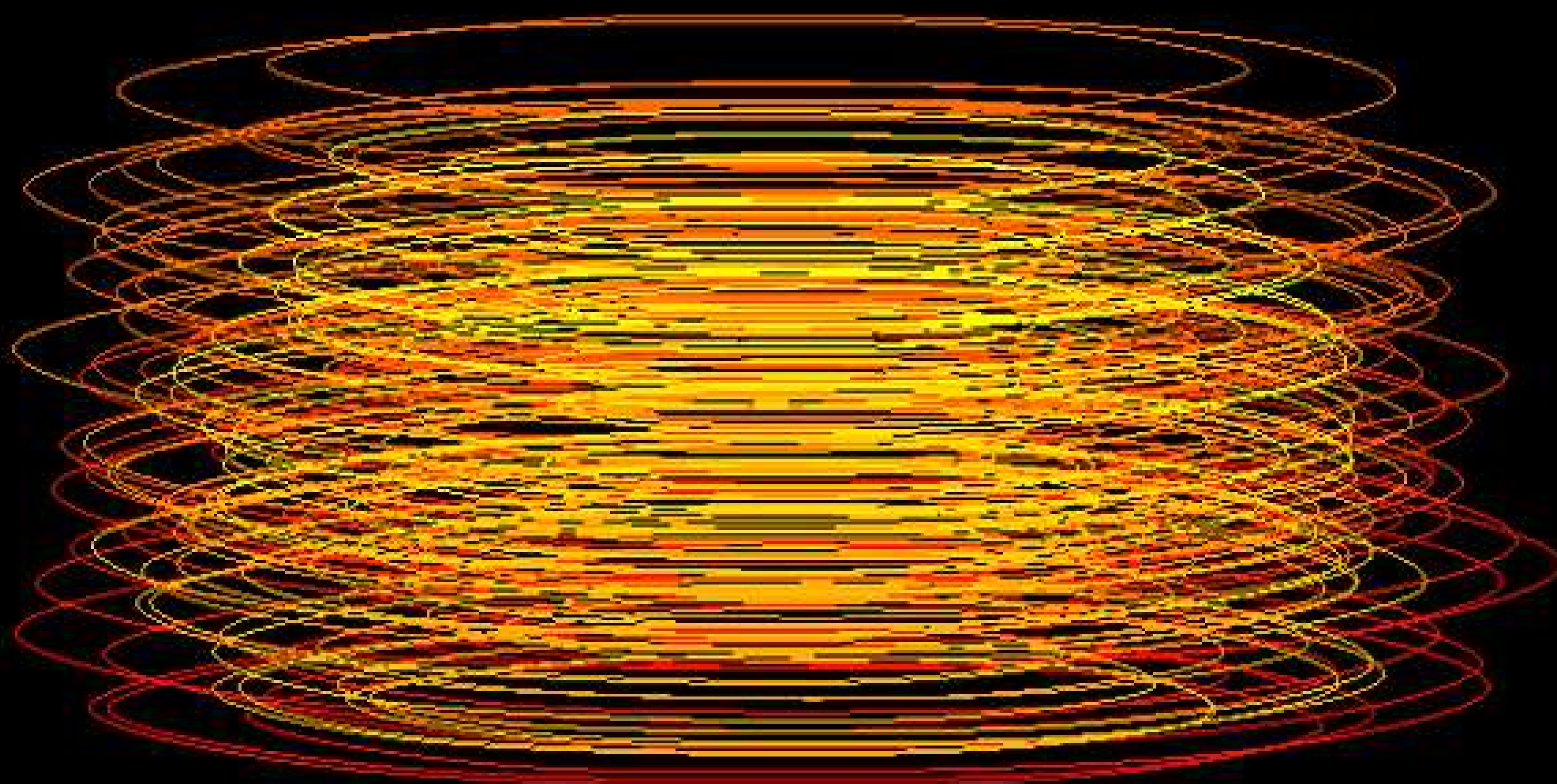
Electrical Impedance, Network and Circuits

(Concepts and Elements)



Chadwick Martel

First Edition, 2012

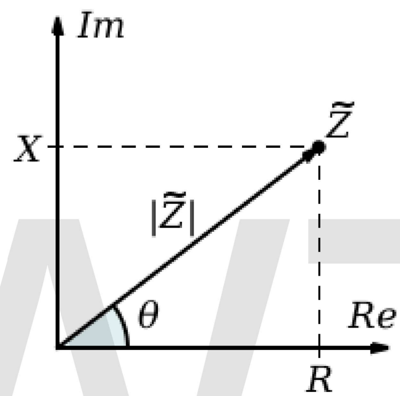
ISBN 978-81-323-1513-1

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

Table of Contents

Chapter- 1

Electrical Impedance



A graphical representation of the complex impedance plane

Electrical impedance, or simply **impedance**, describes a measure of opposition to alternating current (AC). Electrical impedance extends the concept of resistance to AC circuits, describing not only the relative amplitudes of the voltage and current, but also the relative phases. When the circuit is driven with direct current (DC) there is no distinction between impedance and resistance; the latter can be thought of as impedance with zero phase angle.

The symbol for impedance is usually Z and it may be represented by writing its magnitude and phase in the form $|Z|\angle\theta$. However, complex number representation is more powerful for circuit analysis purposes. The term *impedance* was coined by Oliver Heaviside in July 1886. Arthur Kennelly was the first to represent impedance with complex numbers in 1893.

Impedance is defined as the frequency domain ratio of the voltage to the current. In other words, it is the voltage–current ratio for a single complex exponential at a particular frequency ω. In general, impedance will be a complex number, with the same units as resistance, for which the SI unit is the ohm. For a sinusoidal current or voltage input, the polar form of the complex impedance relates the amplitude and phase of the voltage and current. In particular,

- The magnitude of the complex impedance is the ratio of the voltage amplitude to the current amplitude.
- The phase of the complex impedance is the phase shift by which the current is ahead of the voltage.

The reciprocal of impedance is admittance (i.e., admittance is the current-to-voltage ratio, and it conventionally carries units of siemens, formerly called mhos).

Complex impedance

Impedance is represented as a complex quantity Z and the term *complex impedance* may be used interchangeably; the polar form conveniently captures both magnitude and phase characteristics,

$$Z = |Z|e^{j\theta}$$

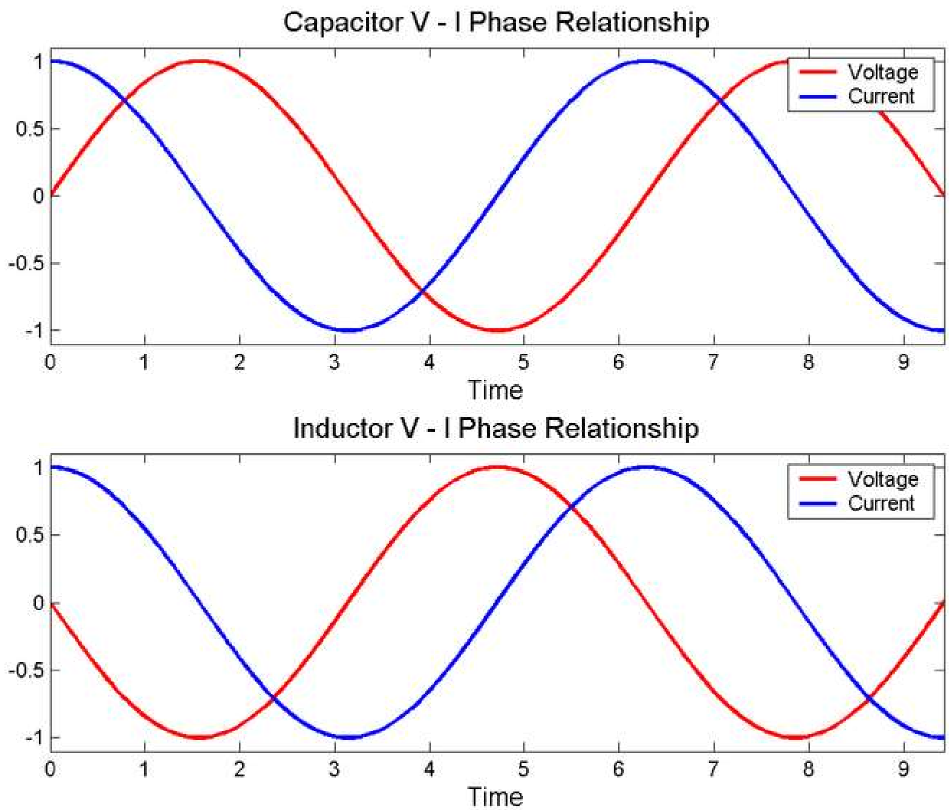
where the magnitude $|Z|$ represents the ratio of the voltage difference amplitude to the current amplitude, while the argument θ gives the phase difference between voltage and current and j is the imaginary unit. In Cartesian form,

$$Z = R + jX$$

where the real part of impedance is the resistance R and the imaginary part is the reactance X.

Where it is required to add or subtract impedances the cartesian form is more convenient, but when quantities are multiplied or divided the calculation becomes simpler if the polar form is used. A circuit calculation, such as finding the total impedance of two impedances in parallel, may require conversion between forms several times during the calculation. Conversion between the forms follows the normal conversion rules of complex numbers.

Device examples



The phase angles in the equations for the impedance of inductors and capacitors indicate that the voltage across a capacitor *lags* the current through it by a phase of $\pi / 2$, while the voltage across an inductor *leads* the current through it by $\pi / 2$. The identical voltage and current amplitudes tell us that the magnitude of the impedance is equal to one.

The impedance of an ideal resistor is purely real and is referred to as a *resistive impedance*:

$$Z_R = R.$$

Ideal inductors and capacitors have a purely imaginary *reactive impedance*:

$$Z_L = j\omega L,$$
$$Z_C = \frac{1}{j\omega C}\ .$$

Note the following identities for the imaginary unit and its reciprocal:

$$j = \cos\left(\frac{\pi}{2}\right) + j\sin\left(\frac{\pi}{2}\right) = e^{j\frac{\pi}{2}},$$
$$\frac{1}{j} = -j = \cos\left(-\frac{\pi}{2}\right) + j\sin\left(-\frac{\pi}{2}\right) = e^{j(-\frac{\pi}{2})}.$$

Thus we can rewrite the inductor and capacitor impedance equations in polar form:

$$Z_L = \omega L e^{j\frac{\pi}{2}},$$
$$Z_C = \frac{1}{\omega C} e^{j(-\frac{\pi}{2})}.$$

The magnitude tells us the change in voltage amplitude for a given current amplitude through the impedance, while the exponential factors give the phase relationship.

Deriving the device specific impedances

What follows below is a derivation of impedance for each of the three basic circuit elements, the resistor, the capacitor, and the inductor. Although the idea can be extended to define the relationship between the voltage and current of any arbitrary signal, these derivations will assume sinusoidal signals, since any arbitrary signal can be approximated as a sum of sinusoids through Fourier Analysis.

Resistor

For a resistor, we have the relation:

$$v_{\text{R}}(t) = i_{\text{R}}(t) R.$$

This is simply a statement of Ohm's Law.

Considering the voltage signal to be

$$v_{\text{R}}(t) = V_p \sin(\omega t)\,,$$

it follows that

$$\frac{v_{\text{R}}(t)}{i_{\text{R}}(t)} = \frac{V_p \sin(\omega t)}{I_p \sin(\omega t)} = R.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across a resistor is R, and that the AC voltage leads the AC current across a resistor by 0 degrees.

This result is commonly expressed as

$$Z_{\text{resistor}} = R\,.$$

Capacitor

For a capacitor, we have the relation:

$$i_{\text{C}}(t) = C\frac{\mathrm{d}v_{\text{C}}(t)}{\mathrm{d}t}.$$

Considering the voltage signal to be

$$v_{\text{C}}(t) = V_p \sin(\omega t)$$

it follows that

$$\frac{\mathrm{d}v_{\text{C}}(t)}{\mathrm{d}t} = \omega V_p \cos(\omega t)\,.$$

And thus

$$\frac{v_{\text{C}}(t)}{i_{\text{C}}(t)} = \frac{V_p \sin(\omega t)}{\omega V_p C \cos(\omega t)} = \frac{\sin(\omega t)}{\omega C \sin\left(\omega t + \frac{\pi}{2}\right)}.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across a capacitor is $\frac{1}{\omega C}$, and that the AC voltage leads the AC current across a capacitor by -90 degrees (or the AC current leads the AC voltage across a capacitor by 90 degrees).

This result is commonly expressed in polar form, as

$$Z_{\text{capacitor}} = \frac{1}{\omega C} e^{-j\frac{\pi}{2}}$$

or, by applying Euler's formula, as

$$Z_{\text{capacitor}} = -j\frac{1}{\omega C} = \frac{1}{j\omega C}\,.$$

Inductor

For the inductor, we have the relation:

$$v_{\text{L}}(t) = L\frac{\mathrm{d}i_{\text{L}}(t)}{\mathrm{d}t}.$$

This time, considering the current signal to be

$$i_{\mathrm{L}}(t) = I_p \sin(\omega t)\,,$$

it follows that

$$\frac{\mathrm{d}i_{\mathrm{L}}(t)}{\mathrm{d}t} = \omega I_p \cos(\omega t)\,.$$

And thus

$$\frac{v_{\mathrm{L}}(t)}{i_{\mathrm{L}}(t)} = \frac{\omega I_p L \cos(\omega t)}{I_p \sin(\omega t)} = \frac{\omega L \sin\left(\omega t + \frac{\pi}{2}\right)}{\sin(\omega t)}.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across an inductor is ωL, and that the AC voltage leads the AC current across an inductor by 90 degrees.

This result is commonly expressed in polar form, as

$$Z_{\text{inductor}} = \omega L e^{j\frac{\pi}{2}}.$$

Or, more simply, using Euler's formula, as

$$Z_{\text{inductor}} = j\omega L.$$

Generalised s-plane impedance

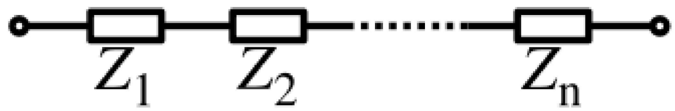
Impedance defined in terms of $j\omega$ can strictly only be applied to circuits which are energised with a steady-state AC signal. The concept of impedance can be extended to a circuit energised with any arbitrary signal by using complex frequency instead of $j\omega$. Complex frequency is given the symbol s and is, in general, a complex number. Signals are expressed in terms of complex frequency by taking the Laplace transform of the time domain expression of the signal. The impedance of the basic circuit elements in this more general notation is as follows:

Element	Impedance expression
Resistor	R
Inductor	sL
Capacitor	$\frac{1}{sC}$

For a DC circuit this simplifies to $s = 0$. For a steady-state sinusoidal AC signal $s = j\omega$.

Series combination

For components connected in series, the current through each circuit element is the same; the total impedance is simply the sum of the component impedances.



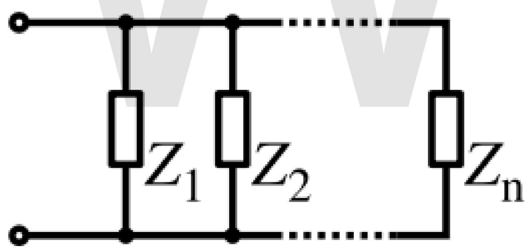
$$Z_{eq} = Z_1 + Z_2 + \cdots + Z_n$$

Or explicitly in real and imaginary terms:

$$Z_{eq} = R + jX = (R_1 + R_2 + \cdots + R_n) + j(X_1 + X_2 + \cdots + X_n)$$

Parallel combination

For components connected in parallel, the voltage across each circuit element is the same; the ratio of currents through any two elements is the inverse ratio of their impedances.



Hence the inverse total impedance is the sum of the inverses of the component impedances:

$$\frac{1}{Z_{eq}} = \frac{1}{Z_1} + \frac{1}{Z_2} + \cdots + \frac{1}{Z_n}$$

or, when n = 2:

$$Z_{eq} = Z_1 \| Z_2 \stackrel{\text{def}}{=} \frac{Z_1 Z_2}{Z_1 + Z_2}$$

The equivalent impedance Z_{eq} can be calculated in terms of the equivalent resistance R_{eq} and reactance X_{eq}.

$$Z_{\text{eq}} = R_{\text{eq}} + jX_{\text{eq}}$$
$$R_{\text{eq}} = \frac{(X_1R_2 + X_2R_1)(X_1 + X_2) + (R_1R_2 - X_1X_2)(R_1 + R_2)}{(R_1 + R_2)^2 + (X_1 + X_2)^2}$$
$$X_{\text{eq}} = \frac{(X_1R_2 + X_2R_1)(R_1 + R_2) - (R_1R_2 - X_1X_2)(X_1 + X_2)}{(R_1 + R_2)^2 + (X_1 + X_2)^2}$$

Measurement

According to Ohm's law the impedance of a device can be calculated by complex division of the voltage and current. The impedance of the device can be calculated by applying a sinusoidal voltage to the device in series with a resistor, and measuring the voltage across the resistor and across the device. Performing this measurement by sweeping the frequencies of the applied signal provides the impedance phase and magnitude.

Impulse impedance spectroscopy

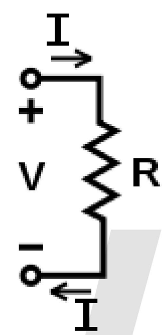
The use of an impulse response may be used in combination with the fast Fourier transform (FFT) to rapidly measure the electrical impedance of various electrical devices. The technique compares well to other methodologies such as network and impedance analyzers while providing additional versatility in the electrical impedance measurement. The technique is theoretically simple, easy to implement and completed with ordinary laboratory instrumentation for minimal cost.

Variable impedance

In general, neither impedance nor admittance can be time varying as they are defined for complex exponentials for $-\infty < t < +\infty$. If the complex exponential voltage–current ratio changes over time or amplitude, the circuit element cannot be described using the frequency domain. However, many systems (e.g., varicaps that are used in radio tuners) may exhibit non-linear or time-varying voltage–current ratios that appear to be LTI for small signals over small observation windows; hence, they can be roughly described as having a time-varying impedance. That is, this description is an approximation; over large signal swings or observation windows, the voltage–current relationship is non-LTI and cannot be described by impedance.

Chapter- 2

Ohm's Law



V, I, and R, the parameters of Ohm's law

Ohm's law states that the current through a conductor between two points is directly proportional to the potential difference or voltage across the two points, and inversely proportional to the resistance between them.

The mathematical equation that describes this relationship is:

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

where I is the current through the resistance in units of amperes, V is the potential difference measured *across* the resistance in units of volts, and R is the resistance of the conductor in units of ohms. More specifically, Ohm's law states that the R in this relation is constant, independent of the current.

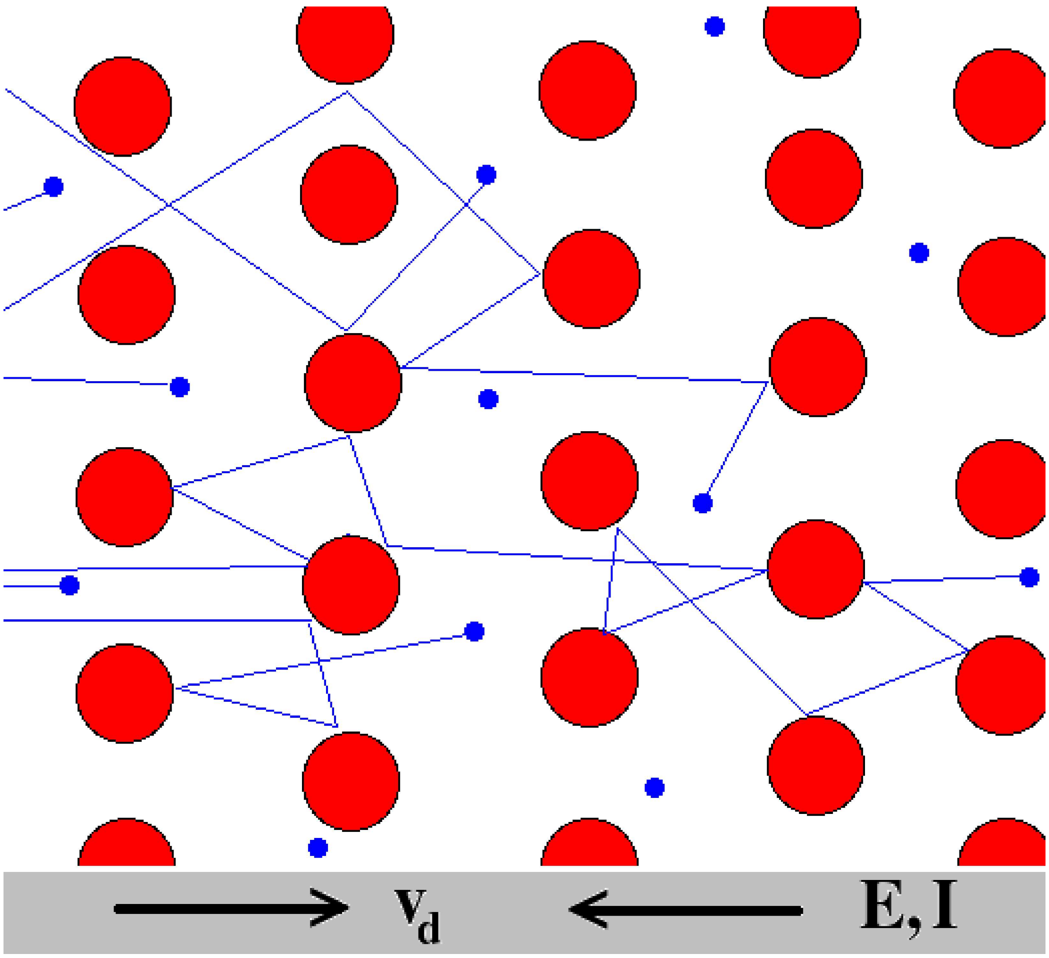
The law was named after the German physicist Georg Ohm, who, in a treatise published in 1827, described measurements of applied voltage and current through simple electrical circuits containing various lengths of wire. He presented a slightly more complex equation than the one above to explain his experimental results. The above equation is the modern form of Ohm's law.

In physics, the term *Ohm's law* is also used to refer to various generalizations of the law originally formulated by Ohm. The simplest example of this is:

$$J = \sigma E,$$

where J is the current density at a given location in a resistive material, E is the electric field at that location, and σ is a material dependent parameter called the conductivity. This reformulation of Ohm's law is due to Gustav Kirchhoff.

Microscopic origins of Ohm's law



Drude Model electrons (shown here in blue) constantly bounce between heavier, stationary crystal ions (shown in red).

The dependence of the current density on the applied electric field is essentially quantum mechanical in nature. A qualitative description leading to Ohm's law can be based upon classical mechanics using the Drude model developed by Paul Drude in 1900.

