BM2203 – Sensors and Measurements

Mr.G.Thiyagarajan

Lecturer

Department of Biomedical Engineering

Measurement:

Measurement is the estimation of the magnitude of some attribute of an object, such as its length or weight, relative to a unit of measurement. Measurement usually involves using a measuring instrument, such as a ruler or scale, which is calibrated to compare the object to some standard, such as a meter or a kilogram. In science, however, where accurate measurement is crucial, a measurement is understood to have three parts: first, the measurement itself, second, the margin of error, and third, the confidence level -- that is, the probability that the actual property of the physical object is within the margin of error. For example, we might measure the length of an object as 2.34 meters plus or minus 0.01 meter, with a 95% level of confidence.

Metrology is the scientific study of measurement. In measurement theory a measurement is an observation that reduces an uncertainty expressed as a quantity. As a verb, measurement is making such observations[1]. It includes the estimation of a physical quantity such as distance, energy, temperature, or time. It could also include such things as assessment of attitudes, values and perception in surveys or the testing of aptitudes of individuals.

In the physical sciences, measurement is most commonly thought of as the ratio of some physical quantity to a standard quantity of the same type, thus a measurement of length is the ratio of a physical length to some standard length, such as a standard meter. Measurements are usually given in terms of a real number times a unit of measurement, for example 2.53 meters, but sometimes measurements use complex numbers, as in measurements of electrical impedance.

Observations and error:

The act of measuring often requires an instrument designed and calibrated for that purpose, such as a thermometer, speedometer, weighing scale, or voltmeter. Surveys and tests are also referred to as "measurement instruments" in academic testing, aptitude testing, voter polls, etc.

Measurements always have errors and therefore uncertainties. In fact, the reduction—not necessarily the elimination—of uncertainty is central to the concept of measurement. Measurement errors are often assumed to be normally distributed about the true value of the measured quantity. Under this assumption, every measurement has three components: the estimate, the error bound, and the probability that the actual magnitude lies within the error bound of the estimate. For example, a measurement of the length of a plank might result in a measurement of 2.53 meters plus or minus 0.01 meter, with a probability of 99%.

Measurement is fundamental in science; it is one of the things that distinguishes science from pseudoscience. It is easy to come up with a theory about nature, hard to come up with a scientific theory that predicts measurements with great accuracy. Measurement is also essential in industry, commerce, engineering, construction, manufacturing, pharmaceutical production, and electronics.

Observational error is the difference between a measured value of quantity and its true value. In statistics, an error is not a "mistake". Variability is an inherent part of things being measured and of the measurement process.

When either randomness or uncertainty modeled by probability theory is attributed to such errors, they are "errors" in the sense in which that term is used in statistics; see errors and residuals in statistics.

Every time we repeat a measurement with a sensitive instrument, we obtain slightly different results. The common statistical model we use is that the error has two additive parts:

systematic error which always occurs (with the same value) when we use the instrument in the same way, and
random error which may vary from observation to observation.

The systematic error is sometimes called statistical bias. It is controlled by very carefully standardized procedures. Part of the education in every science is how to use the standard instruments of the discipline.

The random error (or random variation) is due to factors which we cannot (or do not) control. It may be too expensive or we may be too ignorant of these factors to control them each time we measure. It may even be that whatever we are trying to measure is changing in time (see dynamic models), or is fundamentally probablistic (as is the case in quantum mechanics -- see Measurement in quantum mechanics). Random error often occurs when instruments are pushed to their limits. For example, it is common for digital balances to exhibit random error in their least significant digit. Three measurements of a single object might read something like 0.9111g, 0.9110g, and 0.9112g.

Instrumentation is an electrical or pneumatic device placed in the field to provide measurement and/or control capabilities for the system.

The simplest measurement instrumentation device is a thermistor. A thermistor is very similar to a typical resistor, except that it greatly varies its resistance depending on its temperature. Therefore this device can easily be used for measurement of temperature in the field. Other temperature-sensitive devices include RTDs, which also change resistance depending on temperature, and thermocouples, which produce a varying voltage when subjected to heat.

Control instrumentation includes devices such as solenoids, Electrically Operated Valves, breakers, relays, etc. These devices are able to change a field parameter, and provide remote control capabilities.

Transmitters are devices which produce an analog signal, usually in the form of a 4-20 mA electrical current signal, although many other options are possible using voltage, frequency, or pressure. This signal can be used to directly control other instruments, or sent to a PLC, DCS, SCADA system or other type of computerized controller, where it can be interpreted into readable values, or used to control other devices and processes in the system.

Instrumentation plays a significant role in both gathering information from the field and changing the field parameters, and as such are a key part of control loops.

Instrumentation can be used to measure certain field parameters (physical values):

These measured values include:

pressure, either differential or static
flow
temperature
level
density
viscosity
radiation
current
voltage
inductance
capacitance
frequency
resistivity
conductivity
chemical composition
chemical properties
various physical properties
force applied by a liquid

Example:-

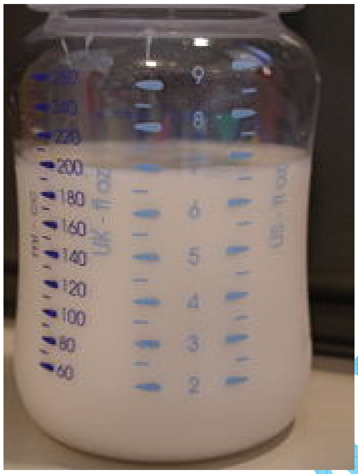
Resistance thermometers, also called **resistance temperature detectors** (**RTD**s), are temperature sensors that exploit the predictable change in electrical resistance of some materials with changing temperature. As they are almost invariably made of platinum, they are often called **platinum resistance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

Standards:

Laws to regulate measurement were originally developed to prevent fraud. However, units of measurement are now generally defined on a scientific basis, and are established by international treaties. In the United States, commercial measurements are regulated by the National Institute of Standards and Technology NIST, a division of the United States Department of Commerce.

Units and systems:

Main articles: Units of measurement and Systems of measurement



A baby bottle that measures in all three measurement systems—Imperial (U.K.), U.S. Customary, and metric.

The definition or specification of precise standards of measurement involves two key features, which are evident in the International System of Units (SI). Specifically, in this system the definition of each of the *base* units makes reference to specific empirical conditions and, with the exception of the kilogram, also to other quantitative attributes. Each *derived* SI unit is defined purely in terms of a relationship involving itself and other units; for example, the unit of velocity is 1 m/s. Due to the fact that derived units make reference to base units, the specification of empirical conditions is an implied component of the definition of all units.

