltage(s) are present. When voltages are removed, the resistance between the two end points and the (nominal) Wiper are undefined. In an operational amplifier circuit, the Off-State impedance of a real potentiometer can help stabilise the DC operating point of the circuit during the power-up stage. This may not be the case when a Digipot is used.

Like their electro-mechanical counterparts, Digipots suffer similar weaknesses. Real potentiometers and Digipots generally have poor tolerances (typically +/- 20%), poor temperature coefficients (many hundreds of ppm per degree C), and a stop resistance that is typically about 0.5-1% of the full scale resistance. Note that Stop Resistance is the residual resistance when the terminal to wiper resistance is set to the minimum value.

Bleeder Resistor

A **bleeder resistor** is a resistor placed in parallel with a high-voltage supply for the purposes of discharging the energy stored in the power source's filter capacitors or other components that store electrical energy when the equipment is turned off.

It is a use for a standard resistor rather than a separate type of component.

Usage

DC power supplies

Power supplies, especially switchmode power supplies, use a bridge rectifier to convert mains AC power into a typical 340 volts DC for the chopper. A large filter capacitor typically stores enough energy at this high voltage to power the load during the zero crossings of the AC input. In fact, the capacitors in many supplies are large enough to support the load during AC outages lasting for a significant fraction of a second. This stored energy is clearly potentially lethal, and without a bleeder resistor it might remain long after the unit has been turned off. With a properly sized bleeder resistor, however, the voltage will quickly decay to safe levels when the supply is switched off, yet not consume too much power while the supply is on.

High voltage supply in television sets

The bleeder resistor commonly found inside a flyback transformer that supplies high voltage for a CRT is valued in the hundreds of megohms range, and can therefore not be measured with the common technician's multimeter.

Instead of a resistor inside the transformer, the focus and screen control array may be used for the same purpose, depending on the application and tolerances of the type of tube it is producing output for.

These bleeders discharge the focus supply, but not the high voltage final anode feed. The CRT itself forms a capacitor that can hold a sizable (and very dangerous) high voltage charge, so it is **always** advisable to momentarily ground a CRT's high voltage terminal before working on the unit.

Failure

The failure of a bleeder resistor prevents the discharge of the capacitors, resulting in dangerous voltages being retained for many days. This is one of several reasons for the typical warning on most equipment: "Warning - No user-serviceable parts inside". An un-suspecting user may get an electrical shock from opened equipment due to failure of a bleeder resistor, or the common practice of not fitting them.

Safe design suggests mounting a bleeder close to a dangerous capacitor, ideally directly to the capacitor terminals, and not through any connectors, so that it is difficult to disconnect the bleeder accidentally.

Despite the presence of a bleeder, it is wise to prove that any potentially dangerous capacitors are discharged, perhaps by shorting their terminals (or through a suitable low resistance for high energy capacitors), before working on any circuit.

Technical considerations

There is always a trade-off between the speed with which the bleeder operates and the amount of power wasted in the bleeder; a faster bleed-down rate wastes more power during normal, power-on operation.

The presence of a bleeder also guarantees a minimum load on the power source, which can help reduce the range of voltage change (regulation) when the normal load is changing and there is no active regulator. Use of a bleeder this way is a common design strategy for power supplies of vacuum tube power amplifiers, for instance.

Dual bleeder

Because of the speed/power tradeoff, high-powered circuits may use two separate bleeder circuits. A fast bleed circuit is switched out during normal operation so that no power is wasted; when power is switched off, the fast bleeder is connected, rapidly bleeding down the voltage. The switch controlling the fast bleeder can fail, either by connecting when it shouldn't (and overheating) or by not connecting when it should (and thereby failing to bleed off the voltage quickly). To avoid the risk of not having an operational bleeder, a secondary, slower (and less lossy) bleeder is usually permanently connected so that there is always some bleed-down capability.

Chapter 7

Attenuator (Electronics)

A 30dB 5W RF-attenuator, DC-18GHz, with N-type coaxial connectors

Coaxial Dynamics 100 Watt power attenuator

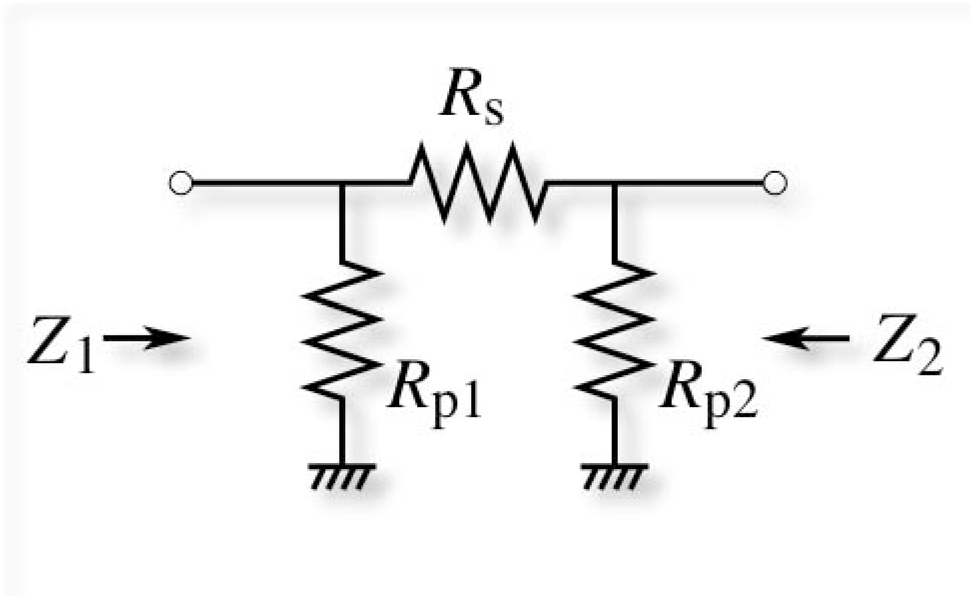
An **attenuator** is an electronic device that reduces the amplitude or power of a signal without appreciably distorting its waveform.

An attenuator is effectively the opposite of an amplifier, though the two work by different methods. While an amplifier provides gain, an attenuator provides loss, or gain less than 1.

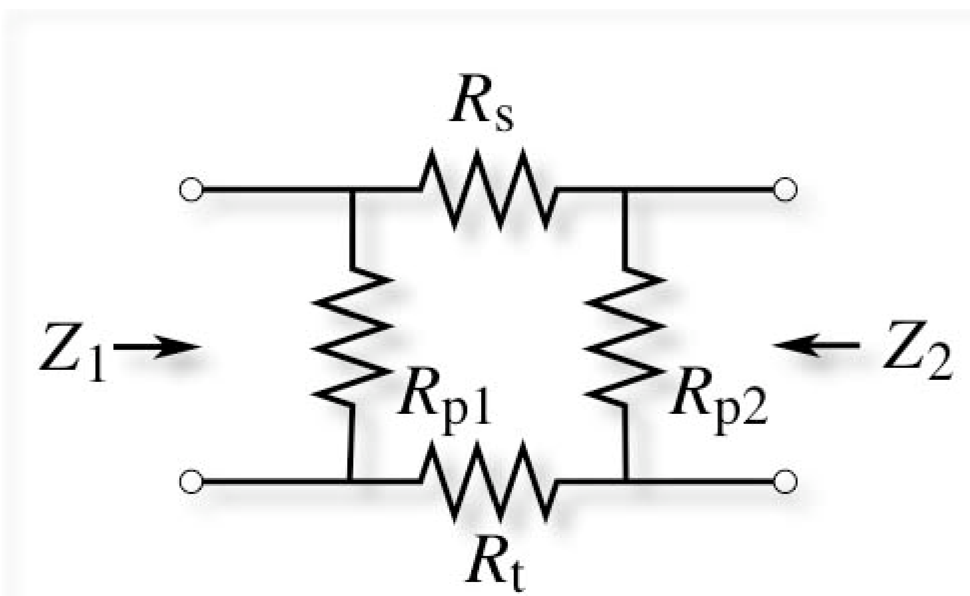
Attenuators are usually passive devices made from simple voltage divider networks. Switching between different resistances forms adjustable stepped attenuators and continuously adjustable ones using potentiometers. For higher frequencies precisely matched low VSWR resistance networks are used.

Fixed attenuators in circuits are used to lower voltage, dissipate power, and to improve impedance matching. In measuring signals, attenuator pads or adaptors are used to lower the amplitude of the signal a known amount to enable measurements, or to protect the measuring device from signal levels that might damage it. Attenuators are also used to 'match' impedances by lowering apparent SWR.