The Drude model treats electrons (or other charge carriers) like pinballs bouncing between the ions that make up the structure of the material. Electrons will be accelerated in the opposite direction to the electric field by the average electric field at their location. With each collision, though, the electron is deflected in a random direction with a velocity that is much larger than the velocity gained by the electric field. The net result is that electrons take a tortuous path due to the collisions, but generally drift in a direction opposing the electric field.

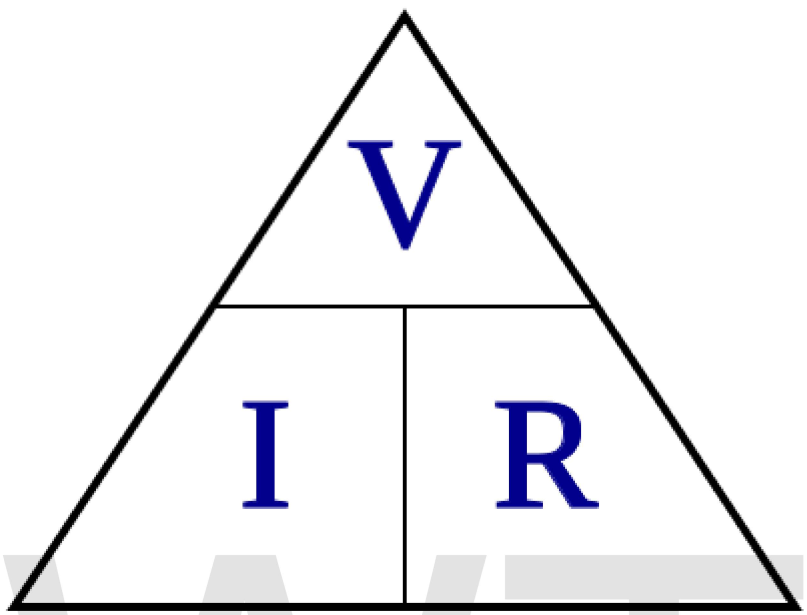
The drift velocity then determines the electric current density and its relationship to $\boldsymbol{E}$ and is independent of the collisions. Drude calculated the average drift velocity from $\boldsymbol{p} = -e\boldsymbol{E}\tau$ where $\boldsymbol{p}$ is the average momentum, $-e$ is the charge of the electron and τ is the average time between the collisions. Since both the momentum and the current density are proportional to the drift velocity, the current density becomes proportional to the applied electric field; this leads to Ohm's law.

Hydraulic analogy

A hydraulic analogy is sometimes used to describe Ohm's Law. Water pressure, measured by pascals (or PSI), is the analog of voltage because establishing a water pressure difference between two points along a (horizontal) pipe causes water to flow. Water flow rate, as in liters per second, is the analog of current, as in coulombs per second. Finally, flow restrictors — such as apertures placed in pipes between points where the water pressure is measured — are the analog of resistors. We say that the rate of water flow through an aperture restrictor is proportional to the difference in water pressure across the restrictor. Similarly, the rate of flow of electrical charge, that is, the electric current, through an electrical resistor is proportional to the difference in voltage measured across the resistor.

Flow and pressure variables can be calculated in fluid flow network with the use of the hydraulic ohm analogy. The method can be applied to both steady and transient flow situations. In the linear laminar flow region, Poiseuille's law describes the hydraulic resistance of a pipe, but in the turbulent flow region the pressure–flow relations become nonlinear.

Circuit analysis



Ohm's law triangle

In circuit analysis, three equivalent expressions of Ohm's law are used interchangeably:

$$I = \frac{V}{R} \quad \text{or} \quad V = IR \quad \text{or} \quad R = \frac{V}{I}$$

Indeed, each is quoted by some sources as the defining relationship of Ohm's law, or all three are quoted, or derived from a proportional form, or even just the two that do not correspond to Ohm's original statement may sometimes be given.

The interchangeability of the equation may be represented by a triangle, where V (voltage) is placed on the top section, the I (current) is placed to the left section, and the R (resistance) is placed to the right. The line that divides the left and right sections indicate multiplication, and the divider between the top and bottom sections indicates division (hence the division bar).

Resistive circuits

Resistors are circuit elements that impede the passage of electric charge in agreement with Ohm's law, and are designed to have a specific resistance value *R*. In a schematic diagram the resistor is shown as a zig-zag symbol. An element (resistor or conductor) that behaves according to Ohm's law over some operating range is referred to as an *ohmic device* (or an *ohmic resistor*) because Ohm's law and a single value for the resistance suffice to describe the behavior of the device over that range.

Ohm's law holds for circuits containing only resistive elements (no capacitances or inductances) for all forms of driving voltage or current, regardless of whether the driving voltage or current is constant (DC) or time-varying such as AC. At any instant of time Ohm's law is valid for such circuits.

Resistors which are in *series* or in *parallel* may be grouped together into a single "equivalent resistance" in order to apply Ohm's law in analyzing the circuit.

Reactive circuits with time-varying signals

When reactive elements such as capacitors, inductors, or transmission lines are involved in a circuit to which AC or time-varying voltage or current is applied, the relationship between voltage and current becomes the solution to a differential equation, so Ohm's law (as defined above) does not directly apply since that form contains only resistances having value R, not complex impedances which may contain capacitance ("C") or inductance ("L").

Equations for time-invariant AC circuits take the same form as Ohm's law, however, the variables are generalized to complex numbers and the current and voltage waveforms are complex exponentials.

In this approach, a voltage or current waveform takes the form Ae^{st}, where t is time, s is a complex parameter, and A is a complex scalar. In any linear time-invariant system, all of the currents and voltages can be expressed with the same s parameter as the input to the system, allowing the time-varying complex exponential term to be canceled out and the system described algebraically in terms of the complex scalars in the current and voltage waveforms.

The complex generalization of resistance is impedance, usually denoted Z; it can be shown that for an inductor,

$$Z = sL$$

and for a capacitor,

$$Z = \frac{1}{sC}.$$

We can now write,

$$\boldsymbol{V} = \boldsymbol{I} \cdot \boldsymbol{Z}$$

where $\boldsymbol{V}$ and $\boldsymbol{I}$ are the complex scalars in the voltage and current respectively and $\boldsymbol{Z}$ is the complex impedance.

This form of Ohm's law, with *Z* taking the place of *R*, generalizes the simpler form. When *Z* is complex, only the real part is responsible for dissipating heat.

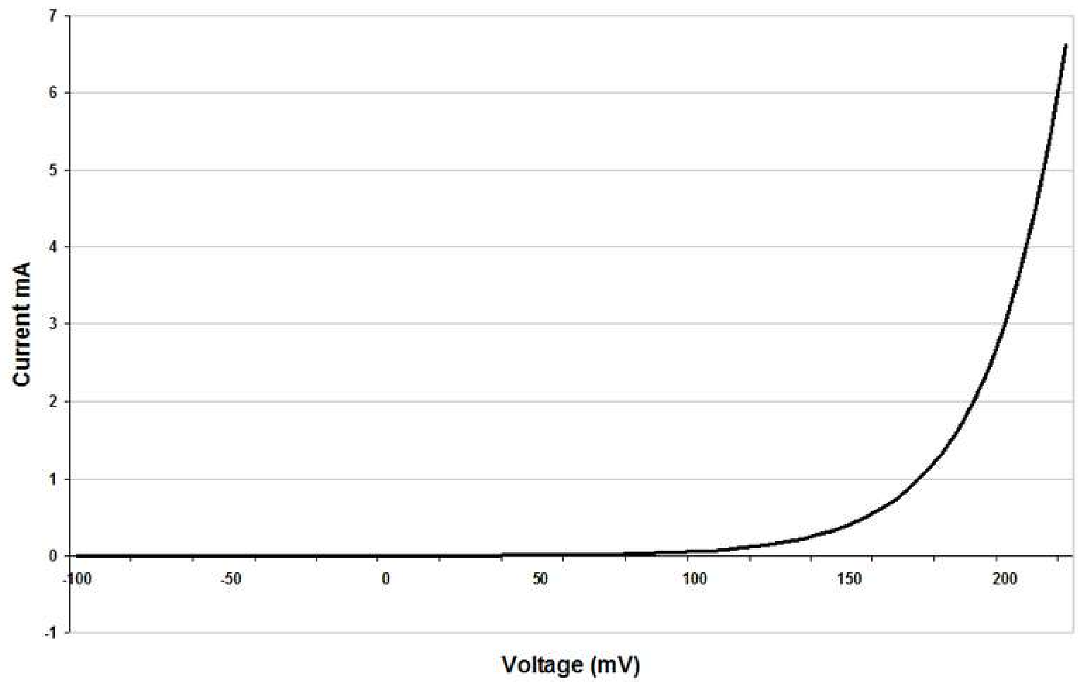
In the general AC circuit, *Z* varies strongly with the frequency parameter *s*, and so also will the relationship between voltage and current.

For the common case of a steady sinusoid, the *s* parameter is taken to be $j\omega$, corresponding to a complex sinusoid $Ae^{j\omega t}$. The real parts of such complex current and voltage waveforms describe the actual sinusoidal currents and voltages in a circuit, which can be in different phases due to the different complex scalars.

Linear approximations

Ohm's law is one of the basic equations used in the analysis of electrical circuits. It applies to both metal conductors and circuit components (resistors) specifically made for this behaviour. Both are ubiquitous in electrical engineering. Materials and components that obey Ohm's law are described as "ohmic" which means they produce the same value for resistance (R = V/I) regardless of the value of V or I which is applied and whether the applied voltage or current is DC (direct current) of either positive or negative polarity or AC (alternating current).

In a true ohmic device, the same value of resistance will be calculated from R = V/I regardless of the value of the applied voltage V. That is, the ratio of V/I is constant, and when current is plotted as a function of voltage the curve is *linear* (a straight line). If voltage is forced to some value V, then that voltage V divided by measured current I will equal R. Or if the current is forced to some value I, then the measured voltage V divided by that current I is also R. Since the plot of I versus V is a straight line, then it is also true that for any set of two different voltages V_1 and V_2 applied across a given device of resistance R, producing currents $I_1 = V_1/R$ and $I_2 = V_2/R$, that the ratio $(V_1-V_2)/(I_1-I_2)$ is also a constant equal to R. The operator "delta" (Δ) is used to represent a difference in a quantity, so we can write $\Delta V = V_1-V_2$ and $\Delta I = I_1-I_2$. Summarizing, for any truly ohmic device having resistance R, $V/I = \Delta V/\Delta I = R$ for any applied voltage or current or for the difference between any set of applied voltages or currents.



Plot of I–V curve of an ideal p-n junction diode at 1µA reverse leakage current. Failure of the device to follow Ohm's law is clearly shown since the curve is not a straight line.

There are, however, components of electrical circuits which do not obey Ohm's law; that is, their relationship between current and voltage (their I–V curve) is *nonlinear*. An example is the p-n junction diode (curve at right). As seen in the figure, the current does not increase linearly with applied voltage for a diode. One can determine a value of current (I) for a given value of applied voltage (V) from the curve, but not from Ohm's law, since the value of "resistance" is not constant as a function of applied voltage. Further, the current only increases significantly if the applied voltage is positive, not negative. The ratio V/I for some point along the nonlinear curve is sometimes called the *static*, or *chordal*, or DC, resistance, but as seen in the figure the value of total V over total I varies depending on the particular point along the nonlinear curve which is chosen. This means the "DC resistance" V/I at some point on the curve is not the same as what would be determined by applying an AC signal having peak amplitude ΔV volts or ΔI amps centered at that same point along the curve and measuring ΔV/ΔI. However, in some diode applications, the AC signal applied to the device is small and it is possible to analyze the circuit in terms of the *dynamic*, *small-signal*, or *incremental* resistance, defined as the one over the slope of the V–I curve at the average value (DC operating point) of the voltage (that is, one over the derivative of current with respect to voltage). For sufficiently small signals, the dynamic resistance allows the Ohm's law small signal resistance to be calculated as approximately one over the slope of a line drawn tangentially to the V-I curve at the DC operating point.

Temperature effects

Ohm's law has sometimes been stated as, "for a conductor in a given state, the electromotive force is proportional to the current produced." That is, that the resistance, the ratio of the applied electromotive force (or voltage) to the current, "does not vary with the current strength." The qualifier "in a given state" is usually interpreted as meaning "at a constant temperature," since the resistivity of materials is usually temperature dependent. Because the conduction of current is related to Joule heating of the conducting body, according to Joule's first law, the temperature of a conducting body may change when it carries a current. The dependence of resistance on temperature therefore makes resistance depend upon the current in a typical experimental setup, making the law in this form difficult to directly verify. Maxwell and others worked out several methods to test the law experimentally in 1876, controlling for heating effects.

Relation to heat conductions

Ohm's principle predicts the flow of electrical charge (i.e. current) in electrical conductors when subjected to the influence of voltage differences; Jean-Baptiste-Joseph Fourier's principle predicts the flow of heat in heat conductors when subjected to the influence of temperature differences.

The same equation describes both phenomena, the equation's variables taking on different meanings in the two cases. Specifically, solving a heat conduction (Fourier) problem with *temperature* (the driving "force") and *flux of heat* (the rate of flow of the driven "quantity", i.e. heat energy) variables also solves an analogous electrical conduction (Ohm) problem having *electric potential* (the driving "force") and *electric current* (the rate of flow of the driven "quantity", i.e. charge) variables.

The basis of Fourier's work was his clear conception and definition of thermal conductivity. He assumed that, all else being the same, the flux of heat is strictly proportional to the gradient of temperature. Although undoubtedly true for small temperature gradients, strictly proportional behavior will be lost when real materials (e.g. ones having a thermal conductivity that is a function of temperature) are subjected to large temperature gradients.

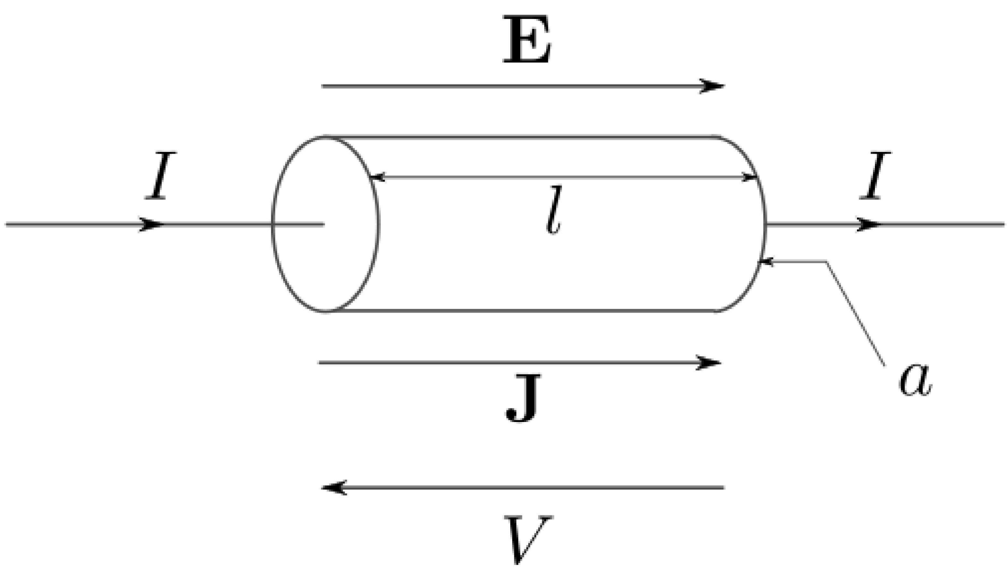
A similar assumption is made in the statement of Ohm's law: other things being alike, the strength of the current at each point is proportional to the gradient of electric potential. The accuracy of the assumption that flow is proportional to the gradient is more readily tested, using modern measurement methods, for the electrical case than for the heat case.

Other versions of Ohm's law

Ohm's law, in the form above, is an extremely useful equation in the field of electrical/electronic engineering because it describes how voltage, current and resistance are interrelated on a "macroscopic" level, that is, commonly, as circuit elements in an electrical circuit. Physicists who study the electrical properties of matter at the microscopic level use a closely related and more general vector equation, sometimes also referred to as Ohm's law, having variables that are closely related to the V, I, and R scalar variables of Ohm's law, but are each functions of position within the conductor. Physicists often use this continuum form of Ohm's Law:

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

where "**E**" is the electric field vector with units of volts per meter (analogous to "V" of Ohm's law which has units of volts), "**J**" is the current density vector with units of amperes per unit area (analogous to "I" of Ohm's law which has units of amperes), and "ρ" (Greek "rho") is the resistivity with units of ohm·meters (analogous to "R" of Ohm's law which has units of ohms). The above equation is sometimes written as **J** = σ**E** where "σ" is the conductivity which is the reciprocal of ρ.



Current flowing through a uniform cylindrical conductor (such as a round wire) with a uniform field applied.

The potential difference between two points is defined as:

$$\Delta V = -\int \mathbf{E} \cdot d\mathbf{l}$$

with $d\mathbf{l}$ the element of path along the integration of electric field vector **E**. If the applied **E** field is uniform and oriented along the length of the conductor as shown in the figure, then defining the voltage V in the usual convention of being opposite in direction to the field (see figure), and with the understanding that the voltage V is measured differentially across the length of the conductor allowing us to drop the Δ symbol, the above vector equation reduces to the scalar equation:

$$V = El \text{ or } E = \frac{V}{l}$$

Since the **E** field is uniform in the direction of wire length, for a conductor having uniformly consistent resistivity ρ, the current density **J** will also be uniform in any cross-sectional area and oriented in the direction of wire length, so we may write:

$$J = \frac{I}{a}$$

Substituting the above 2 results (for *E* and *J* respectively) into the continuum form shown at the beginning of this section:

$$\frac{V}{l} = \frac{I}{a}\rho \qquad \text{or} \qquad V = I\rho\frac{l}{a}$$

The electrical resistance of a uniform conductor is given in terms of resistivity by:

$$R = \rho\frac{l}{a}$$

where l is the length of the conductor in SI units of meters, a is the cross-sectional area (for a round wire $a = \pi r^2$ if r is radius) in units of meters squared, and ρ is the resistivity in units of ohm·meters.

After substitution of *R* from the above equation into the equation preceding it, the continuum form of Ohm's law for a uniform field (and uniform current density) oriented along the length of the conductor reduces to the more familiar form:

$$V = IR$$

A perfect crystal lattice, with low enough thermal motion and no deviations from periodic structure, would have no resistivity, but a real metal has crystallographic defects, impurities, multiple isotopes, and thermal motion of the atoms. Electrons scatter from all of these, resulting in resistance to their flow.

The more complex generalized forms of Ohm's law are important to condensed matter physics, which studies the properties of matter and, in particular, its electronic structure. In broad terms, they fall under the topic of constitutive equations and the theory of transport coefficients.

Magnetic effects

The continuum form of the equation is only valid in the reference frame of the conducting material. If the material is moving at velocity **v** relative to a magnetic field **B**, a term must be added as follows:

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{J}\rho$$

History

In January 1781, before Georg Ohm's work, Henry Cavendish experimented with Leyden jars and glass tubes of varying diameter and length filled with salt solution. He measured the current by noting how strong a shock he felt as he completed the circuit with his body. Cavendish wrote that the "velocity" (current) varied directly as the "degree of electrification" (voltage). He did not communicate his results to other scientists at the time, and his results were unknown until Maxwell published them in 1879.

Ohm did his work on resistance in the years 1825 and 1826, and published his results in 1827 as the book *Die galvanische Kette, mathematisch bearbeitet* (The galvanic Circuit investigated mathematically). He drew considerable inspiration from Fourier's work on heat conduction in the theoretical explanation of his work. For experiments, he initially used voltaic piles, but later used a thermocouple as this provided a more stable voltage source in terms of internal resistance and constant potential difference. He used a galvanometer to measure current, and knew that the voltage between the thermocouple terminals was proportional to the junction temperature. He then added test wires of varying length, diameter, and material to complete the circuit. He found that his data could be modeled through the equation

$$x = \frac{a}{b + l},$$

where x was the reading from the galvanometer, l was the length of the test conductor, a depended only on the thermocouple junction temperature, and b was a constant of the entire setup. From this, Ohm determined his law of proportionality and published his results.

Ohm's law was probably the most important of the early quantitative descriptions of the physics of electricity. We consider it almost obvious today. When Ohm first published his work, this was not the case; critics reacted to his treatment of the subject with hostility. They called his work a "web of naked fancies" and the German Minister of Education proclaimed that "a professor who preached such heresies was unworthy to teach science." The prevailing scientific philosophy in Germany at the time, led by Hegel, asserted that experiments need not be performed to develop an understanding of nature because nature is so well ordered, and that scientific truths may be deduced through reasoning alone. Also, Ohm's brother Martin, a mathematician, was battling the German educational system. These factors hindered the acceptance of Ohm's work, and his work did not become widely accepted until the 1840s. Fortunately, Ohm received recognition for his contributions to science well before he died.

In the 1850s, Ohm's law was known as such, and was widely considered proved, and alternatives such as "Barlow's law" discredited, in terms of real applications to telegraph system design, as discussed by Samuel F. B. Morse in 1855.

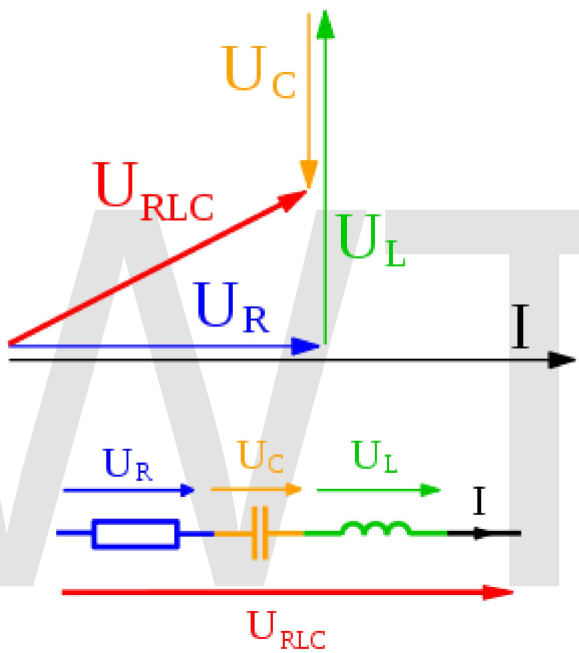
While the old term for electrical conductance, the mho (the inverse of the resistance unit ohm), is still used, a new name, the siemens, was adopted in 1971, honoring Ernst Werner von Siemens. The siemens is preferred in formal papers.