SI

Main article: International System of Units

The International System of Units (abbreviated **SI** from the French language name *Système International d'Unités*) is the modern, revised form of the metric system. It is the world's most widely used system of units, both in everyday commerce and in science. The SI was developed in 1960 from the metre-kilogram-second (MKS) system, rather than the centimetre-gram-second (CGS) system, which, in turn, had many variants. At its development the SI also introduced several newly named units that were previously not a part of the metric system.

There are two types of SI units, base and derived units. Base units are the simple measurements for time, length, mass, temperature, amount of substance, electric current, and light intensity. Derived units are made up of base units, for example density is kg/m^3.

Mean:

In statistics, ***mean*** has two related meanings:

the arithmetic mean (and is distinguished from the geometric mean or harmonic mean).
the expected value of a random variable, which is also called the *population mean*.

It is sometimes stated that the 'mean' means average. This is incorrect if "mean" is taken in the specific sense of "arithmetic mean" as there are different types of averages: the mean, median, and mode. For instance, average house prices almost always use the median value for the average.

For a real-valued random variable X, the mean is the expectation of X. Note that not every probability distribution has a defined mean (or variance); see the Cauchy distribution for an example.

For a data set, the mean is the sum of the observations divided by the number of observations. The mean is often quoted along with the standard deviation: the mean describes the central location of the data, and the standard deviation describes the spread.

An alternative measure of dispersion is the mean deviation, equivalent to the average absolute deviation from the mean. It is less sensitive to outliers, but less mathematically tractable.

As well as statistics, means are often used in geometry and analysis; a wide range of means have been developed for these purposes, which are not much used in statistics. These are listed below.

Arithmetic mean

Main article: Arithmetic mean

The *arithmetic mean* is the "standard" average, often simply called the "mean".

The **mean** may often be confused with the median or mode. The mean is the arithmetic average of a set of values, or distribution; however, for skewed distributions, the mean is not necessarily the same as the middle value (median), or the most likely (mode). For example, mean income is skewed upwards by a small number of people with very large incomes, so that the majority have an income lower than the mean. By contrast, the median income is the level at which half the population is below and half is above. The mode income is the most likely income, and favors the larger number of people with lower incomes. The median or mode are often more intuitive measures of such data.

That said, many skewed distributions are best described by their mean - such as the Exponential and Poisson distributions.

For example, the arithmetic mean of six values: 34, 27, 45, 55, 22, 34 is:

$34+27+45+55+22+34/6=36.167$

Standard deviation:

In probability and statistics, the **standard deviation** of a probability distribution, random variable, or population or multiset of values is a measure of the spread of its values. The standard deviation is usually denoted with the letter σ (lower case sigma). It is defined as the square root of the variance.

To understand standard deviation, keep in mind that variance is the *average* of the squared differences between data points and the mean. Variance is tabulated in units squared. Standard deviation, being the square root of that quantity, therefore measures the spread of data about the mean, measured in the same units as the data.

Stated more formally, the standard deviation is the root mean square (RMS) deviation of values from their arithmetic mean.

For example, in the population {4, 8}, the mean is 6 and the deviations from mean are {−2, 2}. Those deviations squared are {4, 4} the average of which (the variance) is 4. Therefore, the standard deviation is 2. In this case 100% of the values in the population are at one standard deviation from the mean.

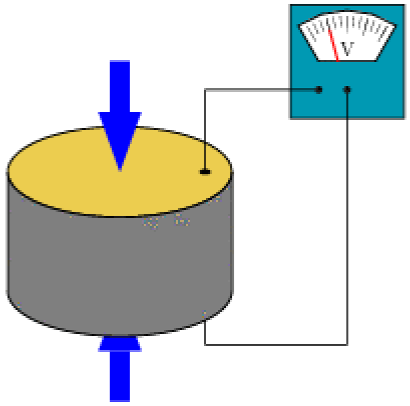
The standard deviation is the most common measure of statistical dispersion, measuring how widely spread the values in a data set are. If many data points are close to the mean, then the standard deviation is small; if many data points are far from the mean, then the standard deviation is large. If all the data values are equal, then the standard deviation is zero.

For a population, the standard deviation can be estimated by a modified standard deviation (s) of a sample.

Transducer

A **transducer** is a device, usually electrical, electronic, electro-mechanical, electromagnetic, photonic, or photovoltaic that converts one type of energy or physical attribute to another for various purposes including measurement or information transfer (for example, pressure sensors).

The term transducer is commonly used in two senses; the sensor, used to detect a parameter in one form and report it in another (usually an electrical or digital signal), and the audio loudspeaker, which converts electrical voltage variations representing music or speech, to mechanical cone vibration and hence vibrates air molecules creating acoustical energy.



Types of transducers:

This list is confined to the narrower definition of the term.

Electromagnetic:
Antenna - converts electromagnetic waves into electric current and vice versa.
Cathode ray tube (CRT) - converts electrical signals into visual form
Fluorescent lamp, light bulb - converts electrical power into visible light
Magnetic cartridge - converts motion into electrical form
Photodetector or Photoresistor (LDR) - converts changes in light levels into resistance changes
Tape head - converts changing magnetic fields into electrical form
Hall effect sensor - converts a magnetic field level into electrical form only.
Electrochemical:

pH probes
Electro-galvanic fuel cell
Electromechanical (electromechanical output devices are generically called actuators):
Electroactive polymers
Galvanometer
MEMS
Rotary motor, linear motor
Vibration powered generator
Potentiometer when used for measuring position
Load cell converts force to mV/V electrical signal using strain gauge
Accelerometer
Strain gauge
String Potentiometer
Air flow sensor

Electroacoustic:
Geophone - convert a ground movement (displacement) into voltage
Gramophone pick-up
Hydrophone - converts changes in water pressure into an electrical form
Loudspeaker, earphone - converts changes in electrical signals into acoustic form
Microphone - converts changes in air pressure into an electrical signal
Piezoelectric crystal - converts pressure changes into electrical form
Tactile transducer
Photoelectric:
Laser diode, light-emitting diode - convert electrical power into forms of light
Photodiode, photoresistor, phototransistor, photomultiplier tube - converts changing light levels into electrical form
Electrostatic:
Electrometer
Thermoelectric:
RTD Resistance Temperature Detector
Thermocouple
Peltier cooler
Thermistor (includes PTC resistor and NTC resistor)
Radioacoustic:
Geiger-Müller tube used for measuring radioactivity.
Receiver (radio)

Temperature :

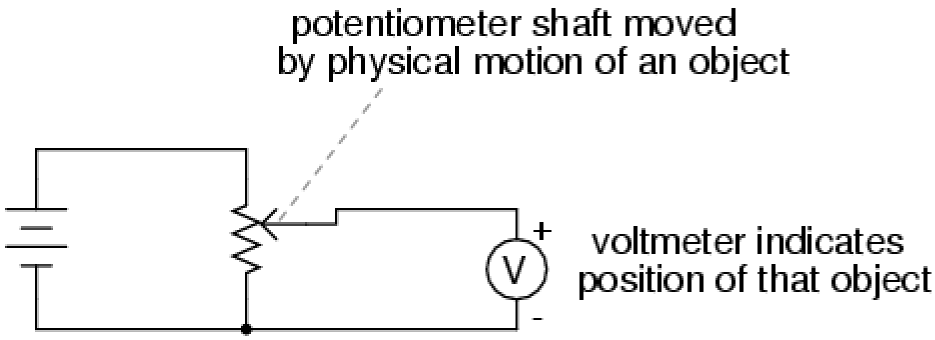
1. **Resistance thermometers**, also called **resistance temperature detectors** (**RTD**s), are temperature sensors that exploit the predictable change in electrical resistance of some materials with changing temperature. As they are almost invariably made of platinum, they are often called **platinum resistance thermometers** (**PRT**s).