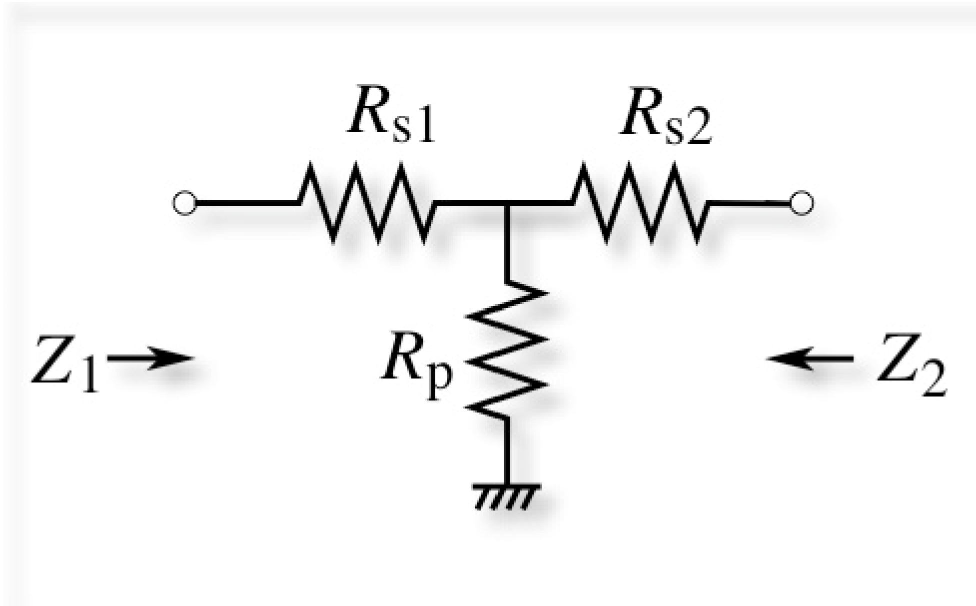
Attenuator circuits



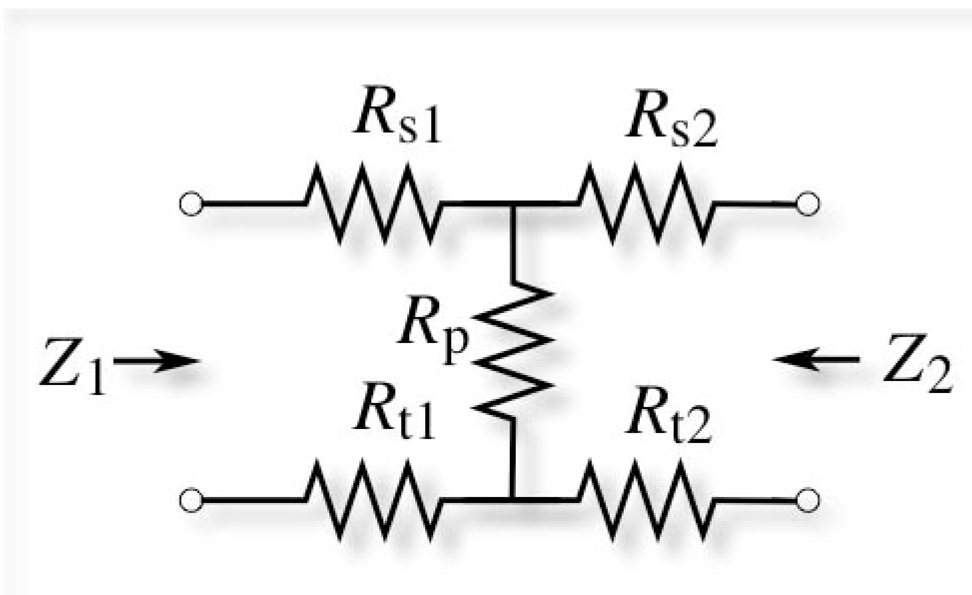
π-type unbalanced attenuator circuit



π-type balanced attenuator circuit



T-type unbalanced attenuator circuit



T-type balanced attenuator circuit

Basic circuits used in attenuators are pi pads (π-type) and T pads. These may be required to be balanced or unbalanced networks depending on whether the line geometry with which they are to be used is balanced or unbalanced. For instance, attenuators used with coaxial lines would be the unbalanced form while attenuators for use with twisted pair are required to be the balanced form.

Four fundamental attenuator circuit diagrams are given in the figures on the left. Since an attenuator circuit consists solely of passive resistor elements, it is linear and reciprocal. If the circuit is also made symmetrical (this is usually the case since it is usually required that the input and output impedances Z_1 and Z_2 are equal) then the input and output ports are not distinguished, but by convention the left and right sides of the circuits are referred to as input and output.

Attenuator characteristics

A RF Microwave Attenuator. Picture courtesy of Herley

Key specifications for attenuators are:

- **Attenuation** expressed in decibels of relative power. As a rule of thumb, a 3dB pad reduces power to one half, 6dB to one fourth, 10dB to one tenth, 20dB to one hundredth, 30dB to one thousandth and so on. For voltage you double the dBs so for example 6dB is half in voltage.
- **Frequency bandwidth**, for example DC-18 GHz
- **Power dissipation** depends on mass and surface area of resistance material as well as possible additional cooling fins.
- **SWR** is the standing wave ratio for input and output ports
- **Accuracy**
- **Repeatability**

RF attenuators

Radio frequency attenuators are typically coaxial in structure with precision connectors as ports and coaxial, microstrip or thin-film internal structure. Above SHF special waveguide structure is required.

Important characteristics are:

- accuracy,
- low SWR,
- flat frequency-response and
- repeatability.

The size and shape of the attenuator depends on its ability to dissipate power. RF attenuators are used as loads for and as known attenuations and protective dissipations of power in measuring RF signals.

Audio attenuators

A line-level attenuator in the preamp or a power attenuator after the power amplifier uses electrical resistance to reduce the amplitude of the signal that reaches the speaker, reducing the volume of the output. A line-level attenuator has lower power handling, such as a 1/2-watt potentiometer or voltage divider and controls preamp level signals, whereas a power attenuator has higher power handling capability, such as 10 watts or more, and is used between the power amplifier and the speaker.

- Power attenuator (guitar)
- Guitar amplifier

Component values for resistive pads and attenuators

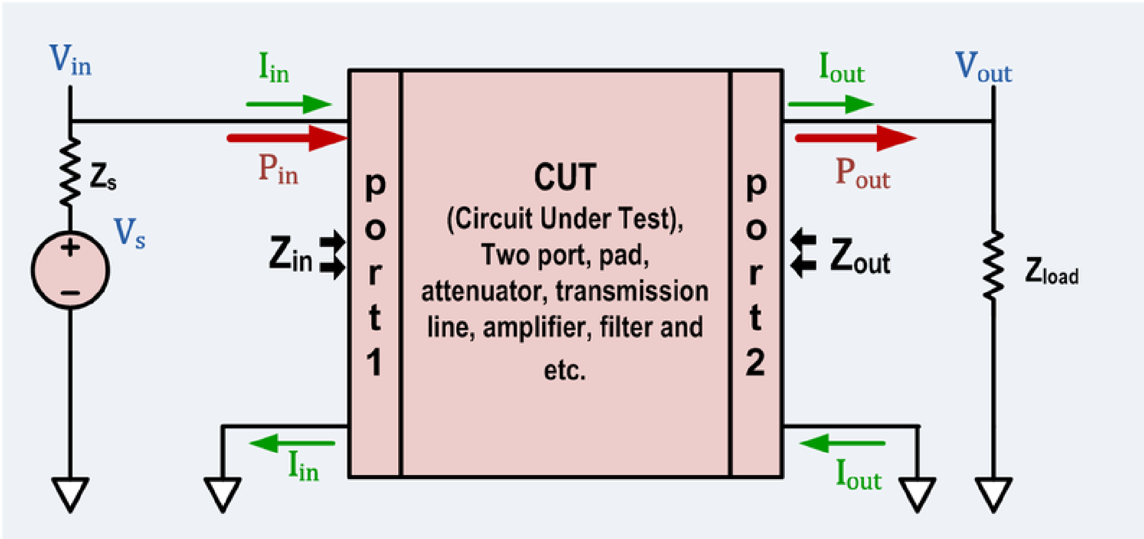
This section concerns pi-pads, T-pads and L-pads made entirely from resistors and terminated on each port with a purely real resistance.

- All impedances, currents, voltages and two-port parameters will be assumed to be purely real. For practical applications, this assumption is often close enough.

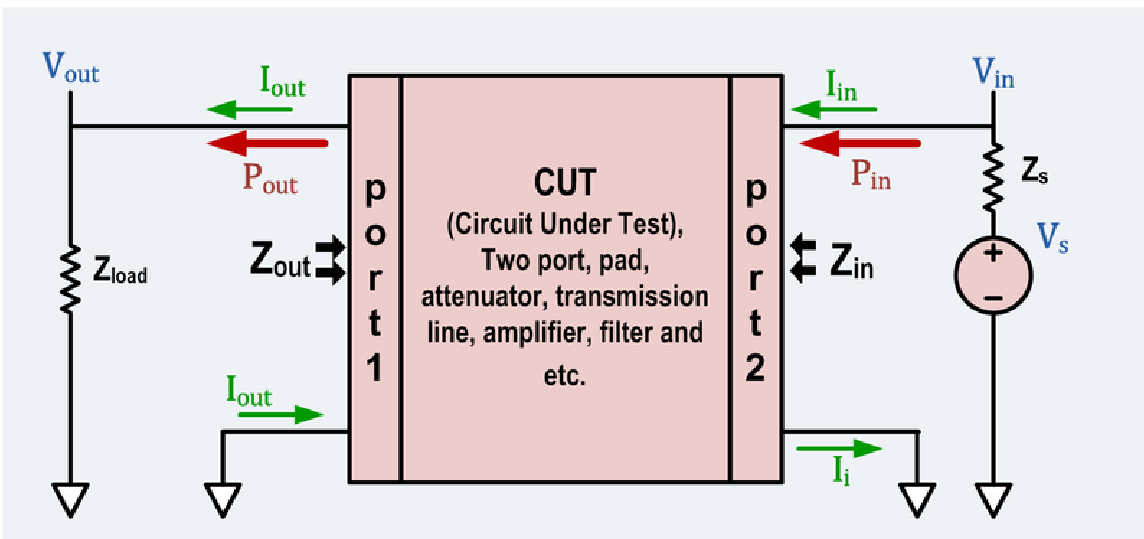
- The pad is designed for a particular load impedance, Z_{Load}, and a particular source impedance, Z_s.