In the 1920s, it was discovered that the current through an ideal resistor actually has statistical fluctuations, which depend on temperature, even when voltage and resistance are exactly constant; this fluctuation, now known as Johnson–Nyquist noise, is due to the discrete nature of charge. This thermal effect implies that measurements of current and voltage that are taken over sufficiently short periods of time will yield ratios of V/I that fluctuate from the value of R implied by the time average or ensemble average of the measured current; Ohm's law remains correct for the average current, in the case of ordinary resistive materials.

Ohm's work long preceded Maxwell's equations and any understanding of frequency-dependent effects in AC circuits. Modern developments in electromagnetic theory and circuit theory do not contradict Ohm's law when they are evaluated within the appropriate limits.

Chapter- 3

Phasor



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

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

Definition

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

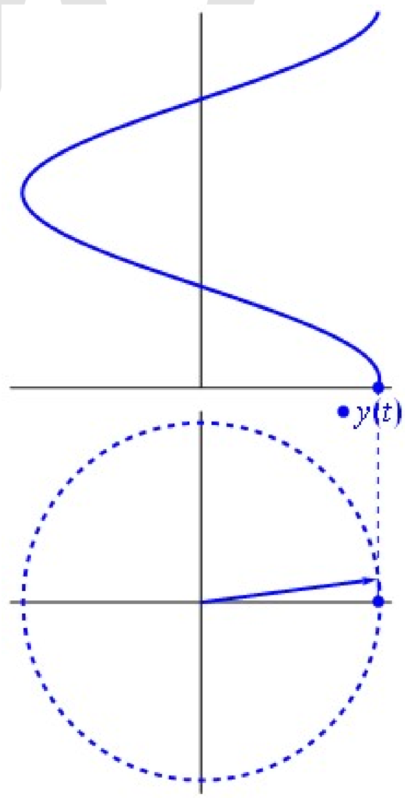
$$A \cdot \cos(\omega t + \theta) = A/2 \cdot e^{i(\omega t + \theta)} + A/2 \cdot e^{-i(\omega t + \theta)},$$

or as the real part of one of the functions:

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

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

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



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

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

Phasor arithmetic

Multiplication by a constant (scalar)

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

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

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

Differentiation and integration

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

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

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

For example, consider the following differential equation for the voltage across the capacitor in an RC circuit:

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

When the voltage source in this circuit is sinusoidal:

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

we may substitute:

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

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

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

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

Solving for the phasor capacitor voltage gives:

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

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

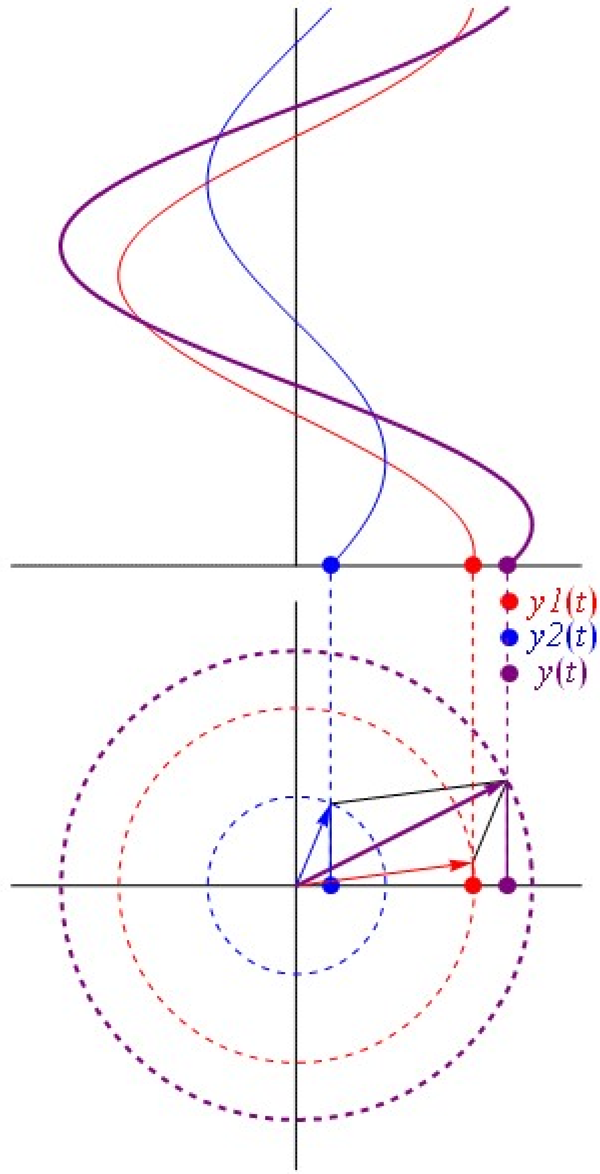
In polar coordinate form, it is:

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

Therefore:

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

Addition



The sum of phasors as addition of rotating vectors

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

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

where:

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

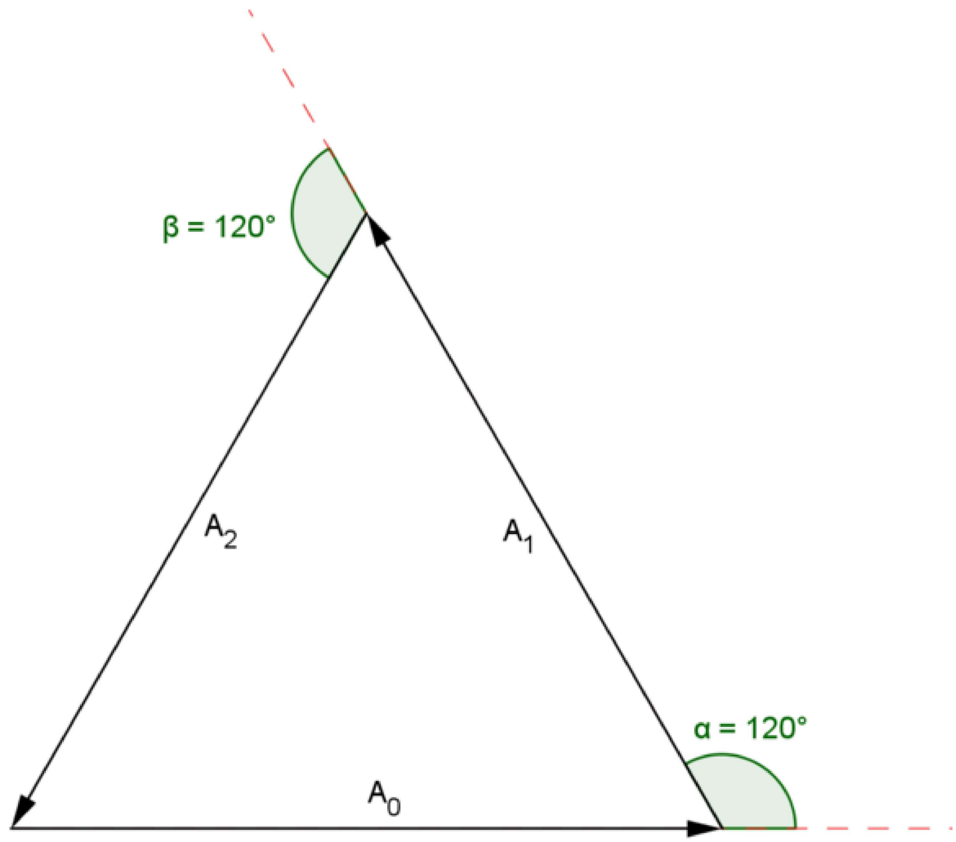
or, via the law of cosines on the complex plane (or the trigonometric identity for angle differences):

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

where $\Delta\theta = \theta_1 - \theta_2$. A key point is that A_3 and θ_3 do not depend on ω or t, which is what makes phasor notation possible. The time and frequency dependence can be suppressed and re-inserted into the outcome as long as the only operations used in between are ones that produce another phasor. In angle notation, the operation shown above is written:

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

Another way to view addition is that two **vectors** with coordinates [A_1 cos(ωt+θ_1), A_1 sin(ωt+θ_1)] and [A_2 cos(ωt+θ_2), A_2 sin(ωt+θ_2)] are added vectorially to produce a resultant vector with coordinates [A_3 cos(ωt+θ_3), A_3 sin(ωt+θ_3)].



Phasor diagram of three waves in perfect destructive interference

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

In other words, what this shows is:

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

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

Phasor diagrams

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

Circuit laws

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

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

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

Power engineering

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

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

Chapter- 4

Electrical Resistance & Reactance

Electrical resistance

A 65-k Ω resistor, as identified by its electronic color code. An ohmmeter could be used to verify this value.

The **electrical resistance** of an object is a measure of its opposition to the passage of an electric current. An object of uniform cross section will have a resistance proportional to its resistivity and length and inversely proportional to its cross-sectional area.

Discovered by Georg Ohm in 1827, electrical resistance shares some conceptual parallels with the mechanical notion of friction. The SI unit of electrical resistance is the ohm (Ω). Resistance's reciprocal quantity is electrical conductance measured in siemens.

The resistance of an object can be defined as the ratio of voltage to current:

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

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, such as resistors, 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

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 aluminum. 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 of a conductor of uniform cross section, therefore, can be computed as

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

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 (Greek: rho) is the electrical resistivity (also called *specific electrical resistance*) of the material, measured in ohm-metres (Ω m). Resistivity is a measure of the material's ability to oppose electric current.

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

If a wire conducts high-frequency alternating current, then the effective cross sectional area of the wire is reduced because of the skin effect. If several conductors are together, then due to proximity effect, the effective resistance of each is higher than if that conductor were alone. These effects are so small for low frequency of ordinary household AC that they should ordinarily be treated as if it were DC resistance.

Ohmmeter

An ohmmeter

An **ohmmeter** is an electrical instrument that measures electrical resistance, the opposition to an electric current. Micro-ohmmeters (microhmmeter or microohmmeter) make low resistance measurements. Megohmmeters (aka megaohmmeter or in the case of a trademarked device Megger) measure large values of resistance. The unit of measurement for resistance is ohms (Ω).

The original design of an ohmmeter provided a small battery to apply a voltage to a resistance. It uses a galvanometer to measure the electric current through the resistance. The scale of the galvanometer was marked in ohms, because the fixed voltage from the battery assured that as resistance is decreased, the current through the meter would increase.

A more accurate type of ohmmeter has an electronic circuit that passes a constant current (I) through the resistance, and another circuit that measures the voltage (V) across the resistance. According to the following equation, derived from Ohm's Law, the value of the resistance (R) is given by:

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

For high-precision measurements the above types of meter are inadequate. This is because the meter's reading is the sum of the resistance of the measuring leads, the contact resistances and the resistance being measured. To reduce this effect, a precision ohmmeter has four terminals, called Kelvin contacts. Two terminals carry the current from the meter, while the other two allow the meter to measure the voltage across the resistor. With this type of meter, any voltage drop due to the resistance of the first pair of leads and their contact resistances is ignored by the meter. This four terminal measurement technique is called Kelvin sensing, after William Thomson, Lord Kelvin, who invented the Kelvin bridge in 1861 to measure very low resistances. The Four-terminal sensing method can also be utilized to conduct accurate measurements of low resistances.

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.

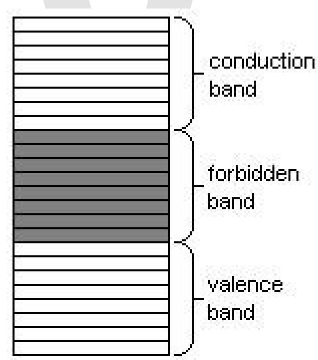
In ionic liquids/electrolytes

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:

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

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.

Electrical reactance

Reactance is the opposition of a circuit element to a change of current, caused by the build-up of electric or magnetic fields in the element. Those fields act to produce counter-emf that is proportional to either the rate of change (time derivative), or accumulation (time integral), of the current. An ideal resistor has zero reactance, while ideal inductors and capacitors consist entirely of reactance with zero resistance.

Analysis

In vector analysis, reactance is the imaginary part of electrical impedance, used to compute amplitude and phase changes of sinusoidal alternating current going through the circuit element. It is denoted by the symbol X. The SI unit of reactance is the same as that of resistance: the ohm.

Both reactance X and resistance R are required to calculate the impedance Z. In some circuits one of these may dominate, but an approximate knowledge of the minor component is useful to determine if it may be neglected.

$$Z = R + jX$$

where

$$j^2 = -1$$

Both the magnitude $|Z|$ and the phase θ of the impedance depend on both the resistance and the reactance.

$$|Z| = \sqrt{ZZ^*} = \sqrt{R^2 + X^2}$$ where Z^* is the complex conjugate of Z

$$\theta = \arctan \frac{X}{R}$$

The magnitude is the ratio of the voltage and current amplitudes, while the phase is the voltage–current phase difference.

- If $X > 0$, the reactance is said to be *inductive*
- If $X = 0$, then the impedance is purely *resistive*
- If $X < 0$, the reactance is said to be *capacitive*

The reciprocal of reactance (that is, 1 / X) is susceptance.

Chapter- 5

Capacitance & Inductance

Capacitance

In electromagnetism and electronics, **capacitance** is the ability of a body to hold an electrical charge. Capacitance is also a measure of the amount of electrical energy stored (or separated) for a given electric potential. A common form of energy storage device is a parallel-plate capacitor. In a parallel plate capacitor, capacitance is directly proportional to the surface area of the conductor plates and inversely proportional to the separation distance between the plates. If the charges on the plates are $+Q$ and $-Q$, and V gives the voltage between the plates, then the capacitance is given by

$$C = \frac{Q}{V}.$$

The SI unit of capacitance is the farad; 1 farad is 1 coulomb per volt.

The energy (measured in joules) stored in a capacitor is equal to the *work* done to charge it. Consider a capacitance C, holding a charge $+q$ on one plate and $-q$ on the other. Moving a small element of charge $\mathrm{d}q$ from one plate to the other against the potential difference $V = q/C$ requires the work $\mathrm{d}W$:

$$\mathrm{d}W = \frac{q}{C}\,\mathrm{d}q$$

where W is the work measured in joules, q is the charge measured in coulombs and C is the capacitance, measured in farads.

The energy stored in a capacitance is found by integrating this equation. Starting with an uncharged capacitance ($q = 0$) and moving charge from one plate to the other until the plates have charge $+Q$ and $-Q$ requires the work W:

$$W_{\text{charging}} = \int_0^Q \frac{q}{C}\,\mathrm{d}q = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}CV^2 = W_{\text{stored}}.$$

Capacitor

Capacitor

Modern capacitors, by a cm rule

Type	Passive
Invented	Ewald Georg von Kleist (October 1745)

Electronic symbol

A typical electrolytic capacitor

A **capacitor** (formerly known as **condenser**) is a passive electronic component consisting of a pair of conductors separated by a dielectric (insulator). When there is a potential difference (voltage) across the conductors, a static electric field develops in the dielectric that stores energy and produces a mechanical force between the conductors. An ideal capacitor is characterized by a single constant value, capacitance, measured in farads. This is the ratio of the electric charge on each conductor to the potential difference between them.

Capacitors are widely used in electronic circuits for blocking direct current while allowing alternating current to pass, in filter networks, for smoothing the output of power supplies, in the resonant circuits that tune radios to particular frequencies and for many other purposes.

The effect is greatest when there is a narrow separation between large areas of conductor, hence capacitor conductors are often called "plates", referring to an early means of construction. In practice the dielectric between the plates passes a small amount of leakage current and also has an electric field strength limit, resulting in a breakdown voltage, while the conductors and leads introduce an undesired inductance and resistance.

History

Battery of four Leyden jars in Museum Boerhaave, Leiden, the Netherlands

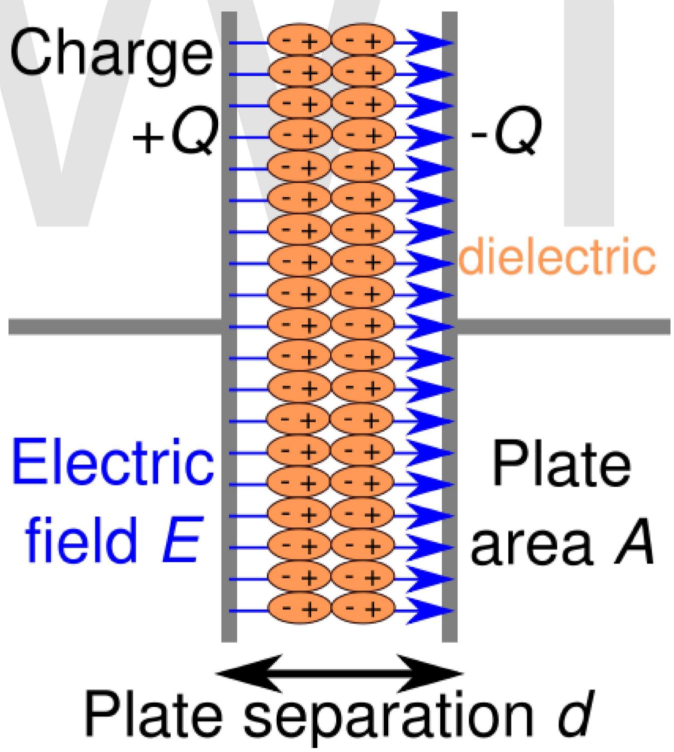
In October 1745, Ewald Georg von Kleist of Pomerania in Germany found that charge could be stored by connecting a high voltage electrostatic generator by a wire to a volume of water in a hand-held glass jar. Von Kleist's hand and the water acted as conductors and the jar as a dielectric (although details of the mechanism were incorrectly identified at the time). Von Kleist found, after removing the generator, that touching the wire resulted in a painful spark. In a letter describing the experiment, he said "I would not take a second shock for the kingdom of France." The following year, the Dutch physicist Pieter van Musschenbroek invented a similar capacitor, which was named the Leyden jar, after the University of Leiden where he worked.

Daniel Gralath was the first to combine several jars in parallel into a "battery" to increase the charge storage capacity. Benjamin Franklin investigated the Leyden jar and "proved" that the charge was stored on the glass, not in the water as others had assumed. He also adopted the term "battery", (denoting the increasing of power with a row of similar units as in a battery of cannon), subsequently applied to clusters of electrochemical cells. Leyden jars were later made by coating the inside and outside of jars with metal foil, leaving a space at the mouth to prevent arcing between the foils. The earliest unit of capacitance was the 'jar', equivalent to about 1 nanofarad.

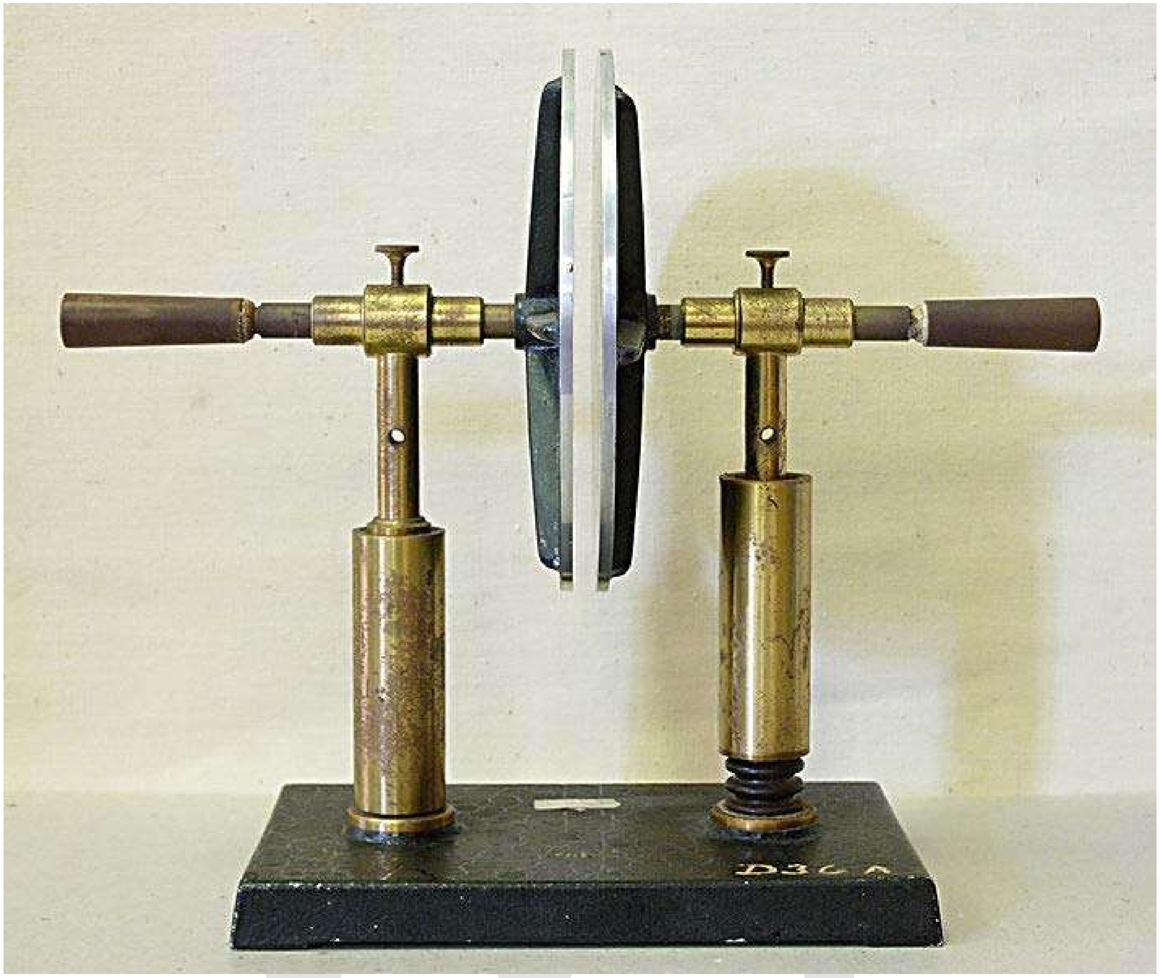
Leyden jars or more powerful devices employing flat glass plates alternating with foil conductors were used exclusively up until about 1900, when the invention of wireless (radio) created a demand for standard capacitors, and the steady move to higher frequencies required capacitors with lower inductance. A more compact construction began to be used of a flexible dielectric sheet such as oiled paper sandwiched between sheets of metal foil, rolled or folded into a small package.

Early capacitors were also known as *condensers*, a term that is still occasionally used today. The term was first used for this purpose by Alessandro Volta in 1782, with reference to the device's ability to store a higher density of electric charge than a normal isolated conductor.

Theory of operation



Charge separation in a parallel-plate capacitor causes an internal electric field. A dielectric (orange) reduces the field and increases the capacitance.



A simple demonstration of a parallel-plate capacitor

A capacitor consists of two conductors separated by a non-conductive region called the dielectric medium though it may be a vacuum or a semiconductor depletion region chemically identical to the conductors. A capacitor is assumed to be self-contained and isolated, with no net electric charge and no influence from any external electric field. The conductors thus hold equal and opposite charges on their facing surfaces, and the dielectric develops an electric field. In SI units, a capacitance of one farad means that one coulomb of charge on each conductor causes a voltage of one volt across the device.