They are slowly replacing the use of thermocouples in many industrial applications below 600 °C.

AC instrumentation transducers

Just as devices have been made to measure certain physical quantities and repeat that information in the form of DC electrical signals (thermocouples, strain gauges, pH probes, etc.), special devices have been made that do the same with AC.

It is often necessary to be able to detect and transmit the physical position of mechanical parts via electrical signals. This is especially true in the fields of automated machine tool control and robotics. A simple and easy way to do this is with a potentiometer: (Figure below)

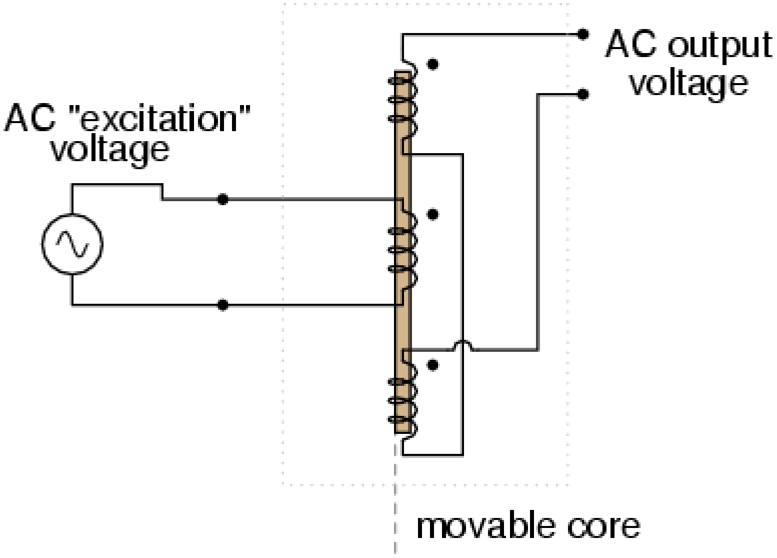


Potentiometer tap voltage indicates position of an object slaved to the shaft.

However, potentiometers have their own unique problems. For one, they rely on physical contact between the “wiper” and the resistance strip, which means they suffer the effects of physical wear over time. As potentiometers wear, their proportional output versus shaft position becomes less and less certain. You might have already experienced this effect when adjusting the volume control on an old radio: when twisting the knob, you might hear “scratching” sounds coming out of the speakers. Those noises are the result of poor wiper contact in the volume control potentiometer.

Also, this physical contact between wiper and strip creates the possibility of arcing (sparking) between the two as the wiper is moved. With most potentiometer circuits, the current is so low that wiper arcing is negligible, but it is a possibility to be considered. If the potentiometer is to be operated in an environment where combustible vapor or dust is present, this potential for arcing translates into a potential for an explosion!

Using AC instead of DC, we are able to completely avoid sliding contact between parts if we use a *variable transformer* instead of a potentiometer. Devices made for this purpose are called LVDT's, which stands for **L**inear **V**ariable **D**ifferential **T**ransformers. The design of an LVDT looks like this: (Figure below)



AC output of linear variable differential transformer (LVDT) indicates core position.

Obviously, this device is a *transformer*: it has a single primary winding powered by an external source of AC voltage, and two secondary windings connected in series-bucking fashion. It is *variable* because the core is free to move between the windings. It is *differential* because of the way the two secondary windings are connected. Being arranged to oppose each other (180° out of phase) means that the output of this device will be the *difference* between the voltage output of the two secondary windings. When the core is centered and both windings are outputting the same voltage, the net result at the output terminals will be zero volts. It is called *linear* because the core's freedom of motion is straight-line.

The AC voltage output by an LVDT indicates the position of the movable core. Zero volts means that the core is centered. The further away the core is from center position, the greater percentage of input ("excitation") voltage will be seen at the output. The phase of the output voltage relative to the excitation voltage indicates which direction from center the core is offset.

The primary advantage of an LVDT over a potentiometer for position sensing is the absence of physical contact between the moving and stationary parts. The core does not contact the wire windings, but slides in and out within a nonconducting tube. Thus, the LVDT does not "wear" like a potentiometer, nor is there the possibility of creating an arc. Excitation of the LVDT is typically 10 volts RMS or less, at frequencies ranging from power line to the high audio (20 kHz) range. One potential disadvantage of the LVDT is its response time, which is mostly dependent on the frequency of the AC voltage source. If very quick response times are desired, the frequency must be higher to allow whatever voltage-sensing circuits enough cycles of AC to determine voltage level as the core is moved. To illustrate the potential problem here, imagine this exaggerated scenario: an LVDT powered by a 60 Hz voltage source, with the core being moved in and out hundreds of times per second. The output of this LVDT wouldn't even look like a sine wave because the core would be moved throughout its range of motion before the AC source voltage could complete a single cycle! It would be almost impossible to determine instantaneous core position if it moves faster than the instantaneous source voltage does.

General description

There are two broad categories, "film" and "wire-wound" types.

Film thermometers have a layer of platinum on a substrate; the layer may be extremely thin, perhaps 1 micrometer. Advantages of this type are relatively low cost and fast response. Such devices have improved in performance although the different expansion rates of the substrate and platinum give
Wire-wound thermometers can have greater accuracy, especially for wide temperature ranges. The coil diameter provides a compromise between mechanical stability and allowing expansion of the wire to minimize strain and consequential drift.

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

How do resistance thermometers work?

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

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

Commercial platinum grades are produced which exhibit a change of resistance of 0.385 ohms/°C (European Fundamental Interval) The sensor is usually made to have a resistance of 100Ω at 0 °C. This is defined in BS EN 60751:1996 (taken from IEC 60751:1995) . The American Fundamental Interval is 0.392 Ω/°C, based on using a purer grade of platinum than the European standard. The American standard is from the Scientific Apparatus Manufacturers Association (SAMA), who are no longer in this standards field.