  - The impedance seen looking into the input port will be Z_S if the output port is terminated by Z_{Load}.
  - The impedance seen looking into the output port will be Z_{Load} if the input port is terminated by Z_S.

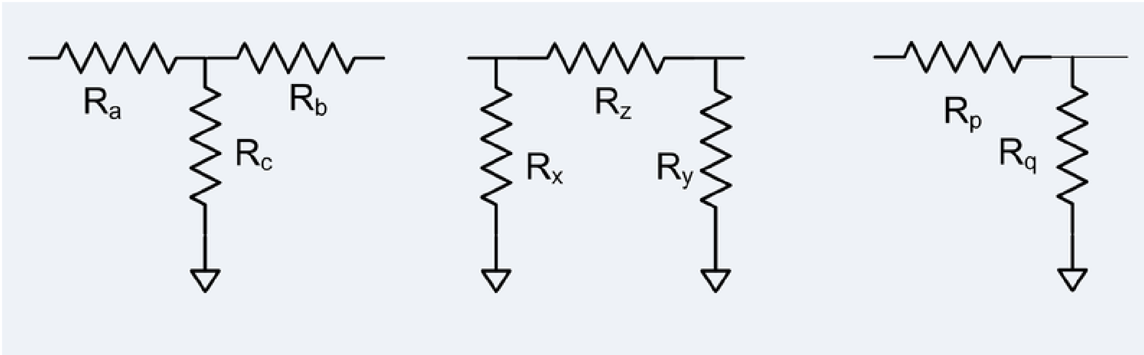
Reference figures for attenuator component calculation



This circuit is used for the general case, all T-pads, all pi-pads and L-pads when the source impedance is greater than or equal to the load impedance.



The L-pad computation assumes that port 1 has the highest impedance. If the highest impedance happens to be the output port, then use this figure.



Unique resistor designations for Tee, Pi and L pads.

The attenuator two-port is generally bidirectional. However here it will be treated as though it were one way. In general, either of the two figures above applies, but the figure on the left (which depicts the source on the left) will be tacitly assumed most of the time. In the case of the L-pad, the right figure will be used if the load impedance is greater than the source impedance.

Each resistor in each type of pad discussed is given a unique designation to decrease confusion.

The L-pad component value calculation assumes that the design impedance for port 1 (on the left) is equal or higher than the design impedance for port 2.

Terms used

- Pad will include pi-pad, T-pad, L-pad, attenuator, and two-port.
- Two-port will include pi-pad, T-pad, L-pad, attenuator, and two-port.
- Input port will mean the input port of the two-port.
- Output port will mean the output port of the two-port.
- Symmetric means a case where the source and load have equal impedance.
- Loss means the ratio of power entering the input port of the pad divided by the power absorbed by the load.
- Insertion Loss means the ratio of power that would be delivered to the load if the load were directly connected to the source divided by the power absorbed by the load when connected through the pad.

Symbols used

Passive, resistive pads and attenuators are bidirectional two-ports, but here they will be treated as unidirectional.

- Z_S = the output impedance of the source.
- Z_{Load} = the input impedance of the load.

- Z_{in} = the impedance seen looking into the input port when Z_{Load} is connected to the output port. Z_{in} is a function of the load impedance.
- Z_{out} = the impedance seen looking into the output port when Z_s is connected to the input port. Z_{out} is a function of the source impedance.
- V_s = source open circuit or unloaded voltage.
- V_{in} = voltage applied to the input port by the source.
- V_{out} = voltage applied to the load by the output port.
- I_{in} = current entering the input port from the source.
- I_{out} = current entering the load from the output port.
- $P_{in} = V_{in} I_{in}$ = power entering the input port from the source.
- $P_{out} = V_{out} I_{out}$ = power absorbed by the load from the output port.
- P_{direct} = the power that would be absorbed by the load if the load were connected directly to the source.
- $L_{pad} = 10 \log_{10} (P_{in} / P_{out})$ always. And if $Z_s = Z_{Load}$ then $L_{pad} = 20 \log_{10} (V_{in} / V_{out})$ also. Note, as defined, Loss ≥ 0 dB
- $L_{insertion} = 10 \log_{10} (P_{direct} / P_{out})$. And if $Z_s = Z_{Load}$ then $L_{insertion} = L_{pad}$.
- Loss ≡ L_{pad}. Loss is defined to be L_{pad}.

Symmetric T pad resistor calculation

$$A = 10^{-Loss/20} \qquad R_a = R_b = Z_S\frac{1-A}{1+A} \qquad R_c = \frac{Z_s^2 - R_b^2}{2R_b}$$

Symmetric pi pad resistor calculation

$$A = 10^{-Loss/20} \qquad R_x = R_y = Z_S\frac{1+A}{1-A} \qquad R_z = \frac{2R_x}{\left(\frac{R_x}{Z_0}\right)^2 - 1}]$$

L-Pad for impedance matching resistor calculation

If a source and load are both resistive (i.e. Z_1 and Z_2 have zero or very small imaginary part) then a resistive L-pad can be used to match them to each other. As shown, either side of the L-pad can be the source or load, but the Z_1 side must be the side with the higher impedance.

$$R_q = \frac{Z_m}{\sqrt{\rho - 1}} \qquad R_p = Z_m\sqrt{\rho - 1} \qquad Loss = 20\log_{10}\left(\sqrt{\rho-1} + \sqrt{\rho}\right) \quad \text{where} \quad \rho = \frac{Z_1}{Z_2} \quad Z_m = \sqrt{Z_1 Z_2}$$

Large positive numbers means loss is large. The loss is a monotonic function of the impedance ratio. Higher ratios require higher loss.

Converting T-pad to pi-pad

$$R_z = \frac{R_aR_b + R_aR_c + R_bR_c}{R_c} \qquad R_x = \frac{R_aR_b + R_aR_c + R_bR_c}{R_b} \qquad R_y = \frac{R_aR_b + R_aR_c + R_bR_c}{R_a}.$$

Converting pi-pad to T-pad

$$R_c = \frac{R_xR_y}{R_x + R_y + R_z} \qquad R_a = \frac{R_zR_x}{R_x + R_y + R_z} \qquad R_b = \frac{R_zR_y}{R_x + R_y + R_z}$$

Conversion between two-ports and pads

T-pad to impedance parameters

The impedance parameters for a passive two-port are

$$V_1 = Z_{11}I_1 + Z_{12}I_2 \qquad V_2 = Z_{21}I_1 + Z_{22}I_2 \qquad with \qquad Z_{12} = Z_{21}$$

It is always possible to represent a resistive t-pad as a two-port. The representation is particularly simple using impedance parameters as follows:

$$Z_{21} = R_c \qquad Z_{11} = R_c + R_a \qquad Z_{22} = R_c + R_b$$

Impedance parameters to T-pad

The preceding equations are trivially invertible, but if the loss is not enough, some of the t-pad components will have negative resistances.

$$R_c = Z_{21} \qquad R_a = Z_{11} - Z_{21} \qquad R_b = Z_{22} - Z_{21}$$

Impedance parameters to pi-pad

These preceding T-pad parameters can be algebraically converted to pi-pad parameters.

$$R_z = \frac{Z_{11}Z_{22} - Z_{21}^2}{Z_{21}} \qquad R_x = \frac{Z_{11}Z_{22} - Z_{21}^2}{Z_{22} - Z_{21}} \qquad R_y = \frac{Z_{11}Z_{22} - Z_{21}^2}{Z_{11} - Z_{21}}$$

Pi-pad to admittance parameters

The admittance parameters for a passive twp port are

$$I_1 = Y_{11}V_1 + Y_{12}V_2 \qquad I_2 = Y_{21}V_1 + Y_{22}V_2 \qquad with \qquad Y_{12} = Y_{21}$$

It is always possible to represent a resistive pi pad as a two-port. The representation is particularly simple using admittance parameters as follows:

$$Y_{21} = \frac{1}{R_z} \qquad Y_{11} = \frac{1}{R_x} + \frac{1}{R_z} \qquad Y_{22} = \frac{1}{R_y} + \frac{1}{R_z}$$

Admittance parameters to pi-pad

The preceding equations are trivially invertible, but if the loss is not enough, some of the pi-pad components will have negative resistances.

$$R_z = \frac{1}{Y_{21}} \qquad R_x = \frac{1}{Y_{11} - Y_{21}} \qquad R_y = \frac{1}{Y_{22} - Y_{21}}$$

General case, determining impedance parameters from requirements

Because the pad is entirely made from resistors, it must have a certain minimum loss to match source and load if they are not equal.