The capacitor is a reasonably general model for electric fields within electric circuits. An ideal capacitor is wholly characterized by a constant capacitance *C*, defined as the ratio of charge ±*Q* on each conductor to the voltage *V* between them:

$$C = \frac{Q}{V}$$

Sometimes charge build-up affects the capacitor mechanically, causing its capacitance to vary. In this case, capacitance is defined in terms of incremental changes:

$$C = \frac{\mathrm{d}q}{\mathrm{d}v}$$

Energy storage

Work must be done by an external influence to "move" charge between the conductors in a capacitor. When the external influence is removed the charge separation persists in the electric field and energy is stored to be released when the charge is allowed to return to its equilibrium position. The work done in establishing the electric field, and hence the amount of energy stored, is given by:

$$W = \int_{q=0}^{Q} V \mathrm{d}q = \int_{q=0}^{Q} \frac{q}{C} \mathrm{d}q = \frac{1}{2} \frac{Q^2}{C} = \frac{1}{2} C V^2 = \frac{1}{2} V Q.$$

Current-voltage relation

The current $i(t)$ through any component in an electric circuit is defined as the rate of flow of a charge $q(t)$ passing through it, but actual charges, electrons, cannot pass through the dielectric layer of a capacitor, rather an electron accumulates on the negative plate for each one that leaves the positive plate, resulting in an electron depletion and consequent positive charge on one electrode that is equal and opposite to the accumulated negative charge on the other. Thus the charge on the electrodes is equal to the integral of the current as well as proportional to the voltage as discussed above. As with any antiderivative, a constant of integration is added to represent the initial voltage $v\ (t_0)$. This is the integral form of the capacitor equation,

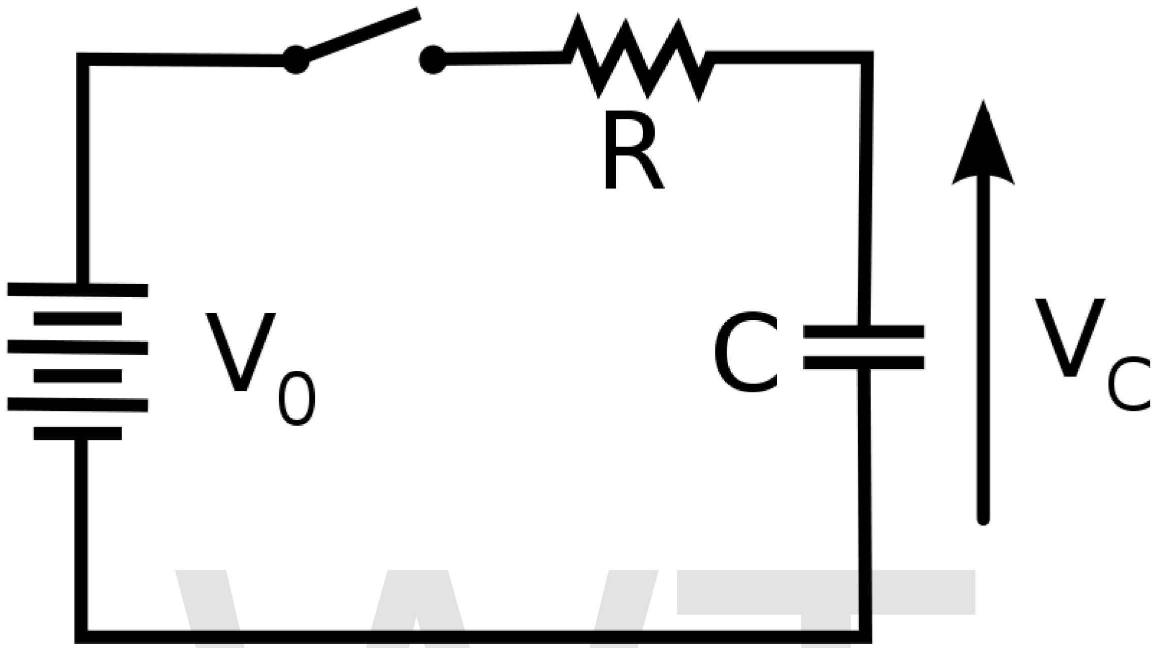
$$v(t) = \frac{q(t)}{C} = \frac{1}{C} \int_{t_0}^{t} i(\tau) \mathrm{d}\tau + v(t_0)$$

Taking the derivative of this, and multiplying by C, yields the derivative form,

$$i(t) = \frac{\mathrm{d}q(t)}{\mathrm{d}t} = C \frac{\mathrm{d}v(t)}{\mathrm{d}t}$$

The dual of the capacitor is the inductor, which stores energy in the magnetic field rather than the electric field. Its current-voltage relation is obtained by exchanging current and voltage in the capacitor equations and replacing C with the inductance L.

DC circuits



A simple resistor-capacitor circuit demonstrates charging of a capacitor

A series circuit containing only a resistor, a capacitor, a switch and a constant DC source of voltage V_0 is known as a *charging circuit*. If the capacitor is initially uncharged while the switch is open, and the switch is closed at $t = 0$, it follows from Kirchhoff's voltage law that

$$V_0 = v_{\text{resistor}}(t) + v_{\text{capacitor}}(t) = i(t)R + \frac{1}{C}\int_0^t i(\tau)\mathrm{d}\tau.$$

Taking the derivative and multiplying by C, gives a first-order differential equation,

$$RC\frac{\mathrm{d}i(t)}{\mathrm{d}t} + i(t) = 0.$$

At $t = 0$, the voltage across the capacitor is zero and the voltage across the resistor is V_0. The initial current is then $i(0) = V_0/R$. With this assumption, the differential equation yields

$$i(t) = \frac{V_0}{R}e^{-t/\tau_0}$$
$$v(t) = V_0\left(1 - e^{-t/\tau_0}\right),$$

where $\tau_0 = RC$ is the *time constant* of the system.

As the capacitor reaches equilibrium with the source voltage, the voltage across the resistor and the current through the entire circuit decay exponentially. The case of *discharging* a charged capacitor likewise demonstrates exponential decay, but with the initial capacitor voltage replacing V_0 and the final voltage being zero.

AC circuits

Impedance, the vector sum of reactance and resistance, describes the phase difference and the ratio of amplitudes between sinusoidally varying voltage and sinusoidally varying current at a given frequency. Fourier analysis allows any signal to be constructed from a spectrum of frequencies, whence the circuit's reaction to the various frequencies may be found. The reactance and impedance of a capacitor are respectively

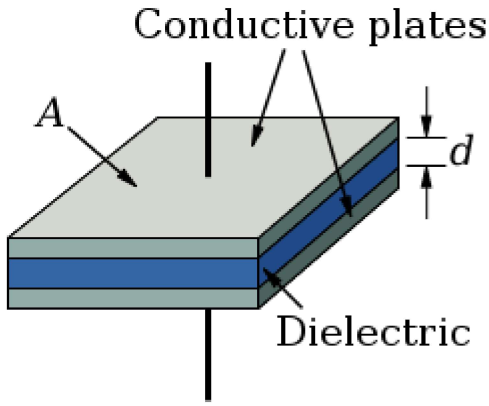
$$X = -\frac{1}{\omega C} = -\frac{1}{2\pi f C}$$
$$Z = \frac{1}{j\omega C} = -\frac{j}{\omega C} = -\frac{j}{2\pi f C}$$

where j is the imaginary unit and ω is the angular velocity of the sinusoidal signal. The $-j$ phase indicates that the AC voltage $V = Z\,I$ lags the AC current by 90°: the positive current phase corresponds to increasing voltage as the capacitor charges; zero current corresponds to instantaneous constant voltage, etc.

Note that impedance decreases with increasing capacitance and increasing frequency. This implies that a higher-frequency signal or a larger capacitor results in a lower voltage amplitude per current amplitude—an AC "short circuit" or AC coupling. Conversely, for very low frequencies, the reactance will be high, so that a capacitor is nearly an open circuit in AC analysis—those frequencies have been "filtered out".

Capacitors are different from resistors and inductors in that the impedance is *inversely* proportional to the defining characteristic, i.e. capacitance.

Parallel plate model



Dielectric is placed between two conducting plates, each of area *A* and with a separation of *d*.

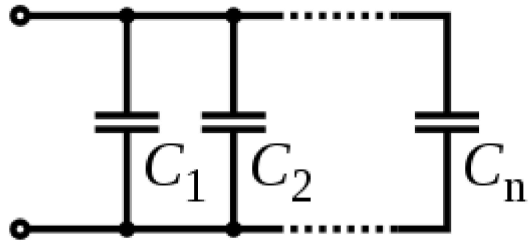
The simplest capacitor consists of two parallel conductive plates separated by a dielectric with permittivity ε (such as air). The model may also be used to make qualitative predictions for other device geometries. The plates are considered to extend uniformly over an area *A* and a charge density $\pm\rho = \pm Q/A$ exists on their surface. Assuming that the width of the plates is much greater than their separation *d*, the electric field near the centre of the device will be uniform with the magnitude $E = \rho/\varepsilon$. The voltage is defined as the line integral of the electric field between the plates

$$V = \int_0^d E \mathrm{d}z = \int_0^d \frac{\rho}{\varepsilon} \mathrm{d}z = \frac{\rho d}{\varepsilon} = \frac{Qd}{\varepsilon A}.$$

Solving this for $C = Q/V$ reveals that capacitance increases with area and decreases with separation

$$C = \frac{\varepsilon A}{d}.$$

The capacitance is therefore greatest in devices made from materials with a high permittivity.



Several capacitors in parallel

Networks

For capacitors in parallel

Capacitors in a parallel configuration each have the same applied voltage. Their capacitances add up. Charge is apportioned among them by size. Using the schematic diagram to visualize parallel plates, it is apparent that each capacitor contributes to the total surface area.

$$C_{eq} = C_1 + C_2 + \cdots + C_n$$

For capacitors in series

Several capacitors in series.

Connected in series, the schematic diagram reveals that the separation distance, not the plate area, adds up. The capacitors each store instantaneous charge build-up equal to that of every other capacitor in the series. The total voltage difference from end to end is apportioned to each capacitor according to the inverse of its capacitance. The entire series acts as a capacitor *smaller* than any of its components.

$$\frac{1}{C_{eq}} = \frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n}$$

Capacitors are combined in series to achieve a higher working voltage, for example for smoothing a high voltage power supply. The voltage ratings, which are based on plate separation, add up. In such an application, several series connections may in turn be connected in parallel, forming a matrix. The goal is to maximize the energy storage utility of each capacitor without overloading it. Series connection is also used to adapt electrolytic capacitors for AC use.

Non-ideal behavior

Capacitors deviate from the ideal capacitor equation in a number of ways. Some of these, such as leakage current and parasitic effects are linear, or can be assumed to be linear, and can be dealt with by adding virtual components to the equivalent circuit of the capacitor. The usual methods of network analysis can then be applied. In other cases, such as with breakdown voltage, the effect is non-linear and normal (i.e., linear) network analysis cannot be used, the effect must be dealt with separately. There is yet another group, which may be linear but invalidate the assumption in the analysis that capacitance is a constant. Such an example is temperature dependence.

Breakdown voltage

Above a particular electric field, known as the dielectric strength E_{ds}, the dielectric in a capacitor becomes conductive. The voltage at which this occurs is called the breakdown voltage of the device, and is given by the product of the dielectric strength and the separation between the conductors,

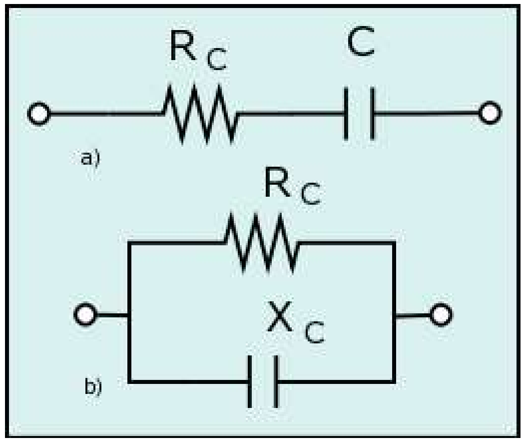
$$V_{\mathrm{bd}} = E_{\mathrm{ds}}d$$

The maximum energy that can be stored safely in a capacitor is limited by the breakdown voltage. Due to the scaling of capacitance and breakdown voltage with dielectric thickness, all capacitors made with a particular dielectric have approximately equal maximum energy density, to the extent that the dielectric dominates their volume.

For air dielectric capacitors the breakdown field strength is of the order 2 to 5 MV/m; for mica the breakdown is 100 to 300 MV/m, for oil 15 to 25 MV/m, and can be much less when other materials are used for the dielectric. The dielectric is used in very thin layers and so absolute breakdown voltage of capacitors is limited. Typical ratings for capacitors used for general electronics applications range from a few volts to 100V or so. As the voltage increases, the dielectric must be thicker, making high-voltage capacitors larger than those rated for lower voltages. The breakdown voltage is critically affected by factors such as the geometry of the capacitor conductive parts; sharp edges or points increase the electric field strength at that point and can lead to a local breakdown. Once this starts to happen, the breakdown will quickly "track" through the dielectric till it reaches the opposite plate and cause a short circuit.

The usual breakdown route is that the field strength becomes large enough to pull electrons in the dielectric from their atoms thus causing conduction. Other scenarios are possible, such as impurities in the dielectric, and, if the dielectric is of a crystalline nature, imperfections in the crystal structure can result in an avalanche breakdown as seen in semi-conductor devices. Breakdown voltage is also affected by pressure, humidity and temperature.

Equivalent circuit



Two equivalent circuits of a real capacitor

An ideal capacitor only stores and releases electrical energy, without dissipating any. In reality, all capacitors have imperfections within the capacitor's material that create resistance. This is specified as the *equivalent series resistance* or **ESR** of a component. This adds a real component to the impedance:

$$R_C = Z + R_{ESR} = \frac{1}{j\omega C} + R_{ESR}$$

As frequency approaches infinity, the capacitive impedance (or reactance) approaches zero and the ESR becomes significant. As the reactance becomes negligible, power dissipation approaches $P_{RMS.} = V_{RMS.}^2 / R_{ESR}$.

Similarly to ESR, the capacitor's leads add *equivalent series inductance* or **ESL** to the component. This is usually significant only at relatively high frequencies. As inductive reactance is positive and increases with frequency, above a certain frequency capacitance will be canceled by inductance. High frequency engineering involves accounting for the inductance of all connections and components.

If the conductors are separated by a material with a small conductivity rather than a perfect dielectric, then a small leakage current flows directly between them. The capacitor therefore has a finite parallel resistance, and slowly discharges over time (time may vary greatly depending on the capacitor material and quality).

Ripple current

Ripple current is the AC component of an applied source (often a switched-mode power supply) whose frequency may be constant or varying. Certain types of capacitors, such as electrolytic tantalum capacitors, usually have a rating for maximum ripple current (both in frequency and magnitude). This ripple current can cause damaging heat to be generated within the capacitor due to the current flow across resistive imperfections in the materials used within the capacitor, more commonly referred to as equivalent series resistance (ESR). For example electrolytic tantalum capacitors are limited by ripple current and generally have the highest ESR ratings in the capacitor family, while ceramic capacitors generally have no ripple current limitation and have some of the lowest ESR ratings.

Capacitance instability

The capacitance of certain capacitors decreases as the component ages. In ceramic capacitors, this is caused by degradation of the dielectric. The type of dielectric and the ambient operating and storage temperatures are the most significant aging factors, while the operating voltage has a smaller effect. The aging process may be reversed by heating the component above the Curie point. Aging is fastest near the beginning of life of the component, and the device stabilizes over time. Electrolytic capacitors age as the electrolyte evaporates. In contrast with ceramic capacitors, this occurs towards the end of life of the component.

Temperature dependence of capacitance is usually expressed in parts per million (ppm) per °C. It can usually be taken as a broadly linear function but can be noticeably non-linear at the temperature extremes. The temperature coefficient can be either positive or negative, sometimes even amongst different samples of the same type. In other words, the spread in the range of temperature coefficients can encompass zero.

Capacitors, especially ceramic capacitors, and older designs such as paper capacitors, can absorb sound waves resulting in a microphonic effect. Vibration moves the plates, causing the capacitance to vary, in turn inducing AC current. Some dielectrics also generate piezoelectricity. The resulting interference is especially problematic in audio applications, potentially causing feedback or unintended recording. In the reverse microphonic effect, the varying electric field between the capacitor plates exerts a physical force, moving them as a speaker. This can generate audible sound, but drains energy and stresses the dielectric and the electrolyte, if any.

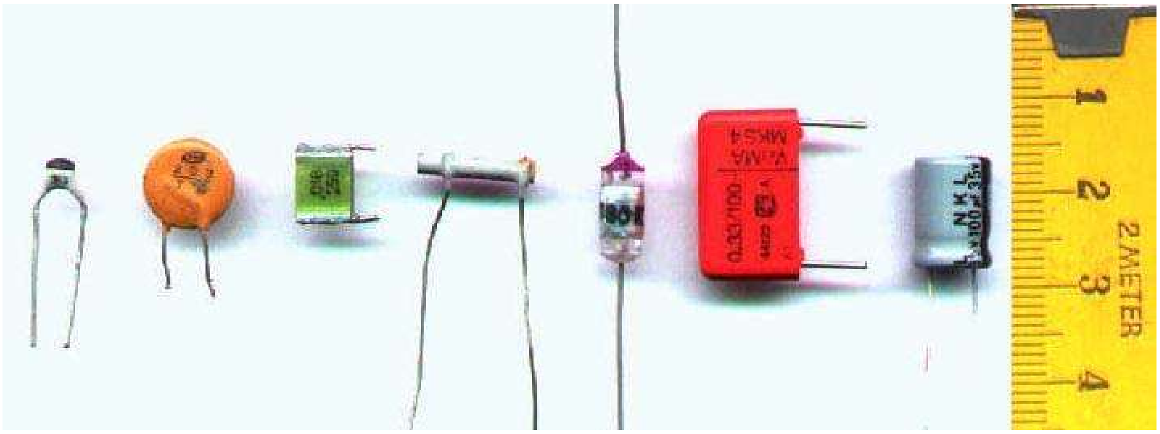
Capacitor types

Practical capacitors are available commercially in many different forms. The type of internal dielectric, the structure of the plates and the device packaging all strongly affect the characteristics of the capacitor, and its applications.

Values available range from very low (picofarad range; while arbitrarily low values are in principle possible, stray (parasitic) capacitance in any circuit is the limiting factor) to about 5 kF supercapacitors.

Above approximately 1 microfarad electrolytic capacitors are usually used because of their small size and low cost compared with other technologies, unless their relatively poor stability, life and polarised nature make them unsuitable. Very high capacity supercapacitors use a porous carbon-based electrode material.

Dielectric materials



Capacitor materials. From left: multilayer ceramic, ceramic disc, multilayer polyester film, tubular ceramic, polystyrene, metalized polyester film, aluminum electrolytic. Major scale divisions are in centimetres.

Most types of capacitor include a dielectric spacer, which increases their capacitance. These dielectrics are most often insulators. However, low capacitance devices are available with a vacuum between their plates, which allows extremely high voltage operation and low losses. Variable capacitors with their plates open to the atmosphere were commonly used in radio tuning circuits. Later designs use polymer foil dielectric between the moving and stationary plates, with no significant air space between them.

In order to maximise the charge that a capacitor can hold, the dielectric material needs to have as high a permittivity as possible, while also having as high a breakdown voltage as possible.

Several solid dielectrics are available, including paper, plastic, glass, mica and ceramic materials. Paper was used extensively in older devices and offers relatively high voltage performance. However, it is susceptible to water absorption, and has been largely replaced by plastic film capacitors. Plastics offer better stability and aging performance, which makes them useful in timer circuits, although they may be limited to low operating temperatures and frequencies. Ceramic capacitors are generally small, cheap and useful for high frequency applications, although their capacitance varies strongly with voltage and they age poorly. They are broadly categorized as class 1 dielectrics, which have predictable variation of capacitance with temperature or class 2 dielectrics, which can operate at higher voltage. Glass and mica capacitors are extremely reliable, stable and tolerant to high temperatures and voltages, but are too expensive for most mainstream applications. Electrolytic capacitors and supercapacitors are used to store small and larger amounts of energy, respectively, ceramic capacitors are often used in resonators, and parasitic capacitance occurs in circuits wherever the simple conductor-insulator-conductor structure is formed unintentionally by the configuration of the circuit layout.

Electrolytic capacitors use an aluminum or tantalum plate with an oxide dielectric layer. The second electrode is a liquid electrolyte, connected to the circuit by another foil plate. Electrolytic capacitors offer very high capacitance but suffer from poor tolerances, high instability, gradual loss of capacitance especially when subjected to heat, and high leakage current. Poor quality capacitors may leak electrolyte, which is harmful to printed circuit boards. The conductivity of the electrolyte drops at low temperatures, which increases equivalent series resistance. While widely used for power-supply conditioning, poor high-frequency characteristics make them unsuitable for many applications. Electrolytic capacitors will self-degrade if unused for a period (around a year), and when full power is applied may short circuit, permanently damaging the capacitor and usually blowing a fuse or causing arcing in rectifier tubes. They can be restored before use (and damage) by gradually applying the operating voltage, often done on antique vacuum tube equipment over a period of 30 minutes by using a variable transformer to supply AC power. Unfortunately, the use of this technique may be less satisfactory for some solid state equipment, which may be damaged by operation below its normal power range, requiring that the power supply first be isolated from the consuming circuits. Such remedies may not be applicable to modern high-frequency power supplies as these produce full output voltage even with reduced input.

Tantalum capacitors offer better frequency and temperature characteristics than aluminum, but higher dielectric absorption and leakage. OS-CON (or OC-CON) capacitors are a polymerized organic semiconductor solid-electrolyte type that offer longer life at higher cost than standard electrolytic capacitors.

Several other types of capacitor are available for specialist applications. Supercapacitors store large amounts of energy. Supercapacitors made from carbon aerogel, carbon nanotubes, or highly porous electrode materials offer extremely high capacitance (up to 5 kF as of 2010) and can be used in some applications instead of rechargeable batteries. Alternating current capacitors are specifically designed to work on line (mains) voltage AC power circuits. They are commonly used in electric motor circuits and are often designed to handle large currents, so they tend to be physically large. They are usually ruggedly packaged, often in metal cases that can be easily grounded/earthed. They also are designed with direct current breakdown voltages of at least five times the maximum AC voltage.

Structure

Capacitor packages: SMD ceramic at top left; SMD tantalum at bottom left; through-hole tantalum at top right; through-hole electrolytic at bottom right. Major scale divisions are cm.