Resistance thermometers require a small current to be passed through in order to determine the resistance. This can cause resistive heating, and manufacturers' limits should always be followed along with heat path considerations in design. Care should also be taken to avoid any strains on the resistance thermometer in its application. Lead wire resistance should be considered, and adopting three and four wire connections can eliminate connection lead resistance effects from measurements - industrial practice is almost universally to use 3-wire connection.

Advantages and limitations

Advantages of platinum resistance thermometers:

High accuracy
Low drift
Wide operating range
Suitability for precision applications

Limitations:

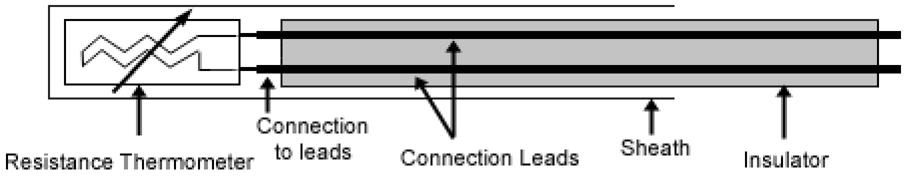
RTDs in industrial applications are rarely used above 660 °C. At temperatures above 660 °C it becomes increasingly difficult to prevent the platinum from becoming contaminated by impurities from the metal sheath of the thermometer. This is why laboratory standard thermometers replace the metal sheath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

Resistance thermometer elements

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

Wire wound in a ceramic insulator - wire spiral within sealed ceramic cylinder, works with temperatures to 850 °C
Wire encapsulated in glass - wire around glass core with glass fused homogenously around, resists vibration, more protection to the detecting wire but smaller usable range
Thin film - platinum film on ceramic substrate, small and inexpensive to mass produce, fast response to temperature change

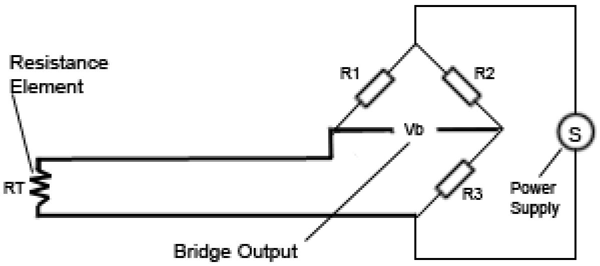
Resistance thermometer construction



These elements nearly always require insulated leads attached. At low temperatures PVC, silicon rubber or PTFE insulators are common to 250°C. Above this, glass fibre or ceramic are used. The measuring point and usually most of the leads require a housing or protection sleeve. This is often a metal alloy which is inert to a particular process. Often more consideration goes in to selecting and designing protection sheaths than sensors as this is the layer that must withstand chemical or physical attack and offer convenient process attachment points.

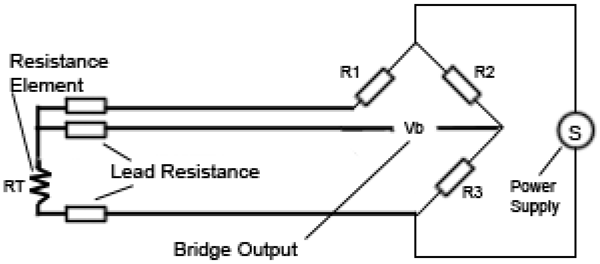
Resistance thermometer wiring configurations

Two-wire configuration



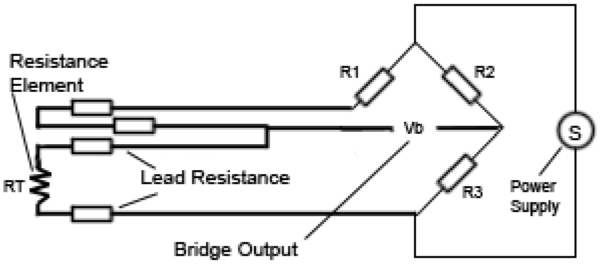
The simplest resistance thermometer configuration uses two wires. It is only used when high accuracy is not required as the resistance of the connecting wires is always included with that of the sensor leading to errors in the signal. Using this configuration you will be able to use 100 metres of cable. This applies equally to balanced bridge and fixed bridge system.

Three-wire configuration

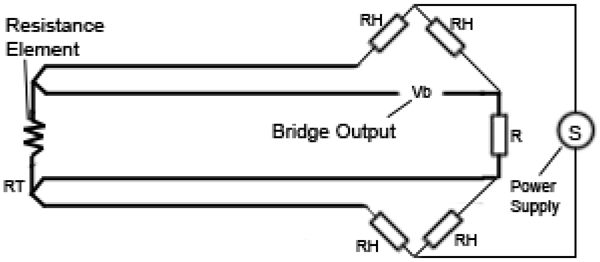


In order to minimize the effects of the lead resistances a three wire configuration can be used. Using this method the two leads to the sensor are on adjoining arms, there is a lead resistance in each arm of the bridge and therefore the lead resistance is cancelled out. High quality connection cables should be used for this type of configuration because an assumption is made that the two lead resistances are the same. This configuration allows for up to 600 meters of cable.

Four-wire configuration



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



2. **Thermocouples**

In electronics and in electrical engineering, **thermocouples** are a widely used type of temperature sensor[1] and can also be used as a means to convert thermal potential difference into electric potential difference.[2] They are cheap[3] and interchangeable, have standard connectors, and can measure a wide range of temperatures. The main limitation is accuracy; system errors of less than one degree Celsius (°C) can be difficult to achieve.

Principle of operations:

In 1821, the German–Estonian physicist Thomas Johann Seebeck discovered that when any conductor (such as a metal) is subjected to a thermal gradient, it will generate a voltage. This is now known as the thermoelectric effect or Seebeck effect. Any attempt to measure this voltage necessarily involves connecting another conductor to the "hot" end. This additional conductor will then also experience the temperature gradient, and develop a voltage of its own which will oppose the original. Fortunately, the magnitude of the effect depends on the metal in use. Using a dissimilar metal to complete the circuit creates a circuit in which the two legs generate different voltages, leaving a small difference in voltage available for measurement. That difference increases with temperature, and can typically be between one and seventy microvolts per degree Celsius (µV/°C) for the modern range of available metal combinations. Certain combinations have become popular as industry standards, driven by cost, availability, convenience, melting point, chemical properties, stability, and output. This coupling of two metals gives the thermocouple its name.