The minimum loss is given by

$$Loss_{min} = 20\ log_{10}\left(\sqrt{\rho - 1} + \sqrt{\rho}\quad\right) \quad \text{where} \quad \rho = \frac{\max[Z_S, Z_{Load}]}{\min[Z_S, Z_{Load}]}$$

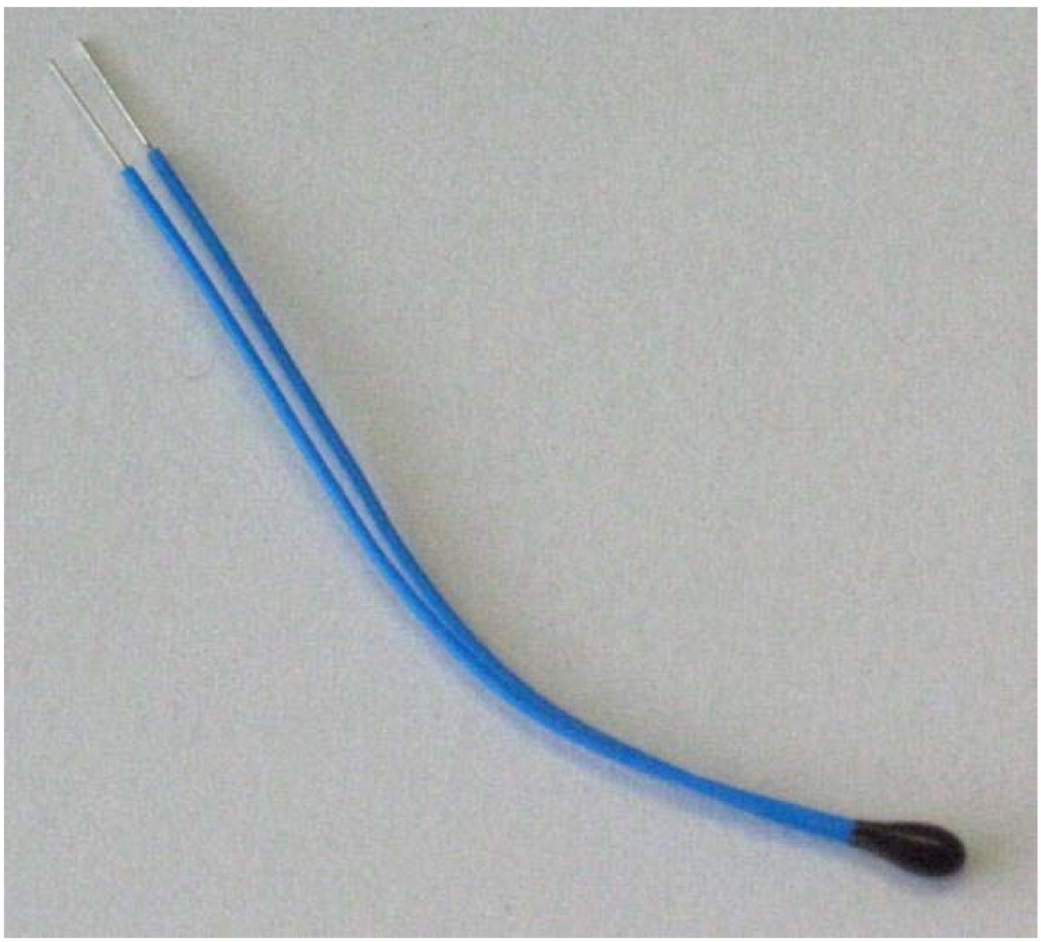
Although a passive matching two-port can have less loss, if it does it will not be convertible to a resistive attenuator pad.

$$A = 10^{-Loss/20} \qquad Z_{11} = Z_S \frac{1 + A^2}{1 - A^2} \qquad Z_{22} = Z_{Load} \frac{1 + A^2}{1 - A^2} \qquad Z_{21} = 2 \frac{A\sqrt{Z_S Z_{Load}}}{1 - A^2}$$

Once these parameters have been determined, they can be implemented as a T or pi pad as discussed above.

Chapter 8

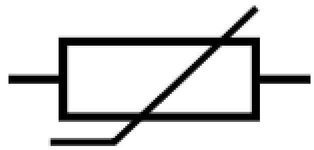
Thermistor



NTC thermistor, bead type, insulated wires

A **thermistor** is a type of resistor whose resistance varies significantly with temperature, more so than in standard resistors. The word is a portmanteau of *thermal* and *resistor*. Thermistors are widely used as inrush current limiters, temperature sensors, self-resetting overcurrent protectors, and self-regulating heating elements.

Thermistors differ from resistance temperature detectors (RTD) in that the material used in a thermistor is generally a ceramic or polymer, while RTDs use pure metals. The temperature response is also different; RTDs are useful over larger temperature ranges, while thermistors typically achieve a higher precision within a limited temperature range [usually −90 °C to 130 °C].



Thermistor symbol

Assuming, as a first-order approximation, that the relationship between resistance and temperature is linear, then:

$$\Delta R = k\Delta T$$

where

ΔR = change in resistance
ΔT = change in temperature
k = first-order temperature coefficient of resistance

Thermistors can be classified into two types, depending on the sign of k. If k is positive, the resistance increases with increasing temperature, and the device is called a positive temperature coefficient (**PTC**) thermistor, or **posistor**. If k is negative, the resistance decreases with increasing temperature, and the device is called a negative temperature coefficient (**NTC**) thermistor. Resistors that are not thermistors are designed to have a k as close to zero as possible(smallest possible k), so that their resistance remains nearly constant over a wide temperature range.

Instead of the temperature coefficient k, sometimes the *temperature coefficient of resistance* α (alpha) or α_T is used. It is defined as

$$\alpha_T = \frac{1}{R(T)}\frac{dR}{dT}.$$

For example, for the common PT100 sensor, α = 0.00385 or 0.385 %/°C. This α_T coefficient should not be confused with the α parameter below.

Steinhart-Hart equation

In practice, the linear approximation (above) works only over a small temperature range. For accurate temperature measurements, the resistance/temperature curve of the device must be described in more detail. The Steinhart-Hart equation is a widely used third-order approximation:

$$\frac{1}{T} = a + b\,\ln(R) + c\,\ln^3(R)$$

where *a*, *b* and *c* are called the Steinhart-Hart parameters, and must be specified for each device. *T* is the temperature in kelvin and *R* is the resistance in ohms. To give resistance as a function of temperature, the above can be rearranged into:

$$R = e^{\left(x-\frac{y}{2}\right)^{\frac{1}{3}} - \left(x+\frac{y}{2}\right)^{\frac{1}{3}}}$$

where

$$y = \frac{a - \frac{1}{T}}{c} \quad \text{and} \quad x = \sqrt{\left(\frac{b}{3c}\right)^3 + \frac{y^2}{4}}$$

The error in the Steinhart-Hart equation is generally less than 0.02 °C in the measurement of temperature. As an example, typical values for a thermistor with a resistance of 3000 Ω at room temperature (25 °C = 298.15 K) are:

$$a = 1.40 \times 10^{-3}$$
$$b = 2.37 \times 10^{-4}$$
$$c = 9.90 \times 10^{-8}$$

B *parameter equation*

NTC thermistors can also be characterised with the *B* parameter equation, which is essentially the Steinhart Hart equation with $a = (1 / T_0) - (1 / B)ln(R_0)$, $b = 1 / B$ and $c = 0$,

$$\frac{1}{T} = \frac{1}{T_0} + \frac{1}{B}\ln\left(\frac{R}{R_0}\right)$$

Where the temperatures are in kelvins and R_0 is the resistance at temperature T_0 (usually 25 °C = 298.15 K). Solving for *R* yields:

$$R = R_0 e^{B(1/T - 1/T_0)}$$

or, alternatively,

$$R = r_\infty e^{B/T}$$

where $r_\infty = R_0 e^{-B/T_0}$. This can be solved for the temperature:

$$T = \frac{B}{\ln\left(R/r_\infty\right)}$$

The B-parameter equation can also be written as $\ln R = B/T + \ln r_\infty$. This can be used to convert the function of resistance vs. temperature of a thermistor into a linear function of ln*R* vs. 1 / *T*. The average slope of this function will then yield an estimate of the value of the *B* parameter.

Conduction model

Many NTC thermistors are made from a pressed disc or cast chip of a semiconductor such as a sintered metal oxide. They work because raising the temperature of a semiconductor increases the number of electrons able to move about and carry charge - it promotes them into the *conduction band*. The more charge carriers that are available, the more current a material can conduct. This is described in the formula:

$$I = n \cdot A \cdot v \cdot e$$

I = electric current (amperes)
n = density of charge carriers (count/m^3)
A = cross-sectional area of the material (m^2)
v = velocity of charge carriers (m/s)
e = charge of an electron ($e = 1.602 \times 10^{-19}$ coulomb)

The current is measured using an ammeter. Over large changes in temperature, calibration is necessary. Over small changes in temperature, if the right semiconductor is used, the resistance of the material is linearly proportional to the temperature. There are many different semiconducting thermistors with a range from about 0.01 kelvin to 2,000 kelvins (−273.14 °C to 1,700 °C).

Most PTC thermistors are of the "switching" type, which means that their resistance rises suddenly at a certain critical temperature. The devices are made of a doped polycrystalline ceramic containing barium titanate ($BaTiO_3$) and other compounds. The dielectric constant of this ferroelectric material varies with temperature. Below the Curie point temperature, the high dielectric constant prevents the formation of potential barriers between the crystal grains, leading to a low resistance. In this region the device has a small negative temperature coefficient. At the Curie point temperature, the dielectric constant drops sufficiently to allow the formation of potential barriers at the grain boundaries, and the resistance increases sharply. At even higher temperatures, the material reverts to NTC behaviour. The equations used for modeling this behaviour were derived by W. Heywang and G. H. Jonker in the 1960s.

Another type of PTC thermistor is the polymer PTC, which is sold under brand names such as "Polyswitch" "Semifuse", and "Multifuse". This consists of a slice of plastic with carbon grains embedded in it. When the plastic is cool, the carbon grains are all in contact with each other, forming a conductive path through the device. When the plastic heats up, it expands, forcing the carbon grains apart, and causing the resistance of the device to rise rapidly. Like the $BaTiO_3$ thermistor, this device has a highly nonlinear resistance/temperature response and is used for switching, not for proportional temperature measurement.