The arrangement of plates and dielectric has many variations depending on the desired ratings of the capacitor. For small values of capacitance (microfarads and less), ceramic disks use metallic coatings, with wire leads bonded to the coating. Larger values can be made by multiple stacks of plates and disks. Larger value capacitors usually use a metal foil or metal film layer deposited on the surface of a dielectric film to make the plates, and a dielectric film of impregnated paper or plastic – these are rolled up to save space. To reduce the series resistance and inductance for long plates, the plates and dielectric are staggered so that connection is made at the common edge of the rolled-up plates, not at the ends of the foil or metalized film strips that comprise the plates.

The assembly is encased to prevent moisture entering the dielectric – early radio equipment used a cardboard tube sealed with wax. Modern paper or film dielectric capacitors are dipped in a hard thermoplastic. Large capacitors for high-voltage use may have the roll form compressed to fit into a rectangular metal case, with bolted terminals and bushings for connections. The dielectric in larger capacitors is often impregnated with a liquid to improve its properties.

Capacitors may have their connecting leads arranged in many configurations, for example axially or radially. "Axial" means that the leads are on a common axis, typically the axis of the capacitor's cylindrical body – the leads extend from opposite ends. Radial leads might more accurately be referred to as tandem; they are rarely actually aligned along radii of the body's circle, so the term is inexact, although universal. The leads (until bent) are usually in planes parallel to that of the flat body of the capacitor, and extend in the same direction; they are often parallel as manufactured.

Small, cheap discoidal ceramic capacitors have existed since the 1930s, and remain in widespread use. Since the 1980s, surface mount packages for capacitors have been widely used. These packages are extremely small and lack connecting leads, allowing them to be soldered directly onto the surface of printed circuit boards. Surface mount components avoid undesirable high-frequency effects due to the leads and simplify automated assembly, although manual handling is made difficult due to their small size.

Mechanically controlled variable capacitors allow the plate spacing to be adjusted, for example by rotating or sliding a set of movable plates into alignment with a set of stationary plates. Low cost variable capacitors squeeze together alternating layers of aluminum and plastic with a screw. Electrical control of capacitance is achievable with varactors (or varicaps), which are reverse-biased semiconductor diodes whose depletion region width varies with applied voltage. They are used in phase-locked loops, amongst other applications.

Capacitor markings

Most capacitors have numbers printed on their bodies to indicate their electrical characteristics. Larger capacitors like electrolytics usually display the actual capacitance together with the unit (for example, **220 μF**). Smaller capacitors like ceramics, however, use a shorthand consisting of three numbers and a letter, where the numbers show the capacitance in pF (calculated as XY x 10^Z for the numbers XYZ) and the letter indicates the tolerance (J, K or M for ±5%, ±10% and ±20% respectively).

Additionally, the capacitor may show its working voltage, temperature and other relevant characteristics.

Example

A capacitor with the text **473K 330V** on its body has a capacitance of 47 x 10^3 pF = 47 nF (±10%) with a working voltage of 330 V.

Applications

Capacitors have many uses in electronic and electrical systems. They are so common that it is a rare electrical product that does not include at least one for some purpose.

Energy storage

A capacitor can store electric energy when disconnected from its charging circuit, so it can be used like a temporary battery. Capacitors are commonly used in electronic devices to maintain power supply while batteries are being changed. (This prevents loss of information in volatile memory.)

Conventional electrostatic capacitors provide less than 360 joules per kilogram of energy density, while capacitors using developing technologies can provide more than 2.52 kilojoules per kilogram.

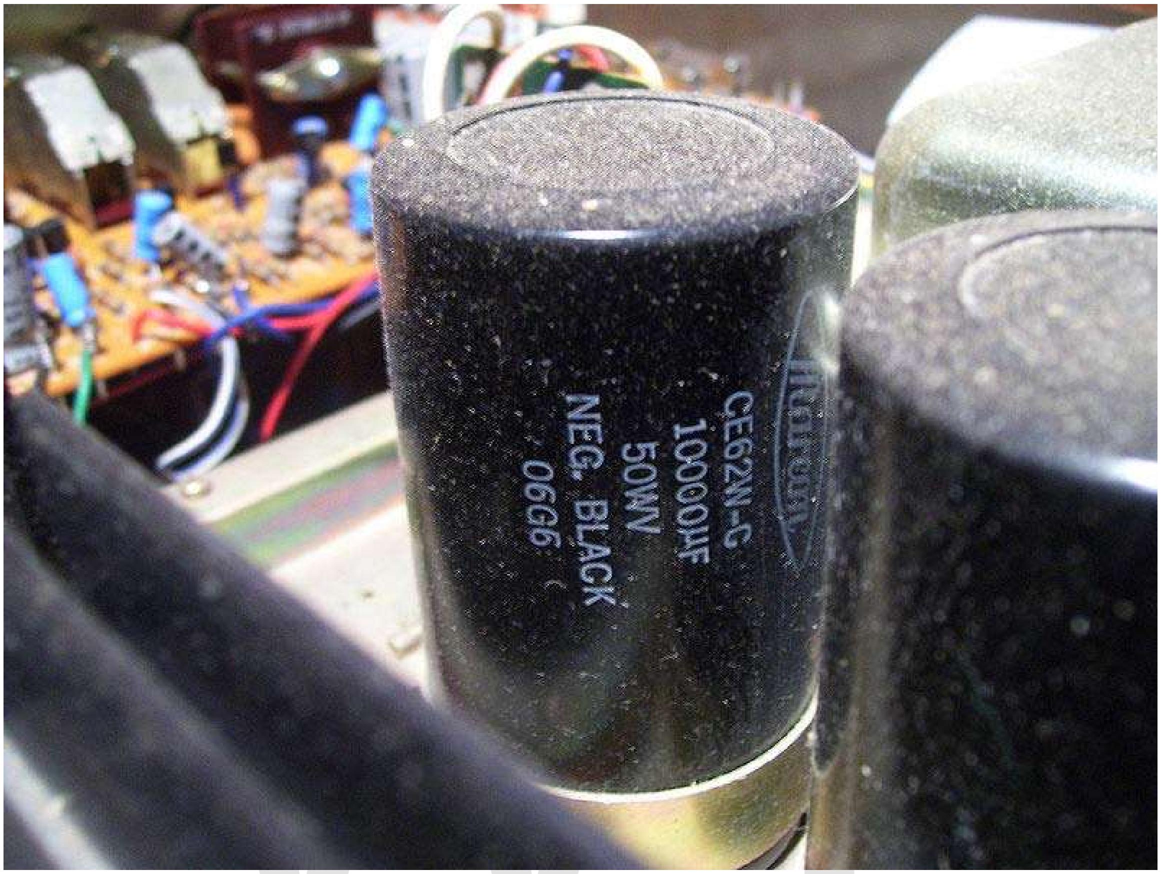
In car audio systems, large capacitors store energy for the amplifier to use on demand. Also for a flash tube a capacitor is used to hold the high voltage. In ceiling fans, capacitors play the important role of storing electrical energy to give the fan enough torque to start spinning.

Pulsed power and weapons

Groups of large, specially constructed, low-inductance high-voltage capacitors (*capacitor banks*) are used to supply huge pulses of current for many pulsed power applications. These include electromagnetic forming, Marx generators, pulsed lasers (especially TEA lasers), pulse forming networks, radar, fusion research, and particle accelerators.

Large capacitor banks (reservoir) are used as energy sources for the exploding-bridgewire detonators or slapper detonators in nuclear weapons and other specialty weapons. Experimental work is under way using banks of capacitors as power sources for electromagnetic armour and electromagnetic railguns and coilguns.

Power conditioning



A 10,000 microfarad capacitor in a TRM-800 amplifier

Reservoir capacitors are used in power supplies where they smooth the output of a full or half wave rectifier. They can also be used in charge pump circuits as the energy storage element in the generation of higher voltages than the input voltage.

Capacitors are connected in parallel with the power circuits of most electronic devices and larger systems (such as factories) to shunt away and conceal current fluctuations from the primary power source to provide a "clean" power supply for signal or control circuits. Audio equipment, for example, uses several capacitors in this way, to shunt away power line hum before it gets into the signal circuitry. The capacitors act as a local reserve for the DC power source, and bypass AC currents from the power supply. This is used in car audio applications, when a stiffening capacitor compensates for the inductance and resistance of the leads to the lead-acid car battery.

Power factor correction

In electric power distribution, capacitors are used for power factor correction. Such capacitors often come as three capacitors connected as a three phase load. Usually, the values of these capacitors are given not in farads but rather as a reactive power in volt-amperes reactive (VAr). The purpose is to counteract inductive loading from devices like electric motors and transmission lines to make the load appear to be mostly resistive. Individual motor or lamp loads may have capacitors for power factor correction, or larger sets of capacitors (usually with automatic switching devices) may be installed at a load center within a building or in a large utility substation.

Supression and coupling

Signal coupling

Because capacitors pass AC but block DC signals (when charged up to the applied dc voltage), they are often used to separate the AC and DC components of a signal. This method is known as *AC coupling* or "capacitive coupling". Here, a large value of capacitance, whose value need not be accurately controlled, but whose reactance is small at the signal frequency, is employed.

Decoupling

A decoupling capacitor is a capacitor used to protect one part of a circuit from the effect of another, for instance to suppress noise or transients. Noise caused by other circuit elements is shunted through the capacitor, reducing the effect they have on the rest of the circuit. It is most commonly used between the power supply and ground. An alternative name is *bypass capacitor* as it is used to bypass the power supply or other high impedance component of a circuit.

Noise filters and snubbers

When an inductive circuit is opened, the current through the inductance collapses quickly, creating a large voltage across the open circuit of the switch or relay. If the inductance is large enough, the energy will generate a spark, causing the contact points to oxidize, deteriorate, or sometimes weld together, or destroying a solid-state switch. A snubber capacitor across the newly opened circuit creates a path for this impulse to bypass the contact points, thereby preserving their life; these were commonly found in contact breaker ignition systems, for instance. Similarly, in smaller scale circuits, the spark may not be enough to damage the switch but will still radiate undesirable radio frequency interference (RFI), which a filter capacitor absorbs. Snubber capacitors are usually employed with a low-value resistor in series, to dissipate energy and minimize RFI. Such resistor-capacitor combinations are available in a single package.

Capacitors are also used in parallel to interrupt units of a high-voltage circuit breaker in order to equally distribute the voltage between these units. In this case they are called grading capacitors.

In schematic diagrams, a capacitor used primarily for DC charge storage is often drawn vertically in circuit diagrams with the lower, more negative, plate drawn as an arc. The straight plate indicates the positive terminal of the device, if it is polarized.

Motor starters

In single phase squirrel cage motors, the primary winding within the motor housing is not capable of starting a rotational motion on the rotor, but is capable of sustaining one. To start the motor, a secondary winding is used in series with a non-polarized *starting capacitor* to introduce a lag in the sinusoidal current through the starting winding. When the secondary winding is placed at an angle with respect to the primary winding, a rotating electric field is created. The force of the rotational field is not constant, but is sufficient to start the rotor spinning. When the rotor comes close to operating speed, a centrifugal switch (or current-sensitive relay in series with the main winding) disconnects the capacitor. The start capacitor is typically mounted to the side of the motor housing. These are called capacitor-start motors, that have relatively high starting torque.

There are also capacitor-run induction motors which have a permanently connected phase-shifting capacitor in series with a second winding. The motor is much like a two-phase induction motor.

Motor-starting capacitors are typically non-polarized electrolytic types, while running capacitors are conventional paper or plastic film dielectric types.

Signal processing

The energy stored in a capacitor can be used to represent information, either in binary form, as in DRAMs, or in analogue form, as in analog sampled filters and CCDs. Capacitors can be used in analog circuits as components of integrators or more complex filters and in negative feedback loop stabilization. Signal processing circuits also use capacitors to integrate a current signal.

Tuned circuits

Capacitors and inductors are applied together in tuned circuits to select information in particular frequency bands. For example, radio receivers rely on variable capacitors to tune the station frequency. Speakers use passive analog crossovers, and analog equalizers use capacitors to select different audio bands.

The resonant frequency *f* of a tuned circuit is a function of the inductance (*L*) and capacitance (*C*) in series, and is given by:

$$f = \frac{1}{2\pi\sqrt{LC}}$$

where *L* is in henries and *C* is in farads.

Sensing

Most capacitors are designed to maintain a fixed physical structure. However, various factors can change the structure of the capacitor, and the resulting change in capacitance can be used to sense those factors.

Changing the dielectric:

The effects of varying the physical and/or electrical characteristics of the **dielectric** can be used for sensing purposes. Capacitors with an exposed and porous dielectric can be used to measure humidity in air. Capacitors are used to accurately measure the fuel level in airplanes; as the fuel covers more of a pair of plates, the circuit capacitance increases.

Changing the distance between the plates:

Capacitors with a flexible plate can be used to measure strain or pressure. Industrial pressure transmitters used for process control use pressure-sensing diaphragms, which form a capacitor plate of an oscillator circuit. Capacitors are used as the sensor in condenser microphones, where one plate is moved by air pressure, relative to the fixed position of the other plate. Some accelerometers use MEMS capacitors etched on a chip to measure the magnitude and direction of the acceleration vector. They are used to detect changes in acceleration, e.g. as tilt sensors or to detect free fall, as sensors triggering airbag deployment, and in many other applications. Some fingerprint sensors use capacitors. Additionally, a user can adjust the pitch of a theremin musical instrument by moving his hand since this changes the effective capacitance between the user's hand and the antenna.

Changing the effective area of the plates:

Capacitive touch switches are now used on many consumer electronic products.

Hazards and safety

Capacitors may retain a charge long after power is removed from a circuit; this charge can cause dangerous or even potentially fatal shocks or damage connected equipment. For example, even a seemingly innocuous device such as a disposable camera flash unit powered by a 1.5 volt AA battery contains a capacitor which may be charged to over 300 volts. This is easily capable of delivering a shock. Service procedures for electronic devices usually include instructions to discharge large or high-voltage capacitors. Capacitors may also have built-in discharge resistors to dissipate stored energy to a safe level within a few seconds after power is removed. High-voltage capacitors are stored with the terminals shorted, as protection from potentially dangerous voltages due to dielectric absorption.

Some old, large oil-filled capacitors contain polychlorinated biphenyls (PCBs). It is known that waste PCBs can leak into groundwater under landfills. Capacitors containing

PCB were labelled as containing "Askarel" and several other trade names. PCB-filled capacitors are found in very old (pre 1975) fluorescent lamp ballasts, and other applications.

High-voltage capacitors may catastrophically fail when subjected to voltages or currents beyond their rating, or as they reach their normal end of life. Dielectric or metal interconnection failures may create arcing that vaporizes dielectric fluid, resulting in case bulging, rupture, or even an explosion. Capacitors used in RF or sustained high-current applications can overheat, especially in the center of the capacitor rolls. Capacitors used within high-energy capacitor banks can violently explode when a short in one capacitor causes sudden dumping of energy stored in the rest of the bank into the failing unit. High voltage vacuum capacitors can generate soft X-rays even during normal operation. Proper containment, fusing, and preventive maintenance can help to minimize these hazards.

High-voltage capacitors can benefit from a pre-charge to limit in-rush currents at power-up of high voltage direct current (HVDC) circuits. This will extend the life of the component and may mitigate high-voltage hazards.

Voltage dependent capacitors

The dielectric constant for a number of very useful dielectrics changes as a function of the applied electrical field, for example ferroelectric materials, so the capacitance for these devices is more complex. For example, in charging such a capacitor the differential increase in voltage with charge is governed by:

$$dQ = C(V)\, dV \ ,$$

where the voltage dependence of capacitance, $C(V)$, stems from the field, which in a large area parallel plate device is given by $\varepsilon = V/d$. This field polarizes the dielectric, which polarization, in the case of a ferroelectric, is a nonlinear *S*-shaped function of field, which, in the case of a large area parallel plate device, translates into a capacitance that is a nonlinear function of the voltage causing the field.

Corresponding to the voltage-dependent capacitance, to charge the capacitor to voltage V an integral relation is found:

$$Q = \int_0^V dV \ C(V) \ ,$$

which agrees with $Q = CV$ only when C is voltage independent.

By the same token, the energy stored in the capacitor now is given by

$$dW = QdV = \left[\int_0^V \ dV' \ C(V') \right] \ dV \ .$$

Integrating:

$$W = \int_0^V dV \int_0^V dV' \, C(V') = \int_0^V dV' \int_{V'}^V dV \, C(V')$$
$$= \int_0^V dV' \, (V - V') \, C(V') \ ,$$

where interchange of the order of integration is used.

The nonlinear capacitance of a microscope probe scanned along a ferroelectric surface is used to study the domain structure of ferroelectric materials.

Another example of voltage dependent capacitance occurs in semiconductor devices such as semiconductor diodes, where the voltage dependence stems not from a change in dielectric constant but in a voltage dependence of the spacing between the charges on the two sides of the capacitor.

Frequency dependent capacitors

If a capacitor is driven with a time-varying voltage that changes rapidly enough, then the polarization of the dielectric cannot follow the signal. As an example of the origin of this mechanism, the internal microscopic dipoles contributing to the dielectric constant cannot move instantly, and so as frequency of an applied alternating voltage increases, the dipole response is limited and the dielectric constant diminishes. A changing dielectric constant with frequency is referred to as dielectric dispersion, and is governed by dielectric relaxation processes, such as Debye relaxation. Under transient conditions, the displacement field can be expressed as:

$$\boldsymbol{D}(t) = \varepsilon_0 \int_{-\infty}^{t} dt' \ \varepsilon_r(t - t') \boldsymbol{E}(t') \ ,$$

indicating the lag in response by the time dependence of ε_r, calculated in principle from an underlying microscopic analysis, for example, of the dipole behavior in the dielectric. The integral extends over the entire past history up to the present time. A Fourier transform in time then results in:

$$\boldsymbol{D}(\omega) = \varepsilon_0 \varepsilon_r(\omega) \boldsymbol{E}(\omega) \ ,$$

where $\varepsilon_r(\omega)$ is now a complex function, with an imaginary part related to absorption of energy from the field by the medium. The capacitance, being proportional to the dielectric constant, also exhibits this frequency behavior. Fourier transforming Gauss's law with this form for displacement field:

$$
\begin{aligned}
I(\omega) &= j\omega Q(\omega) = j\omega \oint_{\Sigma} \boldsymbol{D}(\boldsymbol{r},\ \omega) \cdot d\boldsymbol{\Sigma} \\
&= [G(\omega) + j\omega C(\omega)]\, V(\omega) = \frac{V(\omega)}{Z(\omega)}\ ,
\end{aligned}
$$

where j is the imaginary unit, $V(\omega)$ is the voltage component at angular frequency ω, $G(\omega)$ is the *real* part of the current, called the *conductance*, and $C(\omega)$ determines the *imaginary* part of the current and is the *capacitance*. $Z(\omega)$ is the complex impedance.

When a parallel-plate capacitor is filled with a dielectric, the measurement of dielectric properties of the medium is based upon the relation:

$$
\varepsilon_r(\omega) = \varepsilon'_r(\omega) - j\varepsilon''_r(\omega) = \frac{1}{j\omega Z(\omega) C_0} = \frac{C(\omega)}{C_0}\ ,
$$

where a single *prime* denotes the real part and a double *prime* the imaginary part, $Z(\omega)$ is the complex impedance with the dielectric present, $C(\omega)$ is the so-called *complex* capacitance with the dielectric present, and C_0 is the capacitance without the dielectric. (Measurement "without the dielectric" in principle means measurement in free space, an unattainable goal inasmuch as even the quantum vacuum is predicted to exhibit nonideal behavior, such as dichroism. For practical purposes, when measurement errors are taken into account, often a measurement in terrestrial vacuum, or simply a calculation of C_0, is sufficiently accurate.)

Using this measurement method, the dielectric constant may exhibit a resonance at certain frequencies corresponding to characteristic response frequencies (excitation energies) of contributors to the dielectric constant. These resonances are the basis for a number of experimental techniques for detecting defects. The *conductance method* measures absorption as a function of frequency. Alternatively, the time response of the capacitance can be used directly, as in *deep-level transient spectroscopy*.

Another example of frequency dependent capacitance occurs with MOS capacitors, where the slow generation of minority carriers means that at high frequencies the capacitance measures only the majority carrier response, while at low frequencies both types of carrier respond.

At optical frequencies, in semiconductors the dielectric constant exhibits structure related to the band structure of the solid. Sophisticated modulation spectroscopy measurement methods based upon modulating the crystal structure by pressure or by other stresses and observing the related changes in absorption or reflection of light have advanced our knowledge of these materials.

Capacitance matrix

The discussion above is limited to the case of two conducting plates, although of arbitrary size and shape. The definition C=Q/V still holds for a single plate given a charge, in which case the field lines produced by that charge terminate as if the plate were at the center of an oppositely charged sphere at infinity.

$C = Q / V$ does not apply when there are more than two charged plates, or when the net charge on the two plates is non-zero. To handle this case, Maxwell introduced his "coefficients of potential". If three plates are given charges Q_1,Q_2,Q_3, then the voltage of plate 1 is given by

$$V_1 = P_{11}Q_1 + P_{12}Q_2 + P_{13}Q_3,$$

and similarly for the other voltages. Maxwell showed that the coefficients of potential are symmetric, so that $P_{12} = P_{21}$, etc. Thus the system can be described by a collection of coefficients known as the "Reciprocal Capacitance Matrix"is used, which is defined as:

$$P_{ij} = \frac{V_i}{Q_j}$$

From this, the mutual capacitance C_m between two objects can be defined by solving for the total charge Q and using $C_m = Q / V$.

$$C_m = \frac{V}{(P_{11} + P_{22}) - (P_{12} + P_{21})}$$

Technically speaking, since no actual device holds perfectly equal and opposite charges on each of the two "plates", it is the mutual capacitance that is reported on capacitors. The collection of coefficients $C_{ij} = Q_i / V_j$ is known as the capacitance matrix and also describes the capacitance of the system.

Self-capacitance

In electrical circuits, the term *capacitance* is usually a shorthand for the *mutual capacitance* between two adjacent conductors, such as the two plates of a capacitor. There also exists a property called *self-capacitance*, which is the amount of electrical charge that must be added to an isolated conductor to raise its *electrical potential* by one volt. The reference point for this potential is a theoretical hollow conducting sphere, of infinite radius, centered on the conductor. Using this method, the self-capacitance of a conducting sphere of radius R is given by:

$$C = 4\pi\varepsilon_0 R$$

Example values of self-capacitance are:

- for the top "plate" of a van de Graaff generator, typically a sphere 20 cm in radius: 20 pF
- the planet Earth: about 709 µF

Elastance

The inverse of capacitance is called elastance. The unit of elastance is the daraf.