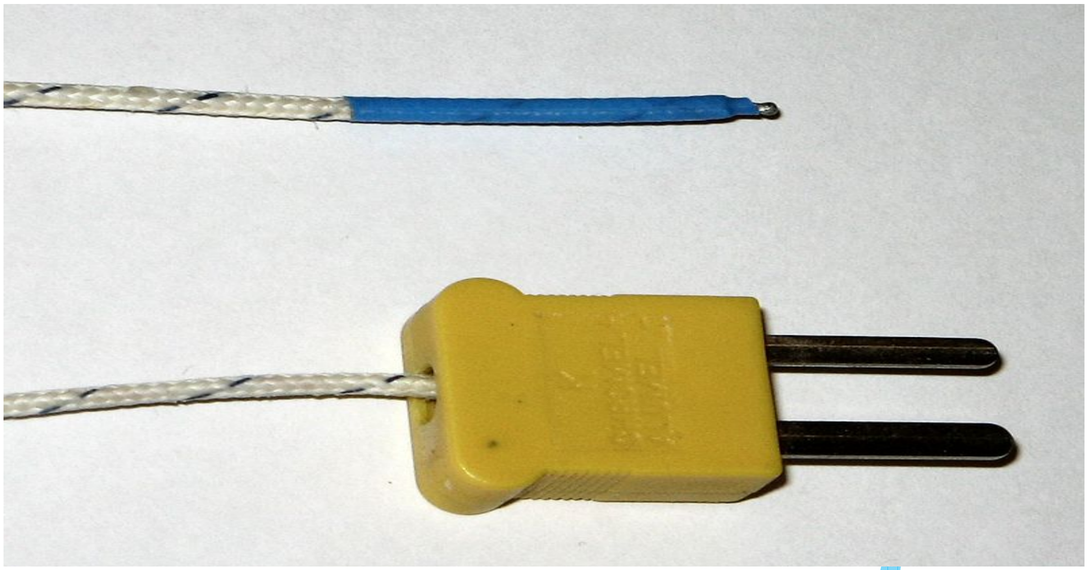
It is important to note that thermocouples measure the temperature difference between two points, not absolute temperature. In traditional applications, one of the junctions—the cold junction—was maintained at a known (reference) temperature, while the other end was attached to a probe.

Having available a known temperature cold junction, while useful for laboratory calibrations, is simply not convenient for most directly connected indicating and control instruments. They incorporate into their circuits an artificial cold junction using some other thermally sensitive device, such as a thermistor or diode, to measure the temperature of the input connections at the instrument, with special care being taken to minimize any temperature gradient between terminals. Hence, the voltage from a known cold junction can be simulated, and the appropriate correction applied. This is known as cold junction compensation.

Additionally, a device can perform cold junction compensation by computation. It can translate device voltages to temperatures by either of two methods. It can use values from look-up tables[4] or approximate using polynomial interpolation.

A thermocouple can produce current, which means it can be used to drive some processes directly, without the need for extra circuitry and power sources. For example, the power from a thermocouple can activate a valve when a temperature difference arises. The electric power generated by a thermocouple is a conversion of the heat energy that one must continuously supply to the hot side of the thermocouple to maintain the electric potential. The flow of heat is necessary because the current flowing through the thermocouple tends to cause the hot side to cool down and the cold side to heat up (the Peltier effect).

Thermocouples can be connected in series with each other to form a thermopile, where all the hot junctions are exposed to the higher temperature and all the cold junctions to a lower temperature. Thus, the voltages of the individual thermocouple add up, which allows for a larger voltage and increased power. With the radioactive decay of transuranic elements providing a heat source this arrangement has been used to power spacecraft on missions too far from the Sun to utilize solar power.



Applications:

Thermocouples are most suitable for measuring over a large temperature range, up to 1800 °C. They are less suitable for applications where smaller temperature differences need to be measured with high accuracy, for example the range 0–100 °C with 0.1 °C accuracy. For such applications, thermistors and resistance temperature detectors are more suitable.

Thermistor:

A **thermistor** is a type of resistor with resistance varying according to its temperature. The word is a combination of *thermal* and *resistor*. Samuel Ruben invented the thermistor in 1930, and was awarded U.S. Patent No. 2,021,491.

Thermistors are widely used as inrush current limiters, temperature sensors, self resetting overcurrent protectors, and self regulating heating elements.

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

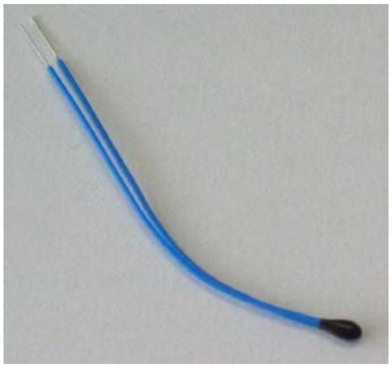
$\Delta R = k\Delta T$

where

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

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

Thermistors differ from resistance temperature detectors in that the material used in a thermistor is generally a ceramic or polymer, while RTDs use pure metals. The temperature response is also different; RTDs are useful over larger temperature ranges.



Applications:

PTC thermistors can be used as current-limiting devices for circuit protection, as replacements for fuses. Current through the device causes a small amount of resistive heating. If the current is large enough to generate more heat than the device can lose to its surroundings, the device heats up, causing its resistance to increase, and therefore causing even more heating. This creates a self-reinforcing effect that drives the resistance upwards, reducing the current and voltage available to the device.

PTC thermistors can be used as heating elements in small temperature-controlled ovens. As the temperature rises, resistance increases, decreasing the current and the heating. The result is a steady state. A typical application is a crystal oven controlling the temperature of the crystal of a high-precision crystal oscillator. Crystal ovens are usually set at the upper limit of the equipment's temperature specification, so they can maintain the temperature by heating.

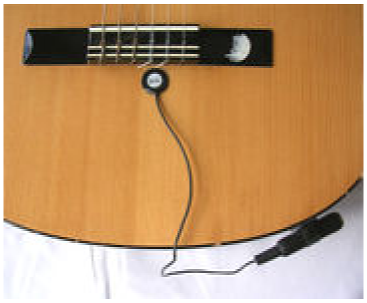
NTC thermistors are used as resistance thermometers in low-temperature measurements of the order of 10 K.

NTC thermistors can be used as inrush-current limiting devices in power supply circuits. They present a higher resistance initially which prevents large currents from flowing at turn-on, and then heat up and become much lower resistance to allow higher current flow during normal operation. These thermistors are usually much larger than measuring type thermistors, and are purposely designed for this application.

NTC thermistors are regularly used in automotive applications. For example they monitor things like coolant temperature and/or oil temperature inside the engine and provide data to the ECU and indirectly the dashboard.

Thermistors are also commonly used in modern digital thermostats and to monitor the temperature of battery packs while charging.