Yet another type of thermistor is a **silistor**, a thermally sensitive silicon resistor. Silistors are similarly constructed and operate on the same principles as other thermistors, but employ silicon as the semiconductive component material.

Self-heating effects

When a current flows through a thermistor, it will generate heat which will raise the temperature of the thermistor above that of its environment. If the thermistor is being used to measure the temperature of the environment, this electrical heating may introduce a significant error if a correction is not made. Alternatively, this effect itself can be exploited. It can, for example, make a sensitive air-flow device employed in a sailplane rate-of-climb instrument, the electronic variometer, or serve as a timer for a relay as was formerly done in telephone exchanges.

The electrical power input to the thermistor is just:

$$P_E = IV$$

where I is current and V is the voltage drop across the thermistor. This power is converted to heat, and this heat energy is transferred to the surrounding environment. The rate of transfer is well described by Newton's law of cooling:

$$P_T = K(T(R) - T_0)$$

where $T(R)$ is the temperature of the thermistor as a function of its resistance R, T_0 is the temperature of the surroundings, and K is the **dissipation constant**, usually expressed in units of milliwatts per degree Celsius. At equilibrium, the two rates must be equal.

$$P_E = P_T$$

The current and voltage across the thermistor will depend on the particular circuit configuration. As a simple example, if the voltage across the thermistor is held fixed, then by Ohm's Law we have $I = V / R$ and the equilibrium equation can be solved for the ambient temperature as a function of the measured resistance of the thermistor:

$$T_0 = T(R) - \frac{V^2}{KR}$$

The dissipation constant is a measure of the thermal connection of the thermistor to its surroundings. It is generally given for the thermistor in still air, and in well-stirred oil. Typical values for a small glass bead thermistor are 1.5 mW/°C in still air and 6.0 mW/°C in stirred oil. If the temperature of the environment is known beforehand, then a thermistor may be used to measure the value of the dissipation constant. For example, the thermistor may be used as a flow rate sensor, since the dissipation constant increases with the rate of flow of a fluid past the thermistor.

Applications

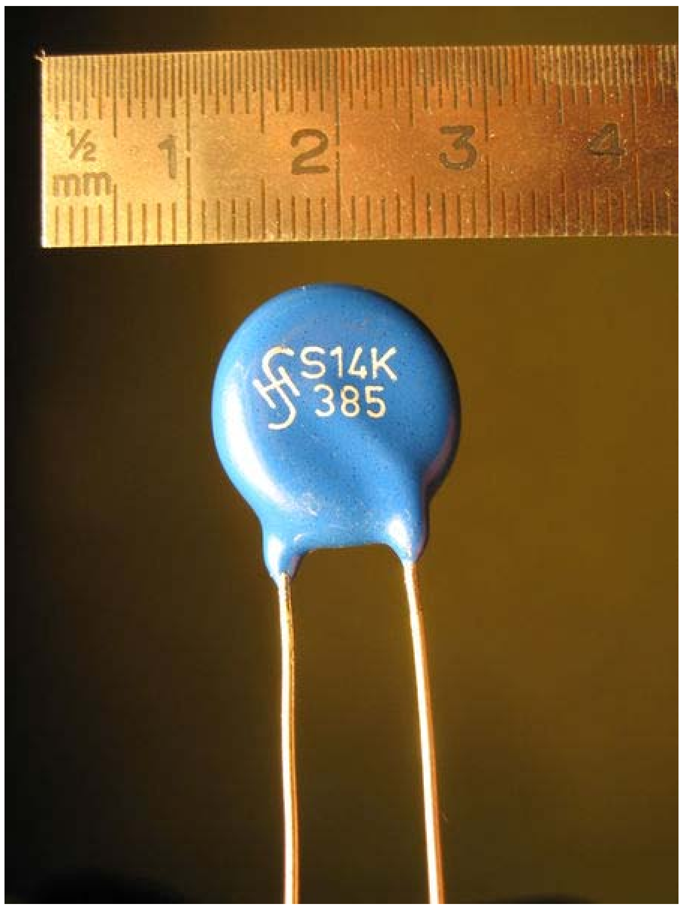
- PTC thermistors can be used as current-limiting devices for circuit protection, as replacements for fuses. Current through the device causes a small amount of resistive heating. If the current is large enough to generate more heat than the device can lose to its surroundings, the device heats up, causing its resistance to increase, and therefore causing even more heating. This creates a self-reinforcing effect that drives the resistance upwards, reducing the current and voltage available to the device.
- PTC thermistors are used as timers in the degaussing coil circuit of most CRT displays and televisions. When the display unit is initially switched on, current flows through the thermistor and degaussing coil. The coil and thermistor are intentionally sized so that the current flow will heat the thermistor to the point that the degaussing coil shuts off in under a second. For effective degaussing, it is necessary that the magnitude of the alternating magnetic field produced by the degaussing coil decreases smoothly and continuously, rather than sharply switching off or decreasing in steps; the PTC thermistor accomplishes this naturally as it heats up. A degaussing circuit using a PTC thermistor is simple, reliable (for its simplicity), and inexpensive.
- NTC thermistors are used as resistance thermometers in low-temperature measurements of the order of 10 K.
- NTC thermistors can be used as inrush-current limiting devices in power supply circuits. They present a higher resistance initially which prevents large currents from flowing at turn-on, and then heat up and become much lower resistance to allow higher current flow during normal operation. These thermistors are usually much larger than measuring type thermistors, and are purposely designed for this application.
- NTC thermistors are regularly used in automotive applications. For example, they monitor things like coolant temperature and/or oil temperature inside the engine and provide data to the ECU and, indirectly, to the dashboard.
- NTC thermistors can be also used to monitor the temperature of an incubator.
- Thermistors are also commonly used in modern digital thermostats and to monitor the temperature of battery packs while charging.

History

The first NTC thermistor was discovered in 1833 by Michael Faraday, who reported on the semiconducting behavior of silver sulfide. Faraday noticed that the resistance of silver sulfide decreased dramatically as temperature increased. Because early thermistors were difficult to produce and applications for the technology were limited, commercial production of thermistors did not begin until the 1930s.

Chapter 9

Varistor



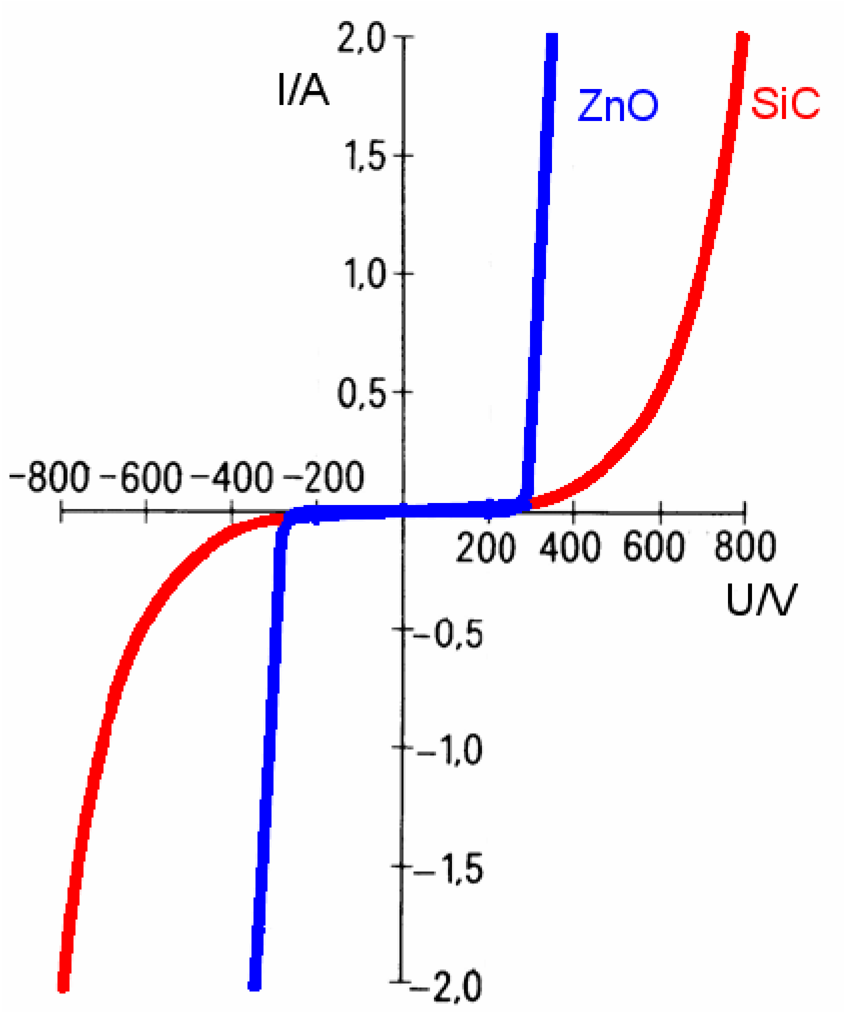
A 385-volt metal oxide varistor

A **varistor** is an electronic component with a "diode-like" nonlinear current–voltage characteristic. The name is a portmanteau of *variable resistor*. Varistors are often used to protect circuits against excessive transient voltages by incorporating them into the circuit in such a way that, when triggered, they will shunt the current created by the high voltage away from the sensitive components. A varistor is also known as *Voltage Dependent Resistor* or **VDR**. A varistor's function is to conduct significantly increased current when voltage is excessive.