Stray capacitance

Any two adjacent conductors can be considered a capacitor, although the capacitance will be small unless the conductors are close together for long. This (often unwanted) effect is termed "stray capacitance". Stray capacitance can allow signals to leak between otherwise isolated circuits (an effect called crosstalk), and it can be a limiting factor for proper functioning of circuits at high frequency.

Stray capacitance is often encountered in amplifier circuits in the form of "feedthrough" capacitance that interconnects the input and output nodes (both defined relative to a common ground). It is often convenient for analytical purposes to replace this capacitance with a combination of one input-to-ground capacitance and one output-to-ground capacitance. (The original configuration — including the input-to-output capacitance — is often referred to as a pi-configuration.) Miller's theorem can be used to effect this replacement. Miller's theorem states that, if the gain ratio of two nodes is 1/K, then an impedance of Z connecting the two nodes can be replaced with a Z/(1-k) impedance between the first node and ground and a KZ/(K-1) impedance between the second node and ground. (Since impedance varies inversely with capacitance, the internode capacitance, C, will be seen to have been replaced by a capacitance of KC from input to ground and a capacitance of (K-1)C/K from output to ground.) When the input-to-output gain is very large, the equivalent input-to-ground impedance is very small while the output-to-ground impedance is essentially equal to the original (input-to-output) impedance.

Capacitance of simple systems

Calculating the capacitance of a system amounts to solving the Laplace equation $\nabla^2\varphi=0$ with a constant potential φ on the surface of the conductors. This is trivial in cases with high symmetry. There is no solution in terms of elementary functions in more complicated cases.

Inductance

Inductance is the property of an electrical circuit measuring the induced electric voltage compared to the rate of change of the electric current in the circuit. This property also is called **self inductance** to discriminate it from **mutual inductance**, describing the voltage induced in one electrical circuit by the rate of change of the electric current in another circuit.

The quantitative definition of the self inductance *L* of an electrical circuit in SI units (webers per ampere, known as henries) is

$$v = L\frac{di}{dt},$$

where *v* denotes the voltage in volts and *i* the current in amperes. This is a linear relation between voltage and current akin to Ohm's law, but with an extra time derivate. The simplest solutions of this equation are a constant current with no voltage or a current changing linearilly in time with a constant voltage.

The term 'inductance' was coined by Oliver Heaviside in February 1886. It is customary to use the symbol *L* for inductance, possibly in honour of the physicist Heinrich Lenz. The SI unit of inductance is the **henry** (H), named after American scientist and magnetic researcher Joseph Henry. 1 H = 1 Wb/A.

Inductance is caused by the magnetic field generated by electric currents according to Ampere's law. To add inductance to a circuit, electronic components called inductors are used, typically consisting of coils of wire to concentrate the magnetic field and to collect the induced voltage. This is analogous to adding capacitance to a circuit by adding capacitors. Capacitance is caused by the electric field generated by electric charge according to Gauss's law.

The generalization to the case of *K* electrical circuits with currents i_m and voltages v_m reads

$$v_m = \sum_{n=1}^{K} L_{m,n}\frac{di_n}{dt}.$$

Inductance here is a symmetric matrix. The diagonal coefficients $L_{m,m}$ are called coefficients of self inductance, the off-diagonal elements are called coefficients of mutual inductance. The coefficients of inductance are constant as long as no magnetizable material with nonlinear characteristics is involved. This is a direct consequence of the linearity of Maxwell's equations in the fields and the current density. The coefficients of inductance become functions of the currents in the nonlinear case.

Derivation from Faraday's law of inductance

The inductance equations above are a consequence of Maxwell's equations. There is a straightforward derivation in the important case of electrical circuits consisting of thin wires.

Consider a system of K wire loops, each with one or several wire turns. The flux linkage of loop m is given by

$$N_m\Phi_m = \sum_{n=1}^{K} L_{m,n} i_n.$$

Here N_m denotes the number of turns in loop m, Φ_m the magnetic flux through this loop, and $L_{m,n}$ are some constants. This equation follows from Ampere's law - magnetic fields and fluxes are linear functions of the currents. By Faraday's law of induction we have

$$v_m = N_m \frac{d\Phi_m}{dt} = \sum_{n=1}^{K} L_{m,n} \frac{di_n}{dt},$$

where v_m denotes the voltage induced in circuit m. This agrees with the definition of inductance above if the coefficients $L_{m,n}$ are identified with the coefficients of inductance. Because the total currents $N_n i_n$ contribute to Φ_m it also follows that $L_{m,n}$ is proportional to the product of turns $N_m N_n$.

Inductance and magnetic field energy

Multiplying the equation for v_m above with $i_m dt$ and summing over m gives the energy transferred to the system in the time interval dt,

$$\sum_{m}^{K} i_m v_m dt = \sum_{m,n=1}^{K} i_m L_{m,n} di_n \stackrel{!}{=} \sum_{n=1}^{K} \frac{\partial W(i)}{\partial i_n} di_n.$$

This must agree with the change of the magnetic field energy W caused by the currents. The integrability condition

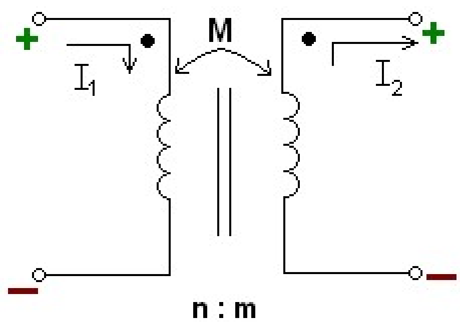
$$\partial^2 W / \partial i_m \partial i_n = \partial^2 W / \partial i_n \partial i_m$$

requires $L_{m,n}=L_{n,m}$. The inductance matrix $L_{m,n}$ thus is symmetric. The integral of the energy transfer is the magnetic field energy as a function of the currents,

$$W(i) = \frac{1}{2} \sum_{m,n=1}^{K} i_m L_{m,n} i_n.$$

This equation also is a direct consequence of the linearity of Maxwell's equations. It is helpful to associate changing electric currents with a build-up or decrease of magnet field energy. The corresponding energy transfer requires or generates a voltage. A mechanical analogy in the K=1 case with magnetic field energy $(1/2)Li^2$ is a body with mass M, velocity u and kinetic energy $(1/2)Mu^2$. The rate of change of velocity (current) multiplied with mass (inductance) requires or generates a force (an electrical voltage).

Coupled inductors



The circuit diagram representation of mutually coupled inductors. The two vertical lines between the inductors indicate a *solid core* that the wires of the inductor are wrapped around. "n:m" shows the ratio between the number of windings of the left inductor to windings of the right inductor. This picture also shows the dot convention.

Mutual inductance occurs when the change in current in one inductor induces a voltage in another nearby inductor. It is important as the mechanism by which transformers work, but it can also cause unwanted coupling between conductors in a circuit.

The mutual inductance, M, is also a measure of the coupling between two inductors. The mutual inductance by circuit i on circuit j is given by the double integral *Neumann formula*.

The mutual inductance also has the relationship:

$$M_{21} = N_1 N_2 P_{21}$$

where

M_{21} is the mutual inductance, and the subscript specifies the relationship of the voltage induced in coil 2 to the current in coil 1.
N_1 is the number of turns in coil 1,
N_2 is the number of turns in coil 2,

P_{21} is the permeance of the space occupied by the flux.

The mutual inductance also has a relationship with the coupling coefficient. The coupling coefficient is always between 1 and 0, and is a convenient way to specify the relationship between a certain orientation of inductor with arbitrary inductance:

$$M = k\sqrt{L_1 L_2}$$

where

k is the *coupling coefficient* and $0 \leq k \leq 1$,
L_1 is the inductance of the first coil, and
L_2 is the inductance of the second coil.

Once the mutual inductance, M, is determined from this factor, it can be used to predict the behavior of a circuit:

$$V_1 = L_1 \frac{dI_1}{dt} - M \frac{dI_2}{dt}$$

where

V_1 is the voltage across the inductor of interest,
L_1 is the inductance of the inductor of interest,
dI_1/dt is the derivative, with respect to time, of the current through the inductor of interest,
dI_2/dt is the derivative, with respect to time, of the current through the inductor that is coupled to the first inductor, and
M is the mutual inductance.

The minus sign arises because of the sense the current I_2 has been defined in the diagram. With both currents defined going into the dots the sign of M will be positive.

When one inductor is closely coupled to another inductor through mutual inductance, such as in a transformer, the voltages, currents, and number of turns can be related in the following way:

$$V_s = \frac{N_s}{N_p} V_p$$

where

V_s is the voltage across the secondary inductor,
V_p is the voltage across the primary inductor (the one connected to a power source),
N_s is the number of turns in the secondary inductor, and

N_p is the number of turns in the primary inductor.

Conversely the current:

$$I_s = \frac{N_p}{N_s} I_p$$

where

I_s is the current through the secondary inductor,
I_p is the current through the primary inductor (the one connected to a power source),
N_s is the number of turns in the secondary inductor, and
N_p is the number of turns in the primary inductor.

Note that the power through one inductor is the same as the power through the other. Also note that these equations don't work if both transformers are forced (with power sources).

When either side of the transformer is a tuned circuit, the amount of mutual inductance between the two windings determines the shape of the frequency response curve. Although no boundaries are defined, this is often referred to as loose-, critical-, and over-coupling. When two tuned circuits are loosely coupled through mutual inductance, the bandwidth will be narrow. As the amount of mutual inductance increases, the bandwidth continues to grow. When the mutual inductance is increased beyond a critical point, the peak in the response curve begins to drop, and the center frequency will be attenuated more strongly than its direct sidebands. This is known as overcoupling.

Calculation techniques

The most general case requires to determine current densities and the magnetic field from Maxwell's equations and Ohm's law. There are, however, simplifications.

- High frequency with complete skin effect. Surface current densities and magnetic field may be obtained by solving the Laplace equation.
- Conductors are thin wires. Self inductance still depends on the wire radius and the distribution of the current in the wire, but this current distribution is more or less constant (on the surface or in the volume of the wire) for a wire radius much smaller than other length scales.

Mutual inductance

The mutual inductance by a filamentary circuit *i* on a filamentary circuit *j* is given by the double integral *Neumann formula*

$$M_{ij} = \frac{\mu_0}{4\pi} \oint_{C_i} \oint_{C_j} \frac{\mathbf{ds}_i \cdot \mathbf{ds}_j}{|\mathbf{R}_{ij}|}$$

The symbol μ_0 denotes the magnetic constant ($4\pi\times10^{-7}$ H/m), 'C_i *and* C_j *are the curves spanned by the wires,* R_{ij} *is the distance between two points.*

Self-inductance

Formally the self-inductance of a wire loop would be given by the above equation with $i = j$. However, *1/R* now becomes infinite and thus the finite radius *a* along with the distribution of the current in the wire must be taken into account. There remain the contribution from the integral over all points where $|R| \geq a/2$ and a correction term,

$$M_{ii} = L \approx \left(\frac{\mu_0}{4\pi} \oint_{C} \oint_{C'} \frac{\mathbf{ds} \cdot \mathbf{ds}'}{|\mathbf{R}|}\right)_{|\mathbf{R}| \geq a/2} + \frac{\mu_0}{2\pi} l Y$$

Here *a* and *l* denote radius and length of the wire, and *Y* is a constant that depends on the distribution of the current in the wire: $Y = 0$ when the current flows in the surface of the wire (skin effect), $Y = 1/4$ when the current is homogeneous across the wire. This approximation is accurate when the wires are long compared to their cross-sectional dimensions. Here is a derivation of this equation.

Method of images

In some cases different current distributions generate the same magnetic field in some section of space. This fact may be used to relate self inductances (method of images). As an example consider the two systems:

- A wire at distance *d/2* in front of a perfectly conducting wall (which is the return)
- Two parallel wires at distance *d*, with opposite current

The magnetic field of the two systems coincides (in a half space). The magnetic field energy and the inductance of the second system thus are twice as large as that of the first system.

Relation between inductance and capacitance

Inductance per length L' and capacitance per length C' are related to each other in the special case of transmission lines consisting of two parallel perfect conductors of arbitrary but constant cross section,

$$L'C' = \varepsilon\mu.$$

Here ε and μ denote dielectric constant and magnetic permeability of the medium the conductors are embedded in. There is no electric and no magnetic field inside the conductors (complete skin effect, high frequency). Current flows down on one line and returns on the other. The signal propagation speed coincides with the propagation speed of electromagnetic waves in the bulk.

Self-inductance of simple electrical circuits in air

The symbol μ_0 denotes the magnetic constant ($4<\pi\times10^{-7}$ H/m). For high frequencies the electric current flows in the conductor surface (skin effect), and depending on the geometry it sometimes is necessary to distinguish low and high frequency inductances. This is the purpose of the constant Y: $Y = 0$ when the current is uniformly distributed over the surface of the wire (skin effect), $Y = 1/4$ when the current is uniformly distributed over the cross section of the wire. In the high frequency case, if conductors approach each other, an additional screening current flows in their surface, and expressions containing Y become invalid. Details for some circuit types are available on another page.

Phasor circuit analysis and impedance

Using phasors, the equivalent impedance of an inductance is given by:

$$Z_L = V/I = jL\omega$$

where

j is the imaginary unit,
L is the inductance,
$\omega = 2\pi f$ is the angular frequency,
f is the frequency and
$L\omega = X_L$ is the inductive reactance.

Nonlinear inductance

Many inductors make use of magnetic materials. These materials over a large enough range exhibit a nonlinear permeability with such effects as saturation. This in-turn makes the resulting inductance a function of the applied current. Faraday's Law still holds but inductance is ambiguous and is different whether you are calculating circuit parameters or magnetic fluxes.

The secant or large-signal inductance is used in flux calculations. It is defined as:

$$L_s(i) \stackrel{\text{def}}{=} \frac{N\Phi}{i} = \frac{\Lambda}{i}$$

The differential or small-signal inductance, on the other hand, is used in calculating voltage. It is defined as:

$$L_d(i) \stackrel{\text{def}}{=} \frac{d(N\Phi)}{di} = \frac{d\Lambda}{di}$$

The circuit voltage for a nonlinear inductor is obtained via the differential inductance as shown by Faraday's Law and the chain rule of calculus.

$$v(t) = \frac{d\Lambda}{dt} = \frac{d\Lambda}{di}\frac{di}{dt} = L_d(i)\frac{di}{dt}$$

There are similar definitions for nonlinear mutual inductances.

Phase relationship

The phase of the voltage across a purely reactive device (a device with a resistance of zero) *lags* the current by $\pi/2$radians for a capacitive reactance and *leads* the current by $\pi/2$radians for an inductive reactance. Note that without knowledge of both the resistance and reactance the relationship between voltage and current cannot be determined.

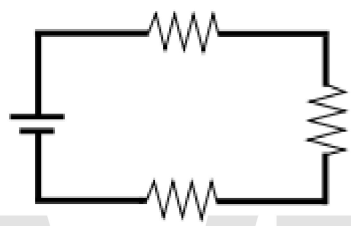
The origin of the different signs for capacitive and inductive reactance is the phase factor in the impedance.

$$\tilde{Z}_C = \frac{1}{\omega C}e^{j(-\frac{\pi}{2})} = j\left(\frac{-1}{\omega C}\right) = jX_C$$
$$\tilde{Z}_L = \omega L e^{j\frac{\pi}{2}} = j\omega L = jX_L$$

For a reactive component the sinusoidal voltage across the component is in quadrature (a $\pi/2$phase difference) with the sinusoidal current through the component. The component alternately absorbs energy from the circuit and then returns energy to the circuit, thus a pure reactance does not dissipate power.

Chapter- 6

Series and Parallel Circuits



A series circuit with a voltage source (such as a battery) and 3 resistors

Components of an electrical circuit or electronic circuit can be connected in many different ways. The two simplest of these are called **series** and **parallel** and occur very frequently. Components connected in series are connected along a single path, so the same current flows through all of the components. Components connected in parallel are connected so the same voltage is applied to each component.

A circuit composed solely of components connected in series is known as a **series circuit**; likewise, one connected completely in parallel is known as a **parallel circuit**.

In a series circuit, the current through each of the components is the same, and the voltage across the components is the sum of the voltages across each component. In a parallel circuit, the voltage across each of the components is the same, and the total current is the sum of the currents through each component.

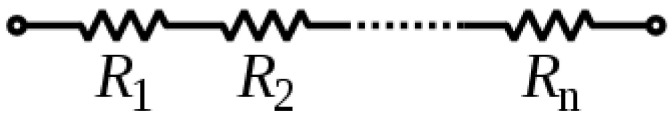
As an example, consider a very simple circuit consisting of four light bulbs and one 6 V battery. If a wire joins the battery to one bulb, to the next bulb, to the next bulb, to the next bulb, then back to the battery, in one continuous loop, the bulbs are said to be in series. If each bulb is wired to the battery in a separate loop, the bulbs are said to be in parallel. If the four light bulbs are connected in series, the same current flows through all of them, and the voltage drop is 1.5 V across each bulb and that may not be sufficient to make them glow. If the light bulbs are connected in parallel, the current flowing through the light bulbs combine to form the current flowing in the battery, while the voltage drop is 6.0 V across each bulb and they all glow.

In a series circuit, every device must function for the circuit to be complete. One bulb burning out in a series circuit breaks the circuit. In parallel circuits, each light has its own circuit, so all but one light could be burned out, and the last one will still function.

Series circuits

Series circuits are sometimes called *current*-coupled or daisy chain-coupled. The current that flows in a series circuit will flow through every component in the circuit. Therefore, all of the components in a series connection carry the same current.

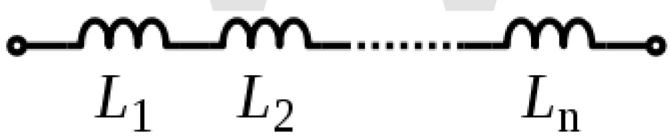
Resistors



$$R_{\text{total}} = R_1 + R_2 + \cdots + R_n$$

Inductors

Inductors follow the same law, in that the total inductance of non-coupled inductors in series is equal to the sum of their individual inductances:



$$L_{\text{total}} = L_1 + L_2 + \cdots + L_n$$

However, in some situations it is difficult to prevent adjacent inductors from influencing each other, as the magnetic field of one device couples with the windings of its neighbours. This influence is defined by the mutual inductance M. For example, if you have two inductors in series, there are two possible equivalent inductances depending on how the magnetic fields of both inductors influence each other.

When there are more than 2 inductors, the mutual inductance between each of them and the way the coils influence each other complicates the calculation. For a larger number of coils the total combined inductance is given by the sum of all mutual inductances between the various coils including the mutual inductance of each given coil with itself, which we term self-inductance or simply inductance. For three coils, there are six mutual inductances M_{12}, M_{13}, M_{23} and M_{21}, M_{31} and M_{32}. There are also the three self-inductances of the three coils: M_{11}, M_{22} and M_{33}.

Therefore

$$L_{\text{total}} = (M_{11} + M_{22} + M_{33}) + (M_{12} + M_{13} + M_{23}) + (M_{21} + M_{31} + M_{32})$$

By reciprocity $M_{ij} = M_{ji}$ so that the last two groups can be combined. The first three terms represent the sum of the self-inductances of the various coils. The formula is easily extended to any number of series coils with mutual coupling. The method can be used to find the self-inductance of large coils of wire of any cross-sectional shape by computing the sum of the mutual inductance of each turn of wire in the coil with every other turn since in such a coil all turns are in series.

Capacitors

Capacitors follow the same law using the reciprocals. The total capacitance of capacitors in series is equal to the reciprocal of the sum of the reciprocals of their individual capacitances:

$$\frac{1}{C_{\text{total}}} = \frac{1}{C_1} + \frac{1}{C_2} + \cdots + \frac{1}{C_n}.$$

The working voltage of a series combination of identical capacitors is equal to the sum of voltage ratings of individual capacitors. This simple relationship only applies if the voltage ratings are equal as well as the capacitances. However, the division of DC voltage between the capacitors is dominated by the leakage resistance of the capacitors, rather than their capacitances, and this has considerable variation. To counter this equalising resistors may be placed in parallel with each capacitor which effectively add to the leakage current. The value of resistor chosen (perhaps a few megohms) is as large as possible, but low enough to ensure that the capacitor leakage current is insignificant compared to the current through the resistor. At DC, the circuit appears as a chain of series identical resistors and equal voltage division between the capacitors is ensured. In high-voltage circuits, the resistors serve an additional function as bleeder resistors.

Switches

Two or more switches in series form a logical AND; the circuit only carries current if all switches are 'on'.

Cells and batteries

A battery is a collection of electrochemical cells. If the cells are connected in series, the voltage of the battery will be the sum of the cell voltages. For example, a 12 volt car battery contains six 2-volt cells connected in series.

Parallel circuits

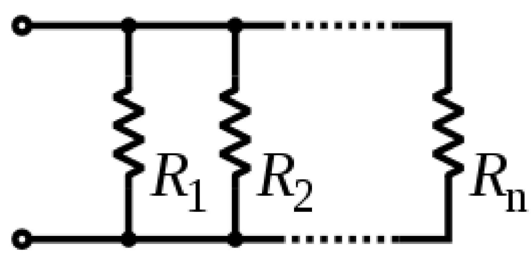
If two or more components are connected in parallel they have the same potential difference (voltage) across their ends. The potential differences across the components are the same in magnitude, and they also have identical polarities. The same voltage is applicable to all circuit components connected in parallel. The total current *I* is the sum of the currents through the individual components, in accordance with Kirchhoff's current law.

Resistors

The current in each individual resistor is found by Ohm's law. Factoring out the voltage gives

$$I_{\text{total}} = V\left(\frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}\right).$$

To find the total resistance of all components, add the reciprocals of the resistances R_i of each component and take the reciprocal of the sum. Total resistance will always be less than the value of the smallest resistance:



$$\frac{1}{R_{\text{total}}} = \frac{1}{R_1} + \frac{1}{R_2} + \cdots + \frac{1}{R_n}.$$

For only two resistors, the unreciprocated expression is reasonably simple:

$$R_{\text{total}} = \frac{R_1 R_2}{R_1 + R_2}.$$

This sometimes goes by the mnemonic "product over sum".

For *N* equal resistors in parallel, the reciprocal sum expression simplifies to:

$$\frac{1}{R_{\text{total}}} = \frac{1}{R} \times N.$$

and therefore to:

$$R_{\text{total}} = \frac{R}{N}.$$

To find the current in a component with resistance R_i, use Ohm's law again:

$$I_i = \frac{V}{R_i}.$$

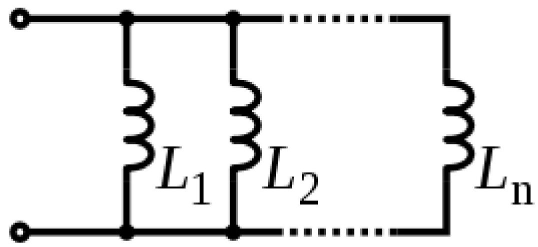
The components divide the current according to their reciprocal resistances, so, in the case of two resistors,

$$\frac{I_1}{I_2} = \frac{R_2}{R_1}.$$

An old term for devices connected in parallel is *multiple*, such as a multiple connection for arc lamps.