Sensors



Piezoelectric disk used as a guitar pickup

Main article: Piezoelectric sensor

The principle of operation of a piezoelectric sensor is that a physical dimension, transformed into a force, acts on two opposing faces of the sensing element. Depending on the design of a sensor, different "modes" to load the piezoelectric element can be used: longitudinal, transversal and shear.

Detection of pressure variations in the form of sound is the most common sensor application, e.g. piezoelectric microphones (sound waves bend the piezoelectric material, creating a changing voltage) and piezoelectric pickups for electrically amplified guitars. A piezo sensor attached to the body of an instrument is known as a contact microphone.

Piezoelectric sensors especially are used with high frequency sound in ultrasonic transducers for medical imaging and also industrial nondestructive testing (NDT).

For many sensing techniques, the sensor can act as both a sensor and an actuator - often the term *transducer* is preferred when the device acts in this dual capacity, but most piezo devices have this property of reversibility whether it is used or not. Ultrasonic transducers, for example, can inject ultrasound waves into the body, receive the returned wave, and convert it to an electrical signal (a voltage). Most medical ultrasound transducers are piezoelectric.

In addition to those mentioned above, various sensor applications include:

Piezoelectric elements are also used in the detection and generation of sonar waves.
Power monitoring in high power applications (e.g. medical treatment, sonochemistry and industrial processing).
Piezoelectric microbalances are used as very sensitive chemical and biological sensors.
Piezos are sometimes used in strain gauges.
Piezoelectric transducers are used in electronic drum pads to detect the impact of the drummer's sticks.
Automotive engine management systems use a piezoelectric transducer to detect detonation, by sampling the vibrations of the engine block.
Ultrasonic piezo sensors are used in the detection of acoustic emissions in acoustic emission testing.

Metal disk with piezoelectric disk attached, used in a buzzer

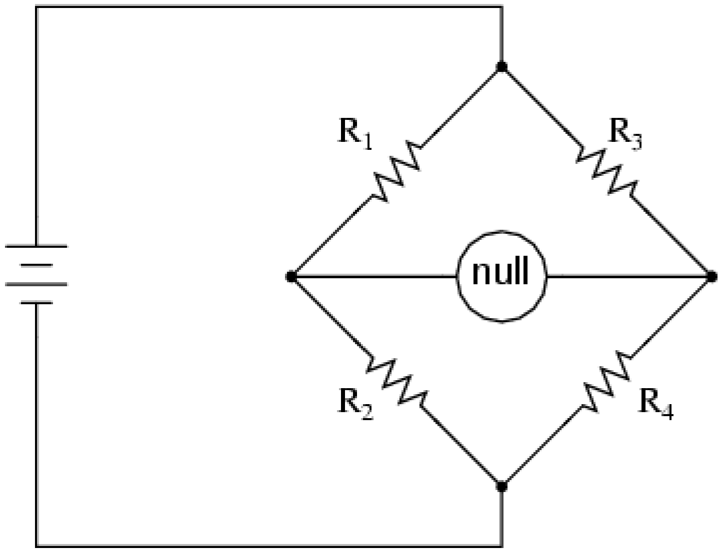
As very high voltages correspond to only tiny changes in the width of the crystal, this width can be changed with better-than-micrometer precision, making piezo crystals the most important tool for positioning objects with extreme accuracy — thus their use in actuators.

Loudspeakers: Voltages are converted to mechanical movement of a piezoelectric polymer film.
Piezoelectric motors: piezoelectric elements apply a directional force to an axle, causing it to rotate. Due to the extremely small distances involved, the piezo motor is viewed as a high-precision replacement for the stepper motor.
Piezoelectric elements can be used in laser mirror alignment, where their ability to move a large mass (the mirror mount) over microscopic distances is exploited to electronically align some laser mirrors. By precisely controlling the distance between mirrors, the laser electronics can accurately maintain optical conditions inside the laser cavity to optimize the beam output.
A related application is the acousto-optic modulator, a device that vibrates a mirror to give the light reflected off it a Doppler shift. This is useful for fine-tuning a laser's frequency.
Atomic force microscopes and scanning tunneling microscopes employ converse piezoelectricity to keep the sensing needle close to the probe.
Inkjet printers: On some inkjet printers, particularly those made by Epson, piezoelectric crystals are used to control the flow of ink from the inkjet head to the paper.
Diesel engines: high-performance common rail diesel engines use piezoelectric fuel injectors, first developed by Robert Bosch LLC, instead of the more common solenoid valve devices.

AC bridge circuits

As we saw with DC measurement circuits, the circuit configuration known as a *bridge* can be a very useful way to measure unknown values of resistance. This is true with AC as well, and we can apply the very same principle to the accurate measurement of unknown impedances.

To review, the bridge circuit works as a pair of two-component voltage dividers connected across the same source voltage, with a *null-detector* meter movement connected between them to indicate a condition of "balance" at zero volts: (Figure below)



A balanced bridge shows a "null", or minimum reading, on the indicator.

Any one of the four resistors in the above bridge can be the resistor of unknown value, and its value can be determined by a ratio of the other three, which are "calibrated," or whose resistances are known to a precise degree. When the bridge is in a balanced condition (zero voltage as indicated by the null detector), the ratio works out to be this:

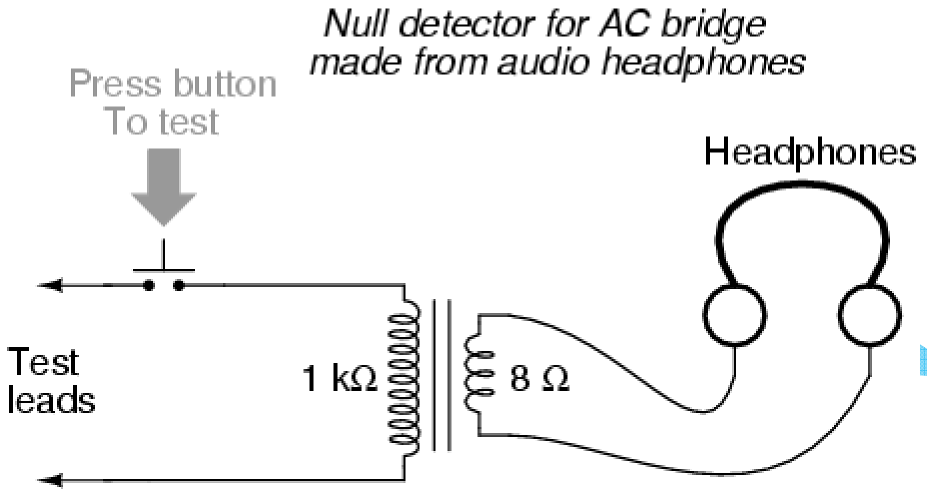
In a condition of **balance:**

$$\frac{R_1}{R_2} = \frac{R_3}{R_4}$$

One of the advantages of using a bridge circuit to measure resistance is that the voltage of the power source is irrelevant. Practically speaking, the higher the supply voltage, the easier it is to detect a condition of imbalance between the four resistors with the null detector, and thus the more sensitive it will be. A greater supply voltage leads to the possibility of increased measurement precision. However, there will be no fundamental error introduced as a result of a lesser or greater power supply voltage unlike other types of resistance measurement schemes.