Note: only non-ohmic variable resistors are usually called varistors*. Other, ohmic types of variable resistor include the potentiometer and the rheostat.*

Metal oxide varistor

The most common type of varistor is the **Metal Oxide Varistor** (MOV). This contains a ceramic mass of zinc oxide grains, in a matrix of other metal oxides (such as small amounts of bismuth, cobalt, manganese) sandwiched between two metal plates (the electrodes). The boundary between each grain and its neighbour forms a diode junction, which allows current to flow in only one direction. The mass of randomly oriented grains is electrically equivalent to a network of back-to-back diode pairs, each pair in parallel with many other pairs. When a small or moderate voltage is applied across the electrodes, only a tiny current flows, caused by reverse leakage through the diode junctions. When a large voltage is applied, the diode junction breaks down due to a combination of thermionic emission and electron tunneling, and a large current flows. The result of this behavior is a highly nonlinear current-voltage characteristic, in which the MOV has a high resistance at low voltages and a low resistance at high voltages.



Varistor current-voltage characteristic

Follow-through current as a result of a lightning strike may generate excessive current that permanently damages a varistor. In general, the primary case of varistor breakdown is localized heating caused as an effect of thermal runaway. This is due to a lack of conformality in individual grain-boundary junctions, which leads to the failure of dominant current paths under thermal stress.

Varistors can absorb part of a surge. How much effect this has on risk to connected equipment depends on the equipment and details of the selected varistor. Varistors do not absorb a significant percentage of a lightning strike, as energy that must be conducted elsewhere is many orders of magnitude greater than what is absorbed by the small device.

A varistor remains non-conductive as a shunt mode device during normal operation when voltage remains well below its "clamping voltage". If a transient pulse (often measured in joules) is too high, the device may melt, burn, vaporize, or otherwise be damaged or destroyed. This (catastrophic) failure occurs when "Absolute Maximum Ratings" in manufacturer's datasheet are significantly exceeded. Varistor degradation is defined by manufacturer's life expectancy charts using curves that relate current, time, and number of transient pulses. A varistor fully degrades typically when its "clamping voltage" has changed by 10%. A fully degraded varistor remains functional (no catastrophic failure) and is not visibly damaged.

Ballpark number for varistor life expectancy is its energy rating. As MOV joules increase, the number of transient pulses increases and the "clamping voltage" during each transient decreases. The purpose of this shunt mode device is to divert a transient so that pulse energy will be dissipated elsewhere. Some energy is also absorbed by the varistor because a varistor is not a perfect conductor. Less energy is absorbed by a varistor, the varistor is more conductive, and its life expectancy increases exponentially as varistor energy rating is increased. Catastrophic failure can be avoided by significantly increasing varistor energy ratings either by using a varistor of higher joules or by connecting more of these shunt mode devices in parallel.

Important parameters are the varistor's energy rating in joules, operating voltage, response time, maximum current, and breakdown (clamping) voltage. Energy rating is often defined using standardized transients such as 8/20 microseconds or 10/1000 microseconds, where 8 microseconds is the transient's front time and 20 microseconds is the time to half value.

To protect communications lines (such as telephone lines) transient suppression devices such as 3 mil carbon blocks (IEEE C62.32), ultra-low capacitance varistors or avalanche diodes are used. For higher frequencies such as radio communication equipment, a gas discharge tube (GDT) may be utilized.

A typical surge protector power strip is built using MOVs. A cheapest kind may use just one varistor, from hot (live, active) to neutral. A better protector would contain at least three varistors; one across each of the three pairs of conductors (hot-neutral, hot-ground, neutral-ground). A power strip protector in the United States should have a UL1449 3rd edition approval so that catastrophic MOV failure would not create a fire hazard.

High voltage varistor

Hazards

While a MOV is designed to conduct significant power for very short durations (≈ 8/20 microseconds), such as caused by lightning strikes, it typically does not have the capacity to conduct sustained energy. Under normal utility voltage conditions, this is not a problem. However, certain types of faults on the utility power grid can result in sustained over-voltage conditions. Examples include a loss of a neutral conductor or shorted lines on the high voltage system. Application of sustained over-voltage to a MOV can cause high dissipation, potentially resulting in the MOV device catching fire. The National Fire Protection Association (NFPA) has documented many cases of catastrophic fires that have been caused by MOV devices in surge suppressors, and has issued bulletins on the issue.

A series connected thermal fuse is one solution to catastrophic MOV failure. Varistors with internal thermal protection are also available.

There are several issues to be noted regarding behavior of transient voltage surge suppressors (TVSS) incorporating MOVs under over-voltage conditions. Depending on the level of conducted current, dissipated heat may be insufficient to cause failure, but may degrade the MOV device and reduce its life expectancy. If excessive current is conducted by a MOV, it may explode inside the case, keeping the load connected but now without any surge protection. A user may have no indication when the surge suppressor has failed. Under the right conditions of over-voltage and line impedance, it may be possible to cause the MOV to burst into flames, the root cause of many fires and the main reason for NFPA's concern. Properly designed TVSS devices should contain the flames, eventually resulting in the opening of a safety fuse.

What varistors don't do

A MOV inside a TVSS device does not provide equipment with complete power protection. In particular, MOV device provide no protection for the connected equipment from sustained over-voltages that may result in damage to that equipment as well as to the protector device.

A varistor provides no equipment protection from inrush current surges (during equipment startup), from overcurrent (created by a short circuit), or from voltage sags (also known as a brownout). A varistor neither senses nor controls such events. Susceptibility of electronic equipment to these other power disturbances is defined by equipment design. Protection from these power disturbances is installed inside that equipment or is provided by other external devices such as an UPS, some voltage regulators and Surge Protectors with built in overvoltage protection that make use of a voltage sensing circuit and a relay for disconnecting the AC input when voltage reaches a danger threshold.

Varistors compared to other transient suppressors

The response time of the MOV is largely ambiguous, as no standard has been officially defined. The sub-nanosecond MOV response claim is based on the material's intrinsic response time, but will be slowed down by other factors such as the inductance of component leads and the mounting method. That response time is also qualified as insignificant when compared to a transient having an 8 µs rise-time, thereby allowing ample time for the device to slowly turn-on. When subjected to a very fast, <1 ns rise-time transient, response times for the MOV are in the 40-60 ns range.

Typical capacitance for consumer-sized (7–20 mm diameter) varistors are in the range of 100-1,000 pF. Smaller, lower-capacitance varistors are available with capacitance of ~1 pF for microelectronic protection, such as in cellular phones. These low-capacitance varistors are, however, unable to withstand large surge currents simply due to their compact PCB-mount size.

Another method for suppressing voltage spikes is the transient voltage suppression diode (TVS). Although diodes do not have as much capacity to conduct large surges as MOVs, diodes are not degraded by smaller surges and can be implemented with a lower "clamping voltage". MOVs degrade from repeated exposure to surges and generally have a higher "clamping voltage" so that leakage does not degrade the MOV. Both types are available over a wide range of voltages. MOVs tend to be more suitable for higher voltages, because they can conduct the higher associated energies at less cost.

Another type of transient suppressor is the gas tube suppressor. This is a type of spark gap that may use air or an inert gas mixture and often, a small amount of radioactive material such as Ni-63, to provide a more consistent breakdown voltage and reduce response time. Unfortunately, these devices may have higher breakdown voltages and longer response times than varistors. However, they can handle significantly higher fault currents and withstand multiple high-voltage hits (for example, from lightning) without significant degradation.

Chapter 10

Resistance Thermometer

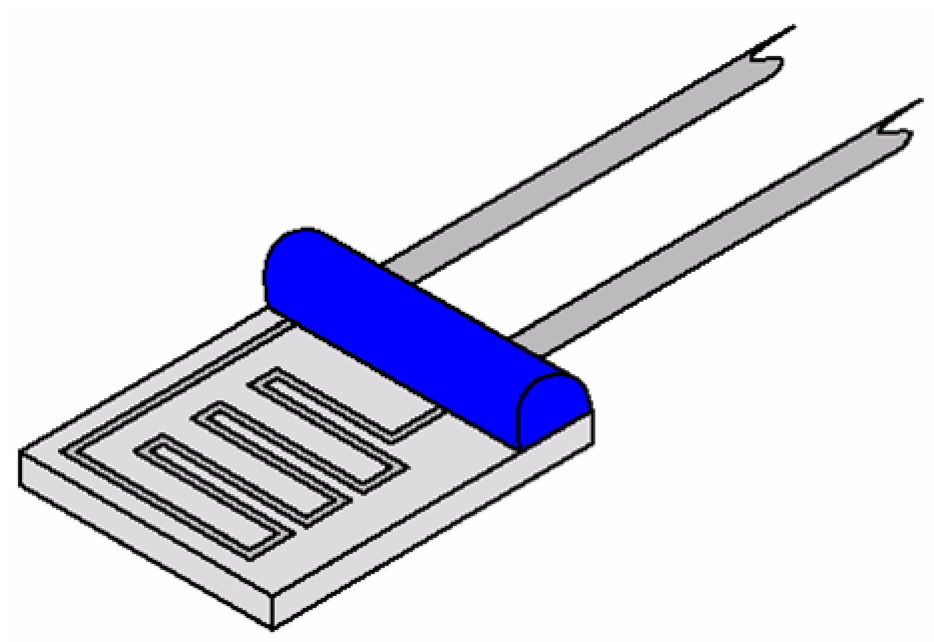
Resistance thermometers, also called **resistance temperature detectors** or **resistive thermal devices** (**RTD**s), are temperature sensors that exploit the predictable change in electrical resistance of some materials with changing temperature. As they are almost invariably made of platinum, they are often called **platinum resistance thermometers** (**PRT**s). They are slowly replacing the use of thermocouples in many industrial applications below 600 °C, due to higher accuracy and repeatability.