Inductors

Inductors follow the same law, in that the total inductance of non-coupled inductors in parallel is equal to the reciprocal of the sum of the reciprocals of their individual inductances:



$$\frac{1}{L_\mathrm{total}} = \frac{1}{L_1} + \frac{1}{L_2} + \cdots + \frac{1}{L_n}.$$

If the inductors are situated in each other's magnetic fields, this approach is invalid due to mutual inductance. If the mutual inductance between two coils in parallel is M, the equivalent inductor is:

$$\frac{1}{L_\mathrm{total}} = \frac{L_1 + L_2 - 2M}{L_1 L_2 - M^2}$$

If $L_1 = L_2$

$$L_{total} = \frac{L + M}{2}$$

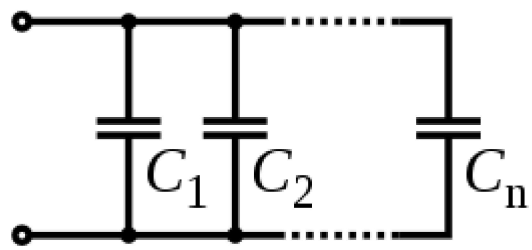
The sign of M depends on how the magnetic fields influence each other. For two equal tightly coupled coils the total inductance is close to that of each single coil. If the polarity of one coil is reversed so that M is negative, then the parallel inductance is nearly zero or the combination is almost non-inductive. We are assuming in the "tightly coupled" case M is very nearly equal to L. However, if the inductances are not equal and the coils are tightly coupled there can be near short circuit conditions and high circulating currents for both positive and negative values of M, which can cause problems.

More than 3 inductors becomes more complex and the mutual inductance of each inductor on each other inductor and their influence on each other must be considered. For three coils, there are three mutual inductances M_{12}, M_{13} and M_{23}. This is best handled by matrix methods and summing the terms of the inverse of the L matrix (3 by 3 in this case).

The pertinent equations are of the form: $$v_i = \sum_j L_{i,j} \frac{di_j}{dt}$$

Capacitors

Capacitors follow the same law using the reciprocals. The total capacitance of capacitors in parallel is equal to the sum of their individual capacitances:



$$C_{\text{total}} = C_1 + C_2 + \cdots + C_n.$$

The working voltage of a parallel combination of capacitors is always limited by the smallest working voltage of an individual capacitor.

Switches

Two or more switches in parallel, form a logical OR; the circuit carries current if at least one switch is 'on'.

Cells and batteries

If the cells of a battery are connected in parallel, the battery voltage will be the same as the cell voltage but the current supplied by each cell will be a fraction of the total current. For example, if a battery contains four cells connected in parallel and delivers a current of 1 ampere, the current supplied by each cell will be 0.25 ampere. Parallel-connected batteries were widely used to power the valve filaments in portable radios but they are now rare.

Applications

Series circuits were formerly used for lighting in electric multiple unit trains. For example, if the supply voltage was 600 volts there might be eight 70 volt bulbs in series (total 560 volts) plus a resistor to drop the remaining 40 volts. Series circuits for train lighting were superseded, first by motor-generators, then by solid state devices.

Series resistance can also be applied to the arrangement of blood vessels within a given organ. Each organ is supplied by a large artery, smaller arteries, arterioles, capillaries, and veins arranged in series. The total resistance is the sum of the individual resistances, as expressed by the following equation: $R_{total} = R_{artery} + R_{arterioles} + R_{capillaries}$. The largest proportion of resistance in this series is contributed by the arterioles.

Parallel resistance is illustrated by the system circulation. Each organ is supplied by an artery that branches off the aorta. The total resistance of this parallel arrangement is expressed by the following equation: $1/R_{total} = 1/R_a + 1/R_b + ... 1/R_n$. R_a, R_b, and R_n are the resistances of the renal, hepatic, and other arteries respectively. The total resistance is less than the resistance of any of the individual arteries.

Chapter- 7

Buffer Amplifier & Impedance Matching

Buffer amplifier

A **buffer amplifier** (sometimes simply called a **buffer**) is one that provides electrical impedance transformation from one circuit to another. Two main types of buffer exist: the **voltage buffer** and the **current buffer**.

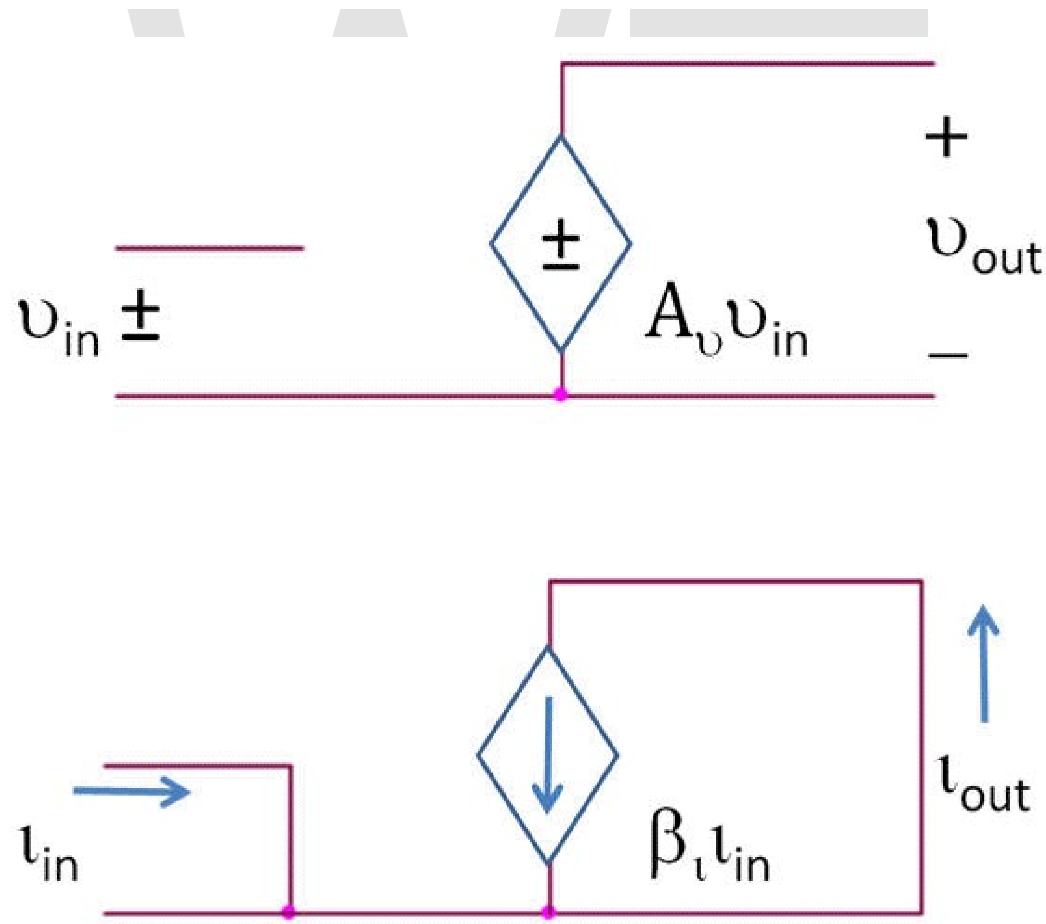


Figure 1: Top: Ideal voltage buffer Bottom: Ideal current buffer

Voltage buffer

A voltage buffer amplifier is used to transfer a voltage from a first circuit, having a high output impedance level, to a second circuit with a low input impedance level. The interposed buffer amplifier prevents the second circuit from loading the first circuit unacceptably and interfering with its desired operation. In the ideal voltage buffer in the diagram, the input resistance is infinite, the output resistance zero (impedance of an ideal voltage source is zero). Other properties of the ideal buffer are: perfect linearity, regardless of signal amplitudes; and instant output response, regardless of the speed of the input signal.

If the voltage is transferred unchanged (the voltage gain A_v is 1), the amplifier is a **unity gain buffer**; also known as a **voltage follower** because the output voltage *follows* or tracks the input voltage. Although the voltage gain of a voltage buffer amplifier may be (approximately) unity, it usually provides considerable current gain and thus power gain. However, it is commonplace to say that it has a gain of 1 (or the equivalent 0 dB), referring to the voltage gain.

As an example, consider a Thévenin source (voltage V_A, series resistance R_A) driving a resistor load R_L. Because of voltage division (also referred to as "loading") the voltage across the load is only $V_A R_L / (R_L + R_A)$. However, if the Thévenin source drives a unity gain buffer such as that in Figure 1 (top, with unity gain), the voltage input to the amplifier is V_A, with *no voltage division* because the amplifier input resistance is infinite. At the output the dependent voltage source delivers voltage $A_v V_A = V_A$ to the load, again without voltage division because the output resistance of the buffer is zero. A Thévenin equivalent circuit of the combined original Thévenin source *and* the buffer is an ideal voltage source V_A with zero Thévenin resistance.

Current buffer

Typically a current buffer amplifier is used to transfer a current from a first circuit, having a low output impedance level, to a second circuit with a high input impedance level. The interposed buffer amplifier prevents the second circuit from loading the first circuit unacceptably and interfering with its desired operation. In the ideal current buffer in the diagram, the input resistance is zero, the output resistance infinite (impedance of an ideal current source is infinite). Again, other properties of the ideal buffer are: perfect linearity, regardless of signal amplitudes; and instant output response, regardless of the speed of the input signal.

For a current buffer, if the current is transferred unchanged (the current gain β_i is 1), the amplifier is again a **unity gain buffer**; this time known as a **current follower** because the output current *follows* or tracks the input current.

As an example, consider a Norton source (current I_A, parallel resistance R_A) driving a resistor load R_L. Because of current division (also referred to as "loading") the current delivered to the load is only $I_A R_A / (R_L + R_A)$. However, if the Norton source drives a unity gain buffer such as that in Figure 1 (bottom, with unity gain), the current input to the amplifier is I_A, with *no current division* because the amplifier input resistance is zero. At the output the dependent current source delivers current $\beta_i I_A = I_A$ to the load, again without current division because the output resistance of the buffer is infinite. A Norton equivalent circuit of the combined original Norton source *and* the buffer is an ideal current source I_A with infinite Norton resistance.

Voltage buffer examples

Op-amp implementation

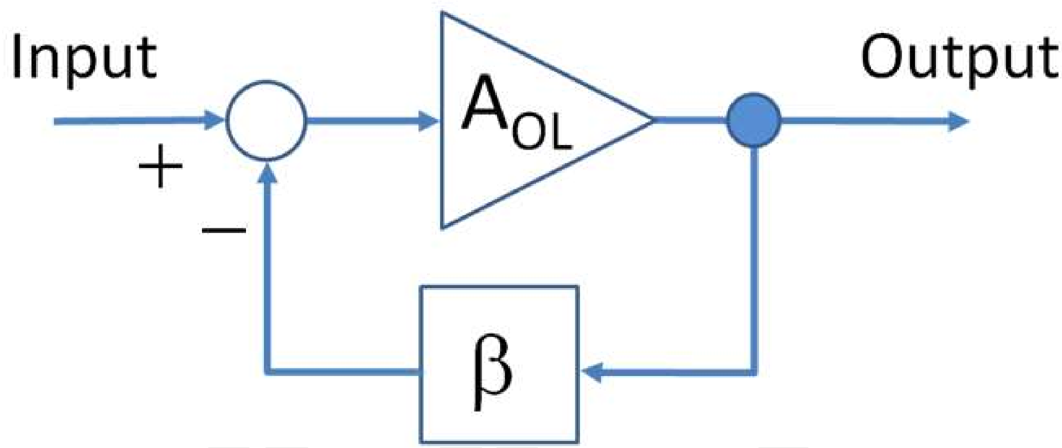


Figure 2: A negative feedback amplifier

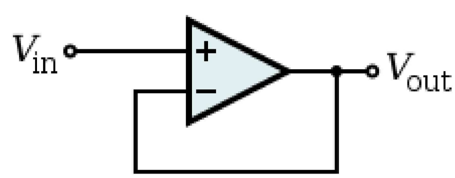


Figure 3. An op-amp–based unity gain buffer amplifier

A unity gain buffer amplifier may be constructed by applying a full series negative feedback (Fig. 2) to an op-amp simply by connecting its output to its inverting input, and connecting the signal source to the non-inverting input (Fig. 3). In this configuration, the entire output voltage (β = 1 in Fig. 2) is placed contrary and in series with the input voltage. Thus the two voltages are subtracted according to KVL and their difference is applied to the op-amp differential input. This connection forces the op-amp to adjust its output voltage simply equal to the input voltage (V_{out} follows V_{in} so the circuit is named op-amp voltage follower).

The importance of this circuit does not come from any change in voltage, but from the input and output impedances of the op-amp. The input impedance of the op-amp is very high (1 MΩ to 10 TΩ), meaning that the input of the op-amp does not load down the source or draw any current from it. Because the output impedance of the op-amp is very low, it drives the load as if it were a perfect voltage source. Both the connections to and from the buffer are therefore bridging connections, which reduce power consumption in the source, distortion from overloading, crosstalk and other electromagnetic interference.

Single-transistor circuits

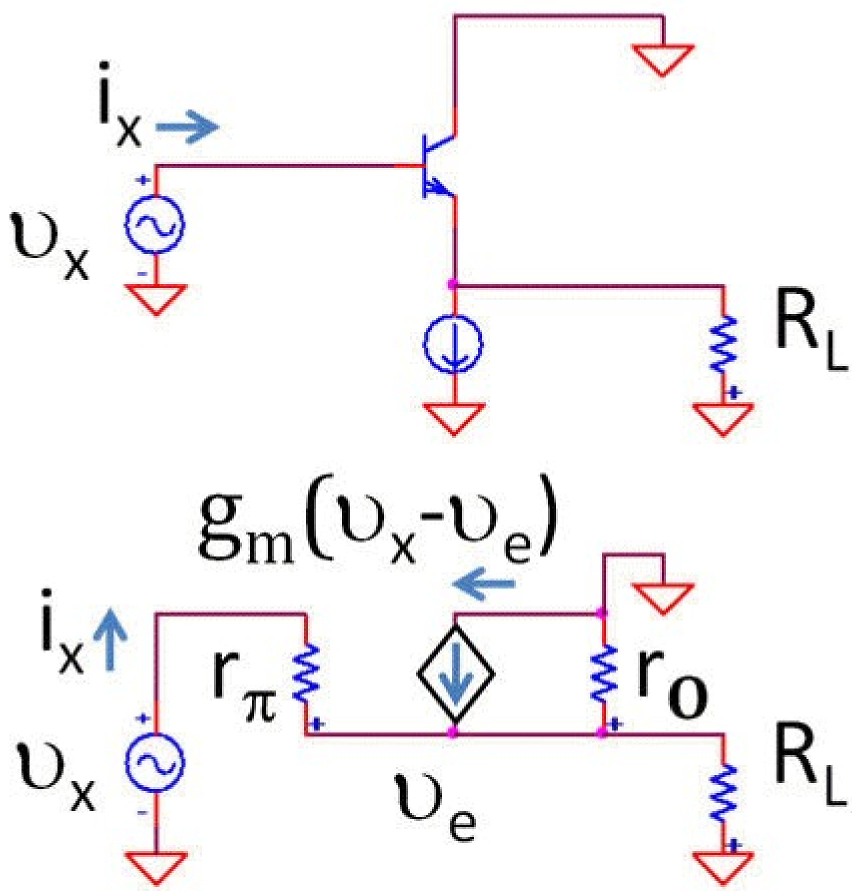


Figure 4: Top: BJT voltage follower Bottom: Small-signal, low-frequency equivalent circuit using hybrid-pi model

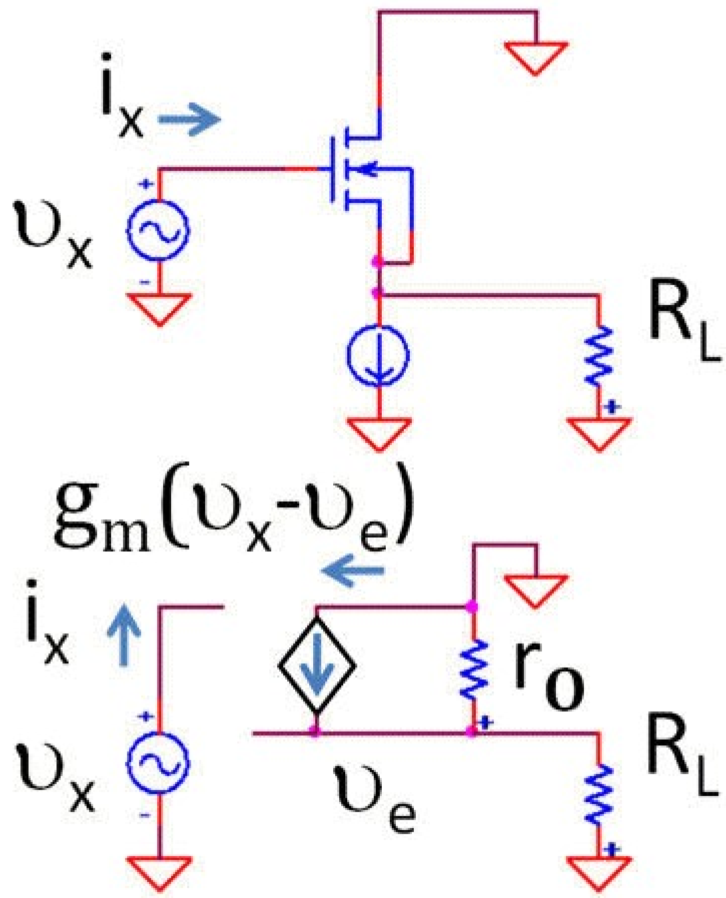


Figure 5: Top: MOSFET voltage follower Bottom: Small-signal, low-frequency equivalent circuit using hybrid-pi model

Other unity gain buffer amplifiers include the bipolar junction transistor in common-collector configuration (called an *emitter follower* because the emitter voltage follows the base voltage, or a *voltage follower* because the output voltage follows the input voltage); the field effect transistor in common-drain configuration (called a source follower because the source voltage follows the gate voltage or, again, a *voltage follower* because the output voltage follows the input voltage); or similar configurations using vacuum tubes (cathode follower), or other active devices. All such amplifiers actually have a gain of slightly less than unity, but the difference is usually small and unimportant.

Impedance transformation using the bipolar voltage follower

Using the small-signal circuit in Figure 4, the impedance seen looking into the circuit is

$$R_{in} = \frac{v_x}{i_x} = r_\pi + (\beta + 1)(r_O // R_L)$$
.

(The analysis uses the relation $g_m r_\pi = (I_C / V_T)\ (V_T / I_B) = \beta$, which follows from the evaluation of these parameters in terms of the bias currents.) Assuming the usual case where $r_O >> R_L$, the impedance is looking into the buffer is larger than the load R_L without the buffer by a factor of (β +1), which is substantial because β is large. The impedance is increased even more by the added r_π, but often $r_\pi << (\beta +1)\ R_L$, so the addition does not make much difference.

Impedance transformation using the MOSFET voltage follower

Using the small-signal circuit in Figure 5, the impedance seen looking into the circuit is no longer R_L but instead is infinite (at low frequencies) because the MOSFET draws no current.

As frequency is increased, the parasitic capacitances of the transistors come into play and the transformed input impedance drops with frequency.

Integrated buffer amplifiers

It is common for a single package to contain several discrete buffer amplifiers. For example, a **hex buffer** is a single package containing 6 discrete buffer amplifiers, and an **octal buffer** is a single package containing 8 discrete buffer amplifiers.

Speaker array amplifiers

The majority of amplifiers used to drive large speaker arrays, such as those used for rock concerts, are unity-gain, high-current amplifiers. Some current amplifiers take the voltage output from Class A/B, B, or tube (valve) amplifiers, while others contain built-in voltage amplifiers as a pre-amp stage. The result is a signal nearly identical to the input signal in terms of voltage, but capable of sending high amounts of current into low impedance speaker arrays where the speakers are wired in parallel.

Current buffer examples

Simple unity gain buffer amplifiers include the bipolar junction transistor in common-base configuration, or the MOSFET in common-gate configuration (called a *current follower* because the output current follows the input current). The current gain of a current buffer amplifier is (approximately) unity.

Single-transistor circuits

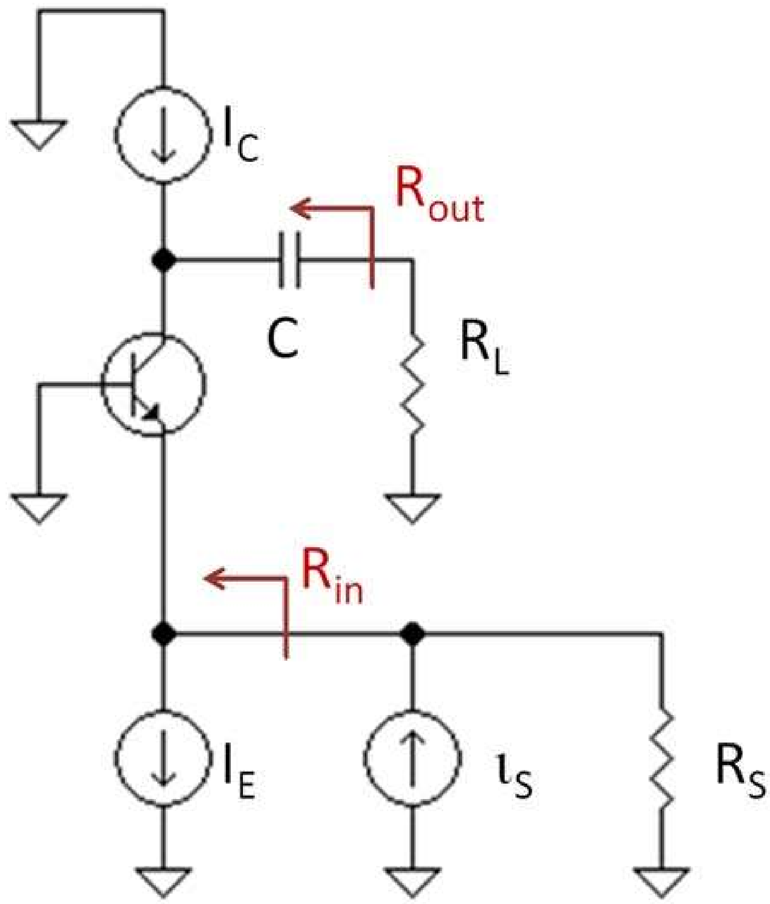


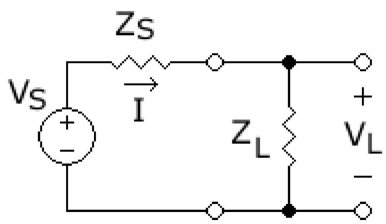
Figure 6: Bipolar current follower biased by current source I_E and with active load I_C

Figure 6 shows a bipolar current buffer biased with a current source (designated I_E for DC emitter current) and driving another DC current source as active load (designated I_C for DC collector current). The AC input signal current i_{in} is applied to the emitter node of the transistor by an AC Norton current source with Norton resistance R_S. The AC output current i_{out} is delivered by the buffer via a large coupling capacitor to load R_L. This coupling capacitor is large enough to be a short-circuit at frequencies of interest.