General description

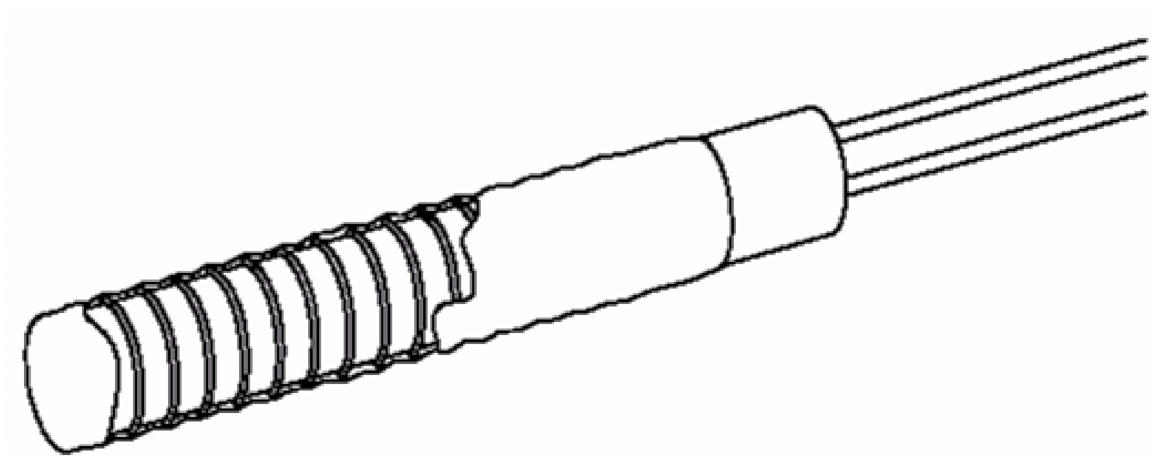
There are many categories; carbon resistors, film, and wire-wound types are the most widely used.

- *Carbon resistors* are widely available and are very inexpensive. They have very reproducible results at low temperatures. They are the most reliable form at extremely low temperatures. They generally do not suffer from significant hysteresis or strain gauge effects. Carbon resistors have been used for many years because of their advantages.

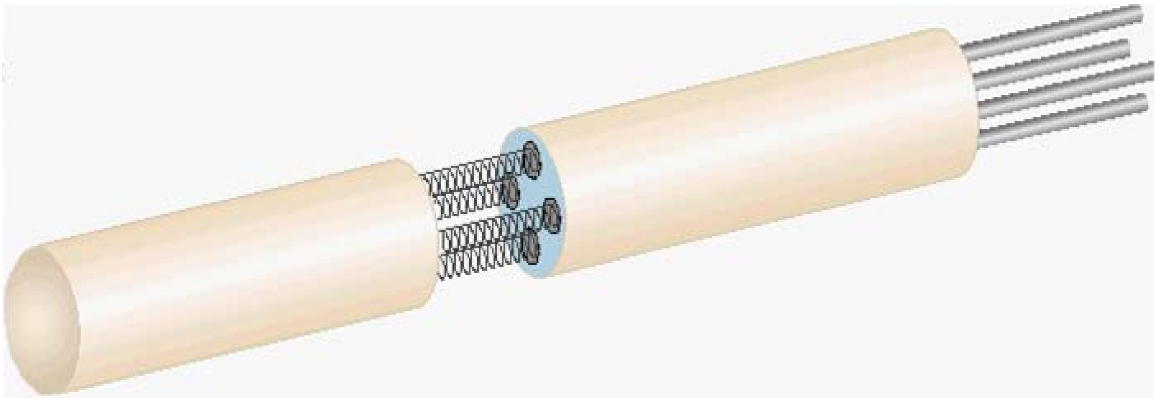
- *Film thermometers* have a layer of platinum on a substrate; the layer may be extremely thin, perhaps one micrometer. Advantages of this type are relatively low cost (the high cost of platinum being offset by the tiny amount required) and fast response. Such devices have improved performance although the different expansion rates of the substrate and platinum give "strain gauge" effects and stability problems.



Wire-wound thermometers can have greater accuracy, especially for wide temperature ranges. The coil diameter provides a compromise between mechanical stability and allowing expansion of the wire to minimize strain and consequential drift.



Coil elements have largely replaced wire-wound elements in industry. This design has a wire coil which can expand freely over temperature, held in place by some mechanical support which lets the coil keep its shape. This design is similar to that of a SPRT, the primary standard upon which ITS-90 is based, while providing the durability necessary for industrial use.



The current international standard which specifies tolerance, and the temperature-to-electrical resistance relationship for platinum resistance thermometers is IEC 751:1983. By far the most common devices used in industry have a nominal resistance of 100 ohms at 0 °C, and are called Pt100 sensors ('Pt' is the symbol for platinum). The sensitivity of a standard 100 ohm sensor is a nominal 0.385 ohm/°C. RTDs with a sensitivity of 0.375 and 0.392 ohm/°C as well as a variety of others are also available.

Function

Resistance thermometers are constructed in a number of forms and offer greater stability, accuracy and repeatability in some cases than thermocouples. While thermocouples use the Seebeck effect to generate a voltage, resistance thermometers use electrical resistance and require a power source to operate. The resistance ideally varies linearly with temperature.

Resistance thermometers are usually made using platinum, because of its linear resistance-temperature relationship and its chemical inertness. The platinum detecting wire needs to be kept free of contamination to remain stable. A platinum wire or film is supported on a former in such a way that it gets minimal differential expansion or other strains from its former, yet is reasonably resistant to vibration. RTD assemblies made from iron or copper are also used in some applications.

Commercial platinum grades are produced which exhibit a change of resistance of 0.00385 ohms/°C (European Fundamental Interval) The sensor is usually made to have a resistance of 100 Ω at 0 °C. This is defined in BS EN 60751:1996 (taken from IEC 60751:1995). The American Fundamental Interval is 0.00392 Ω/°C, based on using a purer grade of platinum than the European standard. The American standard is from the Scientific Apparatus Manufacturers Association (SAMA), who are no longer in this standards field. As a result the "American standard" is hardly the standard even in the US.

Measurement of resistance requires a small current to be passed through the device under test. This can cause resistive heating, causing significant loss of accuracy if manufacturers' limits are not respected, or the design does not properly consider the heat path. Mechanical strain on the resistance thermometer can also cause inaccuracy. Lead wire resistance can also be a factor; adopting three- and four-wire, instead of two-wire, connections can eliminate connection lead resistance effects from measurements; three-wire connection is sufficient for most purposes and almost universal industrial practice. Four-wire connections are used for the most precise applications.

Advantages and limitations

Advantages of platinum resistance thermometers:

- High accuracy
- Low drift
- Wide operating range
- Suitable for precision applications

Limitations:

- RTDs in industrial applications are rarely used above 660 °C. At temperatures above 660 °C it becomes increasingly difficult to prevent the platinum from becoming contaminated by impurities from the metal sheath of the thermometer. This is why laboratory standard thermometers replace the metal sheath with a glass construction. At very low temperatures, say below -270 °C (or 3 K), due to the fact that there are very few phonons, the resistance of an RTD is mainly determined by impurities and boundary scattering and thus basically independent of temperature. As a result, the sensitivity of the RTD is essentially zero and therefore not useful.
- Compared to thermistors, platinum RTDs are less sensitive to small temperature changes and have a slower response time. However, thermistors have a smaller temperature range and stability.

Sources of error:

The common error sources of a PRT are:

- *Interchangeability:* the “closeness of agreement” between the specific PRT's Resistance vs. Temperature relationship and a predefined Resistance vs. Temperature relationship, commonly defined by IEC 60751.
- *Insulation Resistance:* Error caused by the inability to measure the actual resistance of element. Current leaks into or out of the circuit through the sheath, between the element leads, or the elements.
- *Stability:* Ability to maintain R vs T over time as a result of thermal exposure.
- *Repeatability:* Ability to maintain R vs T under the same conditions after experiencing thermal cycling throughout a specified temperature range.
- *Hysteresis:* Change in the characteristics of the materials from which the RTD is built due to exposures to varying temperatures.
- *Stem Conduction:* Error that results from the PRT sheath conducting heat into or out of the process.
- *Calibration/Interpolation:* Errors that occur due to calibration uncertainty at the cal points, or between cal point due to propagation of uncertainty or curve fit errors.

- *Lead Wire:* Errors that occur because a 4 wire or 3 wire measurement is not used, this is greatly increased by higher gauge wire.
  - 2 wire connection adds lead resistance in series with PRT element.
  - 3 wire connection relies on all 3 leads having equal resistance.
- *Self Heating:* Error produced by the heating of the PRT element due to the power applied.
- *Time Response:* Errors are produced during temperature transients because the PRT cannot respond to changes fast enough.
- *Thermal EMF:* Thermal EMF errors are produced by the EMF adding to or subtracting from the applied sensing voltage, primarily in DC systems.

RTDs vs Thermocouples

The two most common ways of measuring industrial temperatures are with resistance temperature detectors (RTDs) and thermocouples. Choice between them is usually determined by four factors.

- *What are the temperature requirements?* If process temperatures are between -200 to 500 °C (-328 to 932 °F), an industrial RTD is the preferred option. Thermocouples have a range of -180 to 2,320 °C (-292 to 4,208 °F), so for temperatures above 500 °C (932 °F) they are the only contact temperature measurement device.

- *What are the time-response requirements?* If the process requires a very fast response to temperature changes—fractions of a second as opposed to seconds (e.g. 2.5 to 10 s)—then a thermocouple is the best choice. Time response is measured by immersing the sensor in water moving at 1 m/s (3 ft/s) with a 63.2% step change.

- *What are the size requirements?* A standard RTD sheath is 3.175 to 6.35 mm (0.1250 to 0.250 in) in diameter; sheath diameters for thermocouples can be less than 1.6 mm (0.063 in).

- *What are the accuracy and stability requirements?* If a tolerance of 2 °C is acceptable and the highest level of repeatability is not required, a thermocouple will serve. RTDs are capable of higher accuracy and can maintain stability for many years, while thermocouples can drift within the first few hours of use.

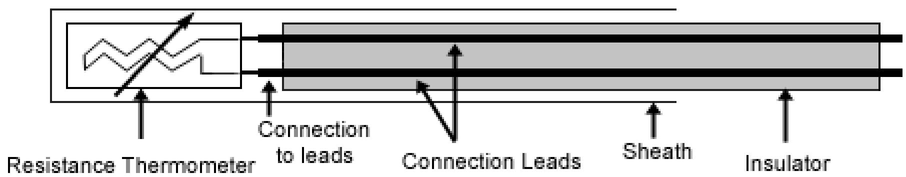
Elements

Resistance thermometer elements are available in a number of forms. The most common are:

- Unsupported wirewound - a wire coil minimally supported within a sealed housing filled with an inert gas. These sensors are used up to 961.78 °C and are used in the SPRT's that define ITS-90
- Wirewound in a ceramic insulator – a wire coil sealed in a ceramic cylinder, works with temperatures to 850 °C

- Wire encapsulated in glass - wire around glass core with glass fused homogeneously around. More protection to the detecting wire than other forms and resists vibration, but smaller usable range
- Thin film - platinum film on ceramic substrate, small and inexpensive to mass-produce, fast response to temperature change, but smaller temperature range and not capable of the highest accuracy

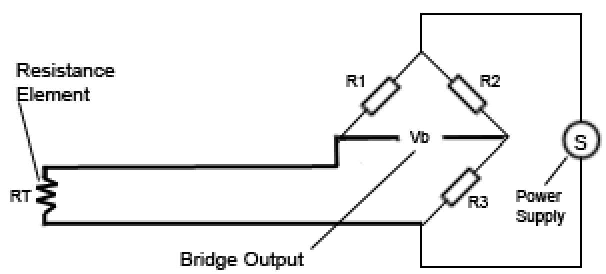
Construction



These elements nearly always require insulated leads attached. At temperatures below about 250 °C PVC, silicon rubber or PTFE insulators are used. Above this, glass fibre or ceramic are used. The measuring point, and usually most of the leads, require a housing or protective sleeve, often made of a metal alloy which is chemically inert to the process being monitored. Selecting and designing protection sheaths can require more care than the actual sensor, as the sheath must withstand chemical or physical attack and provide convenient attachment points.

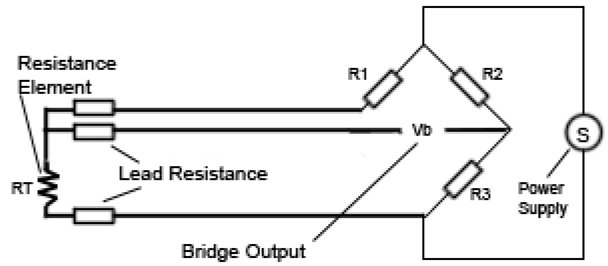
Wiring configurations

Two-wire configuration



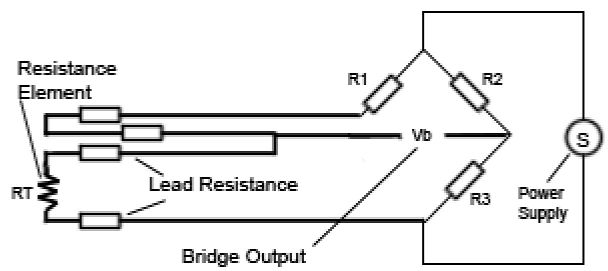
The simplest resistance thermometer configuration uses two wires. It is only used when high accuracy is not required, as the resistance of the connecting wires is added to that of the sensor, leading to errors of measurement. This configuration allows use of 100 meters of cable. This applies equally to balanced bridge and fixed bridge system.

Three-wire configuration

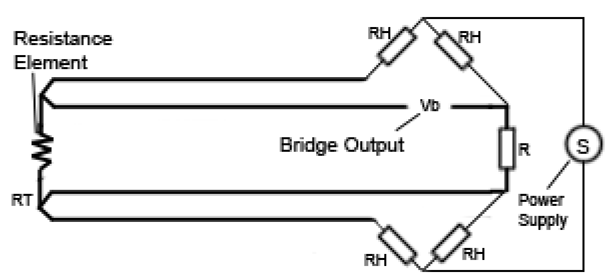


In order to minimize the effects of the lead resistances, a three-wire configuration can be used. Using this method the two leads to the sensor are on adjoining arms. There is a lead resistance in each arm of the bridge so that the resistance is cancelled out, so long as the two lead resistances are accurately the same. This configuration allows up to 600 meters of cable.

Four-wire configuration



The four-wire resistance thermometer configuration increases the accuracy and reliability of the resistance being measured: the resistance error due to lead wire resistance is zero. In the diagram above a standard two-terminal RTD is used with another pair of wires to form an additional loop that cancels out the lead resistance. The above Wheatstone bridge method uses a little more copper wire and is not a perfect solution. Below is a better configuration, four-wire Kelvin connection. It provides full cancellation of spurious effects; cable resistance of up to 15 Ω can be handled.



History

The application of the tendency of electrical conductors to increase their electrical resistance with rising temperature was first described by Sir William Siemens at the Bakerian Lecture of 1871 before the Royal Society of Great Britain. The necessary methods of construction were established by Callendar, Griffiths, Holborn and Wein between 1885 and 1900.

Standard resistance thermometer data

Temperature sensors are usually supplied with thin-film elements. The resisting elements are rated in accordance with BS EN 60751:2008 as:

Tolerance Class	Valid Range
Tolerance class F 0.3	-50 to +500 °C
Tolerance class F 0.15	-30 to +300 °C
Tolerance class F 0.1	0 to +150 °C

Resistance thermometer elements can be supplied which function up to 1000 °C. The relation between temperature and resistance is given by the Callendar-Van Dusen equation,

$$R_T = R_0\left[1 + AT + BT^2 + CT^3(T-100)\right]\ (-200\ °\mathrm{C} < T < 0\ °\mathrm{C}),$$
$$R_T = R_0\left[1 + AT + BT^2\right]\ (0\ °\mathrm{C} \leq T < 850\ °\mathrm{C}).$$

Here, R_T is the resistance at temperature T, R_0 is the resistance at 0 °C, and the constants (for an alpha=0.00385 platinum RTD) are

$$A = 3.9083 \times 10^{-3}\ °\mathrm{C}^{-1}$$
$$B = -5.775 \times 10^{-7}\ °\mathrm{C}^{-2}$$
$$C = -4.183 \times 10^{-12}\ °\mathrm{C}^{-4}.$$

Since the B and C coefficients are relatively small, the resistance changes almost linearly with the temperature.

Values for various popular resistance thermometers

Values for various popular resistance thermometers

Temperature in °C	Pt100 in Ω	Pt1000 in Ω	PTC in Ω	NTC in Ω	NTC in Ω	NTC in Ω	NTC in Ω	NTC in Ω

	Typ: 404	Typ: 501	Typ: 201	Typ: 101	Typ: 102	Typ: 103	Typ: 104	Typ: 105
−50	80.31	803.1	1032					
−45	82.29	822.9	1084					
−40	84.27	842.7	1135			50475		
−35	86.25	862.5	1191			36405		
−30	88.22	882.2	1246			26550		
−25	90.19	901.9	1306		26083	19560		
−20	92.16	921.6	1366		19414	14560		
−15	94.12	941.2	1430		14596	10943		
−10	96.09	960.9	1493		11066	8299		
−5	98.04	980.4	1561	31389	8466			
0	100.00	1000.0	1628	23868	6536			
5	101.95	1019.5	1700	18299	5078			
10	103.90	1039.0	1771	14130	3986			

15	105.85	1058.5	1847	10998		
20	107.79	1077.9	1922	8618		
25	109.73	1097.3	2000	6800	15000	
30	111.67	1116.7	2080	5401	11933	
35	113.61	1136.1	2162	4317	9522	
40	115.54	1155.4	2244	3471	7657	
45	117.47	1174.7	2330		6194	
50	119.40	1194.0	2415		5039	
55	121.32	1213.2	2505		4299	27475
60	123.24	1232.4	2595		3756	22590
65	125.16	1251.6	2689			18668
70	127.07	1270.7	2782			15052
75	128.98	1289.8	2880			12932
80	130.89	1308.9	2977			10837
85	132.80	1328.0	3079			9121

90	134.70	1347.0	3180	7708
95	136.60	1366.0	3285	6539
100	138.50	1385.0	3390	
105	140.39	1403.9		
110	142.29	1422.9		
150	157.31	1573.1		
200	175.84	1758.4		