Because the transistor output resistance connects input and output sides of the circuit, there is a (very small) backward voltage feedback from the output to the input so this circuit is not unilateral. In addition, for the same reason, the input resistance depends (slightly) upon the output load resistance, and the output resistance depends significantly on the input driver resistance.

Impedance matching

In electronics, **impedance matching** is the practice of designing the input impedance of an electrical load or the output impedance of its corresponding signal source in order to maximize the power transfer and minimize reflections from the load.



In the case of a complex source impedance Z_S and load impedance Z_L, matching is obtained when

$$Z_S = Z_L^*$$

where * indicates the complex conjugate.

The concept of impedance matching was originally developed for electrical power, but can be applied to any other field where a form of energy (not necessarily electrical) is transferred between a source and a load.

An alternative to impedance matching is impedance bridging, where the load impedance is chosen to be much larger than the source impedance and maximizing voltage transfer, rather than power, is the goal.

Explanation

The term impedance is used for the resistance of a system to an energy source. For constant signals, this resistance can also be constant. For varying signals, it usually changes with frequency. The energy involved can be electrical, mechanical, magnetic or even thermal. The concept of electrical impedance is perhaps the most commonly known. Electrical impedance, like electrical resistance, is measured in ohms. In general, impedance has a complex value, which means that loads generally have a resistance to the source that is in phase with a sinusoidal source signal **and** reactance that is out of phase with a sinusoidal source signal. The total impedance (symbol: Z) is the vector sum of the resistance (symbol: R; a real number) and the reactance (symbol: X; an imaginary number).

In simple cases, such as low-frequency or direct-current power transmission, the reactance is negligible or zero and the impedance can be considered a pure resistance, expressed as a real number. In the following summary, we will consider the general case when the resistance and reactance are both significant, and also the special case in which the reactance is negligible.

Reflectionless or broadband matching

Impedance matching to minimize reflections and maximize power transfer over a (relatively) large bandwidth (also called **reflectionless matching** or **broadband matching**) is the most commonly used. To prevent all reflections of the signal back into the source, the load (which must be totally resistive) must be matched exactly to the source impedance (which again must be totally resistive). In this case, if a transmission line is used to connect the source and load together, it must also be the same impedance: $Z_{\text{load}} = Z_{\text{line}} = Z_{\text{source}}$, where Z_{line} is the characteristic impedance of the transmission line. Although source and load should each be totally resistive for this form of matching to work, the more general term 'impedance' is still used to describe the source and load characteristics. Any and all reactance actually present in the source or the load will affect the 'match'.

Complex conjugate matching

This is used in cases in which the source and load are reactive. This form of impedance matching can only maximize the power transfer between a reactive source and a reactive load at a *single* frequency. In this case,

$$Z_{\text{load}} = Z_{\text{source}}^{*}$$

(where * indicates the complex conjugate).

If the signals are kept within the narrow frequency range for which the matching network was designed, reflections (in this narrow frequency band only) are also minimized. For the case of purely resistive source and load impedances, all reactance terms are zero and the formula above reduces to

$$Z_{\text{load}} = Z_{\text{source}}$$

as would be expected.

Power transfer

Whenever a source of power *with a fixed output impedance*, such as an electric signal source, a radio transmitter, or even mechanical sound (e.g., a loudspeaker) operates into a load, the maximum possible power is delivered to the load when the impedance of the load (load impedance or input impedance) is equal to the *complex conjugate* of the impedance of the source (that is, its internal impedance or output impedance). For two impedances to be complex conjugates, their resistances must be equal, and their reactances must be equal in magnitude but of opposite sign.

In low-frequency or DC systems, or in systems with purely resistive sources and loads, the reactances are zero, or small enough to be ignored. In this case, maximum power transfer occurs when the resistance of the load is equal to the resistance of the source.

Impedance matching is not always desirable. For example, if a source with a low impedance is connected to a load with a high impedance, then the power that can pass through the connection is limited by the higher impedance, but the electrical voltage transfer is higher and less prone to corruption than if the impedances had been matched. This maximum voltage connection is a common configuration called **impedance bridging** or **voltage bridging** and is widely used in signal processing. In such applications, delivering a high voltage (to minimize signal degradation during transmission and/or to consume less power by reducing currents) is often more important than maximum power transfer.

In older audio systems, reliant on transformers and passive filter networks, and based on the telephone system, the source and load resistances were matched at 600 ohms. One reason for this was to maximize power transfer, as there were no amplifiers available that could restore lost signal. Another reason was to ensure correct operation of the hybrid transformers used at central exchange equipment to separate outgoing from incoming speech so that these could be amplified or fed to a four-wire circuit. Most modern audio circuits, on the other hand, use active amplification and filtering, and they can use voltage bridging connections for best accuracy.

Strictly speaking, impedance matching only applies when both source and load devices are linear, however useful matching may be obtained between nonlinear devices with certain operating ranges.

Impedance matching devices

Adjusting the source impedance or the load impedance, in general, is called "impedance matching".

There are three possible ways to improve an impedance mismatch, all of which are called "impedance matching":

- devices intended to present an apparent load to the source of $R_{\text{load}} = R_{\text{source}}^*$ (complex conjugate matching). Given a source with a fixed voltage and fixed source impedance, the maximum power theorem says this is the only way to extract the maximum power from the source.
- devices intended to present an apparent load of $R_{\text{load}} = R_{\text{line}}$ (complex impedance matching), to avoid echoes. Given a transmission line source with a fixed source impedance, this "reflectionless impedance matching" at the end of the transmission line is the only way to avoid reflecting echoes back to the transmission line.
- devices intended to present an apparent source resistance as close to zero as possible, or presenting an apparent source voltage as high as possible. This is the

only way to maximize energy efficiency, and so it is used at the beginning of electrical power lines. Such an impedance bridging connection also minimizes distortion and electromagnetic interference, and so it is also used in modern audio amplifiers and signal processing devices.

There are a variety of devices that are used between some source of energy and some load that perform "impedance matching".

To match electrical impedances, engineers use combinations of transformers, resistors, inductors, capacitors and transmission lines.

These passive and active impedance matching devices are optimized for different applications, and are called baluns, antenna tuners (sometimes called ATUs or roller coasters because of their appearance), acoustic horns, matching networks, and terminators.

Transformers

Transformers are sometimes used to match the impedances of circuits with different impedances. A transformer converts alternating current at one voltage to the same waveform at another voltage. The power input to the transformer and output from the transformer is the same (except for conversion losses). The side with the lower voltage is at low impedance, because this has the lower number of turns, and the side with the higher voltage is at a higher impedance as it has more turns in its coil.

Resistive network

Resistive impedance matches are easiest to design. They limit the power deliberately, and are used to transfer low-power signals, such as unamplified audio or radio frequency signals in a radio receiver. Almost all digital circuits use resistive impedance matching which is usually built into the structure of the switching element.

Stepped transmission line

Most lumped element devices can match a specific range of load impedance. For example, in order to match an inductive load into a real impedance, a capacitor needs be used. And if the load impedance becomes capacitive for some reason, the matching element must be replaced by an inductor. In many practical cases however, there is a need to use the same circuit to match a broad range of load impedance, thus simplify the circuit design. This issue was addressed by the stepped transmission line where multiple serially placed quarter wave dielectric slugs are used to vary transmission line's characteristic impedance. By controlling the position of each individual element, a broad range of load impedance can be matched without having to reconnect the circuit.

Some special situations - such as radio tuners and transmitters - use tuned filters such as stubs to match impedances at specific frequencies. These can distribute different frequencies to different places in the circuit.

In addition, there is the closely related idea of

- power factor correction devices intended to cancel out the reactive and nonlinear characteristics of a load at the end of a power line. This causes the load seen by the power line to be purely resistive. For a given true power required by a load, this minimizes the true current supplied through the power lines, and so minimizes the power wasted in the resistance of those power lines.

For example, a maximum power point tracker is used to extract the maximum power from a solar panel, and efficiently transfer it to batteries, the power grid, or other loads. The maximum power theorem applies to its "upstream" connection to the solar panel, so it emulates a load resistance equal to the solar panel source resistance. However, the maximum power theorem does not apply to its "downstream" connection, so that connection is an impedance bridging connection—it emulates a high-voltage, low-resistance source, to maximize efficiency.

L-section

One simple electrical impedance matching network requires one capacitor and one inductor. One reactance is in parallel with the source (or load) and the other is in series with the load (or source). If a reactance is in parallel *with the source*, the effective network matches from high impedance to low impedance. The L-section is inherently a narrowband matching network.

The analysis is as follows. Consider a real source impedance of R_1 and real load impedance of R_2. If a reactance X_1 is in parallel with the source impedance, the combined impedance can be written as:

$$\frac{jR_1X_1}{R_1 + jX_1}$$

If the imaginary part of the above impedance is completely canceled by the series reactance, the real part is

$$R_2 = \frac{R_1X_1^2}{R_1^2 + X_1^2}$$

Solving for X_1

$$X_1 = \sqrt{\frac{R_2 R_1^2}{R_1 - R_2}}$$

If $R_1 \gg R_2$ the above equation can be approximated as

$$X_1 \approx \sqrt{R_1 R_2}$$

The inverse connection, impedance step up, is simply the reverse, e.g. reactance in series with the source. The magnitude of the impedance ratio is limited by reactance losses such as the Q of the inductor. Multiple L-sections can be wired in cascade to achieve higher impedance ratios or greater bandwidth. Transmission line matching networks can be modeled as infinitely many L-sections wired in cascade. Optimal matching circuits can be designed for a particular system with the use of the Smith chart.

Transmission lines

Impedance bridging is unsuitable for RF connections because it causes power to be reflected back to the source from the boundary between the high impedance and the low impedance. The reflection creates a standing wave, which leads to further power waste. In these systems, impedance matching is essential.

In electrical systems involving transmission lines, such as radio and fiber optics, where the length of the line is large compared to the wavelength of the signal (the signal changes rapidly compared to the time it takes to travel from source to load), the impedances at each end of the line must be matched to the transmission line's characteristic impedance, Z_0 to prevent reflections of the signal at the ends of the line from causing echoes. (When the length of the line is short compared to the wavelength, impedance mismatch is the basis of transmission line impedance transformers, a topic that is addressed in the previous section.) In radio-frequency (RF) systems, a common value for source and load impedances is 50 ohms. A typical RF load is a quarter-wave ground plane antenna (37 ohms with an ideal ground plane but can be matched to 50 ohms by using a modified ground plane or a coaxial matching section, i.e. part or all the feeder being of higher impedance).

In a transmission line, a wave travels from the source along the line. Suppose the wave hits a boundary (an abrupt change in impedance). Some of the wave is reflected back, while some keeps moving onwards. (Assume there's only one boundary.)

At the boundary, the two waves on the source side of the boundary (with impedance Z_1) will be equal to the waves on the load side (with impedance Z_2). The derivatives will also be equal. Using that equality, we solve for all wave functions, getting a reflection coefficient:

$$\Gamma = \frac{Z_2 - Z_1}{Z_1 + Z_2}$$

The purpose of a transmission line is to get the maximum amount of energy to the other end of the line, or to transmit information with minimal error, so the reflection should be as small as possible. This is achieved by matching the impedances Z_1 and Z_2 so that they are equal ($\Gamma = 0$).

An electromagnetic wave consists of energy being transmitted down the transmission line. This energy is in two forms, an electric field and a magnetic field, which fluctuate constantly, with a continuing exchange between electrical and magnetic energy. The electric field is due to the voltage over the cross section of the line, perpendicular to the direction the wave is flowing. The magnetic field is due to the current flowing parallel to the direction of the wave.

Assume that voltage and current vary as sine waves. Inside the transmission line, the law of conservation of energy applies: the sum of magnetic and electric energy must always be the same (ignoring the effect of the small amount of energy converted to heat). This means that if the voltage is changing rapidly, the current must also change rapidly.

Now consider two moments: 1). when the current is zero and the voltage is maximum; 2). when the current is maximum and the voltage is zero. The amount of energy stored in the electric field at 1). must be exactly the same as the amount of energy stored in the magnetic field at 2). The ratio between voltage and current at 1). and 2). determines the impedance (Z) of the line:

$$Z_0 = \frac{V}{I}$$

At a boundary, for example, where the line is connected to the receiver, the law of conservation of charge applies. The current just before the boundary must be the same as just after. However, if the circuit at the receiver has a different impedance, Z_L, than the line, the voltage will be $V_L = Z_L I$ at the receiver, which is not the same as the original incident voltage V_0^+.

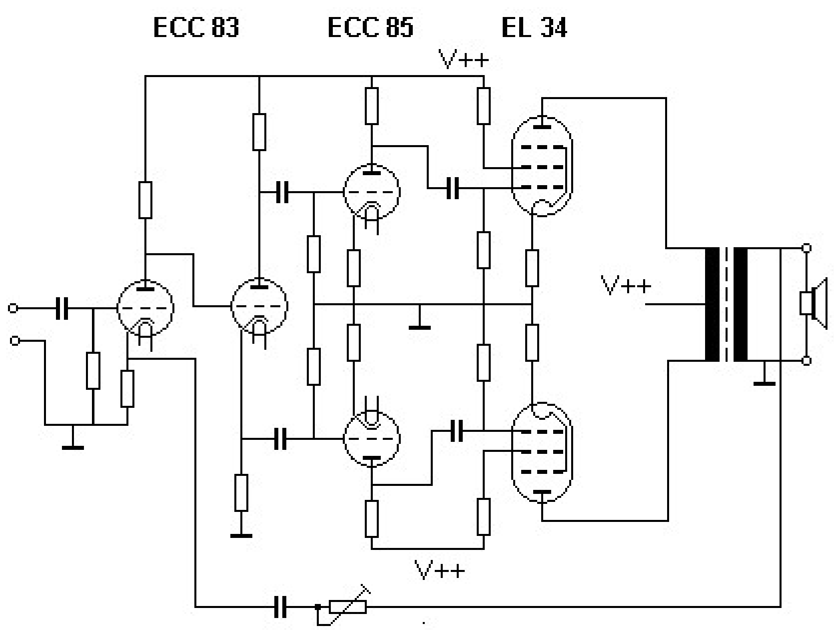
To achieve the voltage difference, an electric field is needed over the boundary. However, energy is needed to form this field, for which a part of the energy of the original wave is used. The remaining energy cannot just 'disappear'; it must go somewhere. Due to the impedance and voltage difference, it cannot go to the other side of the boundary. There remains only one way to go for this energy: back into the transmission line, as a reflection. The voltage of this reflected wave, V_0^-, is calculated from the incident voltage V_0^+ and the reflection coefficient, Γ (from the formula above):

$$V_0^- = \Gamma V_0^+$$

Electrical examples

Telephone systems

Telephone systems also use matched impedances to minimise echoes on long distance lines. This is related to transmission lines theory. Matching also enables the telephone *hybrid coil* (2 to 4 wire conversion) to operate correctly. As the signals are sent and received on the same two-wire circuit to the central office (or exchange), cancellation is necessary at the telephone earpiece so that excessive sidetone is not heard. All devices used in telephone signal paths are generally dependent on using matched cable, source and load impedances. In the local loop, the impedance chosen is 600 ohm (nominal). Terminating networks are installed at the exchange to try to offer the best match to their subscriber lines. Each country has its own standard for these networks but they are all designed to approximate to about 600 ohms over the voice frequency band.



Typical push-pull audio tube power amplifier matched to the loudspeaker with an impedance matching transformer.

Loudspeaker amplifiers

Many modern solid state audio amplifiers do not use matched impedances, because semiconductor based amplifiers do not have output transformers. The driver amplifier has a low output impedance, such as < 0.1 ohm, and the loudspeaker usually has an input impedance of 4, 8, or 16 ohms, which is many times larger than the former. This type of connection is impedance bridging, and it provides better damping of the loudspeaker cone to minimize distortion. The misconception arises from tube audio amplifiers, which required impedance matching for proper, reliable operation. Most of these had output transformer taps to approximately match the amplifier output to typical loudspeaker impedances.

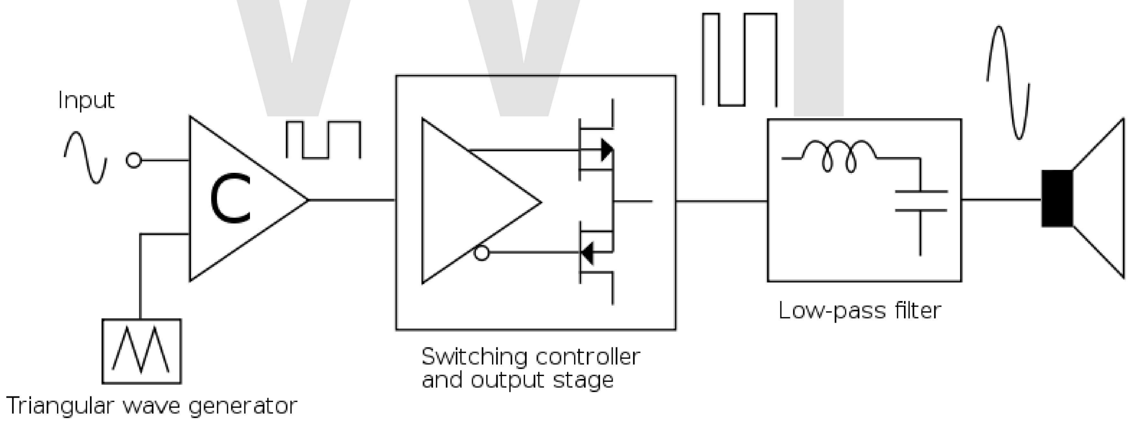
The output transformer in the vacuum tube based amplifiers has two basic functions:

1. Separation of the AC component, which contains the audio signals, from the DC component, supplied by the power supply, in the anode circuit of vacuum tube based power stage. A loudspeaker should not be subjected to DC current.
2. Matching of the low impedance of popular loudspeakers to the high internal resistance of power pentodes such as the EL34.

The impedance of the loudspeaker on the secondary coil of the transformer will be transformed to a higher impedance on the primary coil in the circuit of the power pentodes by the square of the turns ratio (impedance scaling factor).
The secondary impedance of the loudspeaker is frequently moved (or "referred") to the primary side after multiplying the components by the impedance scaling factor $\left(\frac{N_P}{N_S}\right)^2$.

- *Np* = the number of windings of the primary coil
- *Ns* = the number of windings of the secondary coil

The required turns ratio can be computed with a given internal resistance of power pentodes in parallel = 3500 Ohm and a given loudspeaker impedance = 4 Ohm:



Class D amplifier with integrator in the endstage

$$\sqrt{\frac{3500}{4}} = 29.6$$

is the turns ratio of the output transformer. A careless selection of the tap of the output transformer to match the impedance of the loudspeaker used will result in a power loss.

The output stage in semiconductor based endstages with mosfets or powertransistors, do have a very low internal resistances. If they are properly balanced, then there is no need for a device as a transformer or a (big) electrolytic capacitor to separate AC current from DC current. In class D amplifiers, based on pulse conversion, is only a integrator necessary to generate the audio signal and to block the sample frequencies

Non-electrical examples

Acoustics

Similar to electrical transmission lines, the impedance matching problem exists when transferring sound energy from one medium to another. If the acoustic impedance of the two media are very different, then most of the sound energy will be reflected or absorbed, rather than transferred across the border.

The gel used in medical ultrasonography helps transfer acoustic energy from the transducer to the body and back again. Without the gel, the "impedance mismatch" in the transducer-to-air and the air-to-body discontinuity reflects almost all the energy, leaving very little to go into the body.

Horns are used like transformers, matching the impedance of the transducer to the impedance of the air. This principle is used in both horn loudspeakers and musical instruments.

Most loudspeaker systems themselves contain impedance matching mechanisms, especially for low frequencies. Because most driver impedances are poorly matched to the impedance of free air at low frequencies, and because of out-of-phase cancellations between output from the front of a speaker cone and from the rear, loudspeaker enclosures serve both to match impedances and prevent the interference. Sound coupling into air from a loudspeaker is related to the ratio of the diameter of the speaker to the wavelength of the sound being reproduced. That is, larger speakers can produce lower frequencies at higher levels than smaller speakers for this reason. Elliptical speakers are a complex case, acting like large speakers lengthwise, and like small speakers crosswise.

Acoustic impedance matching (or the lack of it) affects the operation of a megaphone, an echo, and soundproofing.

Optics

A similar effect occurs when light (or any electromagnetic wave) hits the interface between two media with different refractive indices. For non-magnetic materials, refractive index is inversely proportional to the material's characteristic impedance. An *optical* or *wave impedance* that depends on the propagation direction can be calculated for each medium, and may be used in the usual transmission line reflection equation

$$r = \frac{Z_2 - Z_1}{Z_1 + Z_2}$$

to calculate the reflection and transmission coefficients for the interface. For non-magnetic dielectrics, this equation is equivalent to the Fresnel equations. Unwanted reflections can be reduced by the use of an anti-reflection optical coating.

Mechanics

If a body of mass *m* collides elastically with a second body, the maximum energy transferred to the second body will occur when the second body has the same mass *m*. For a head-on collision, with equal masses, the energy of the first body will be completely transferred to the second body. In this case, the masses act as "mechanical impedances" which must be matched. If m_1 and m_2 are the masses of the moving and the stationary body respectively, and *P* is the momentum of the system, which remains constant throughout the collision, then the energy of the second body after the collision will be E_2:

$$E_2 = \frac{2P^2 m_2}{(m_1 + m_2)^2}$$

which is analogous to the power transfer equation in the above "mathematical proof" section.

These principles are useful in the application of highly energetic materials (explosives). If an explosive charge is placed upon a target, the sudden release of energy causes compression waves to propagate through the target radially from the point charge contact. When the compression waves reach areas of high acoustic impedance mismatch (like the other side of the target), tension waves reflect back and create spalling. The greater the mismatch, the greater the effect of creasing and spalling will be. A charge initiated against a wall with air behind it will do more damage to the wall than a charge initiated against a wall with soil behind it.