Impedance bridges work the same, only the balance equation is with *complex* quantities, as both magnitude and phase across the components of the two dividers must be equal in order for the null detector to indicate "zero." The null detector, of course, must be a device capable of detecting very small AC voltages. An oscilloscope is often used for this, although very sensitive electromechanical meter movements and even headphones (small speakers) may be used if the source frequency is within audio range.

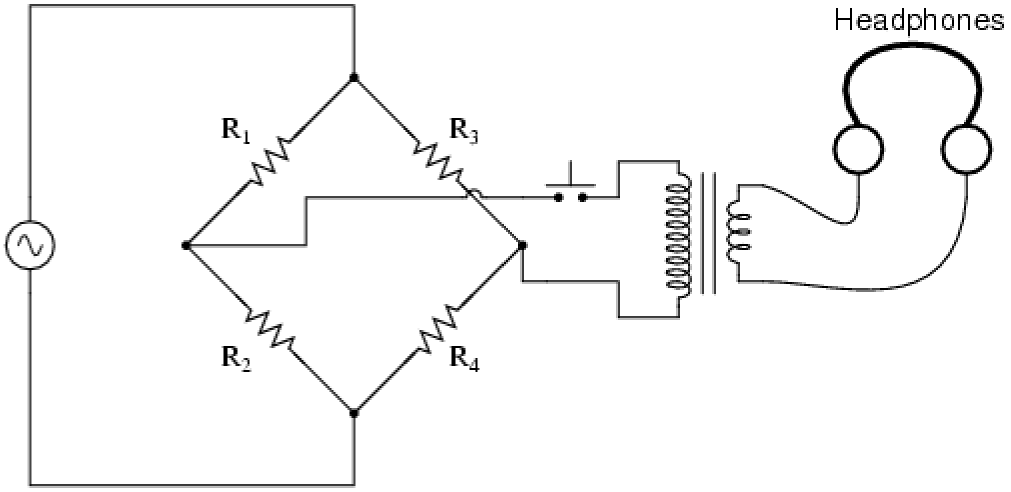
One way to maximize the effectiveness of audio headphones as a null detector is to connect them to the signal source through an impedance-matching transformer. Headphone speakers are typically low-impedance units (8 Ω), requiring substantial current to drive, and so a step-down transformer helps "match" low-current signals to the impedance of the headphone speakers. An audio output transformer works well for this purpose: (Figure below)



“Modern” low-Ohm headphones require an impedance matching transformer for use as a sensitive null detector.

Using a pair of headphones that completely surround the ears (the “closed-cup” type), I've been able to detect currents of less than 0.1 µA with this simple detector circuit. Roughly equal performance was obtained using two different step-down transformers: a small power transformer (120/6 volt ratio), and an audio output transformer (1000:8 ohm impedance ratio). With the pushbutton switch in place to interrupt current, this circuit is usable for detecting signals from DC to over 2 MHz: even if the frequency is far above or below the audio range, a “click” will be heard from the headphones each time the switch is pressed and released.

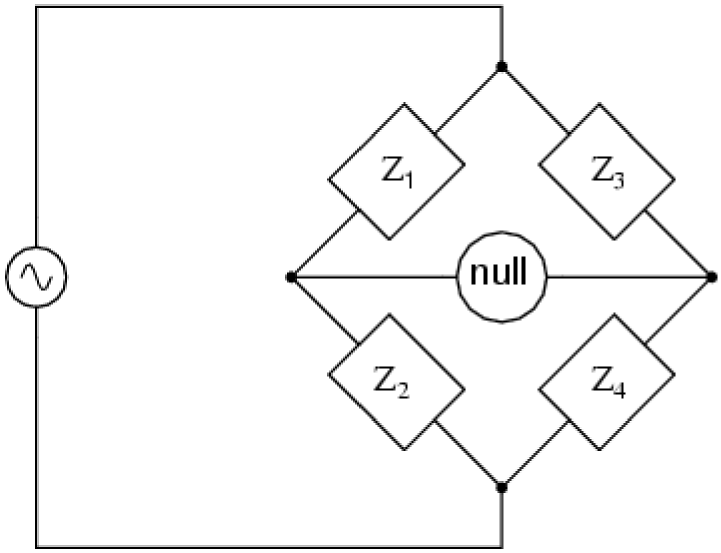
Connected to a resistive bridge, the whole circuit looks like Figure below.



Bridge with sensitive AC null detector.

Listening to the headphones as one or more of the resistor “arms” of the bridge is adjusted, a condition of balance will be realized when the headphones fail to produce “clicks” (or tones, if the bridge's power source frequency is within audio range) as the switch is actuated.

When describing general AC bridges, where *impedances* and not just resistances must be in proper ratio for balance, it is sometimes helpful to draw the respective bridge legs in the form of box-shaped components, each one with a certain impedance: (Figure below)



Generalized AC impedance bridge: Z = nonspecific complex impedance.

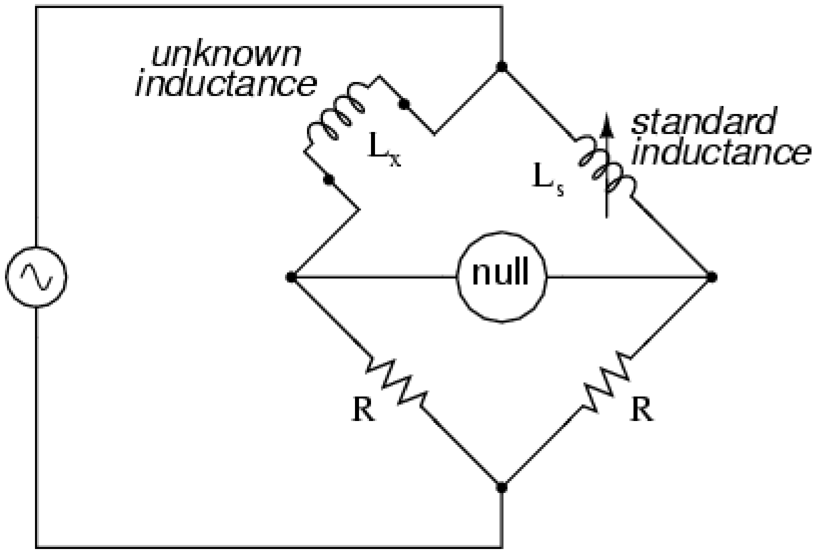
For this general form of AC bridge to balance, the impedance ratios of each branch must be equal:

$$\frac{Z_1}{Z_2} = \frac{Z_3}{Z_4}$$

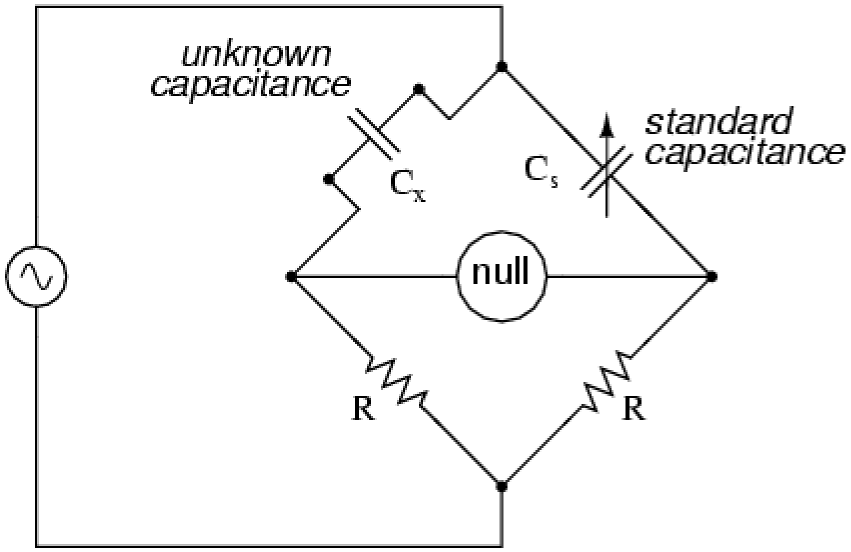
Again, it must be stressed that the impedance quantities in the above equation *must* be complex, accounting for both magnitude and phase angle. It is insufficient that the impedance magnitudes alone be balanced; without phase angles in balance as well, there will still be voltage across the terminals of the null detector and the bridge will not be balanced.

Bridge circuits can be constructed to measure just about any device value desired, be it capacitance, inductance, resistance, or even "Q." As always in bridge measurement circuits, the unknown quantity is always "balanced" against a known standard, obtained from a high-quality, calibrated component that can be adjusted in value until the null detector device indicates a condition of balance. Depending on how the bridge is set up, the unknown component's value may be determined directly from the setting of the calibrated standard, or derived from that standard through a mathematical formula.

A couple of simple bridge circuits are shown below, one for inductance (Figure below) and one for capacitance: (Figure below)



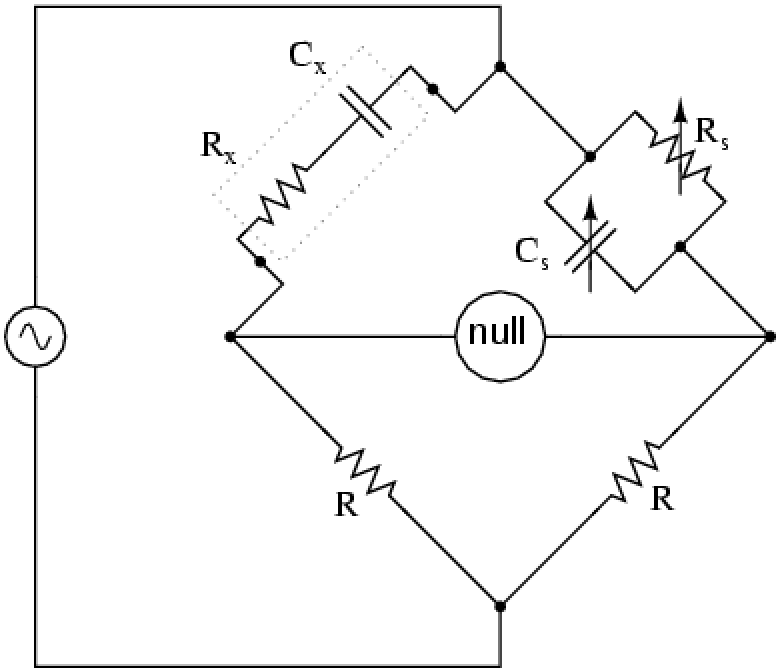
Symmetrical bridge measures unknown inductor by comparison to a standard inductor.



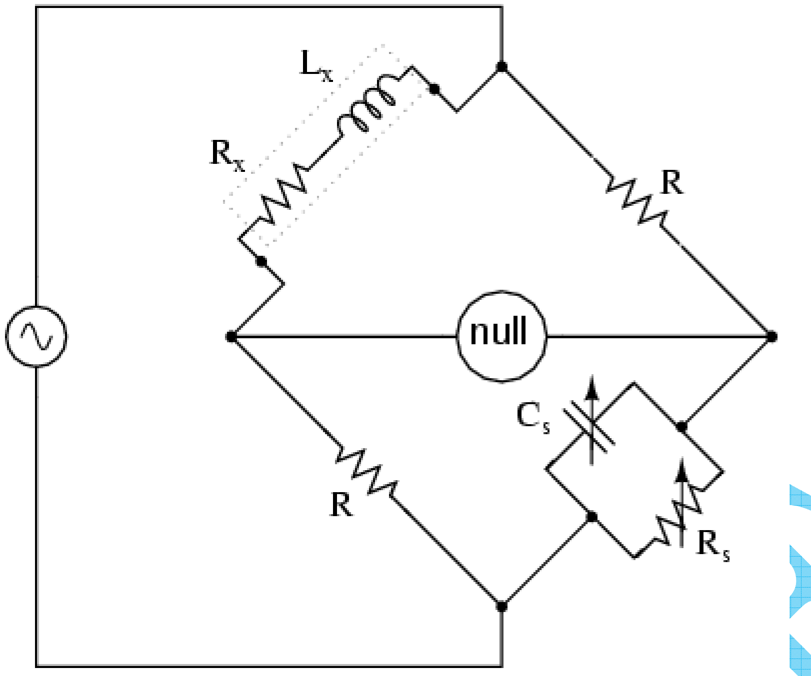
Symmetrical bridge measures unknown capacitor by comparison to a standard capacitor.

Simple "symmetrical" bridges such as these are so named because they exhibit symmetry (mirror-image similarity) from left to right. The two bridge circuits shown above are balanced by adjusting the calibrated reactive component (L_s or C_s). They are a bit simplified from their real-life counterparts, as practical symmetrical bridge circuits often have a calibrated, variable resistor in series or parallel with the reactive component to balance out stray resistance in the unknown component. But, in the hypothetical world of perfect components, these simple bridge circuits do just fine to illustrate the basic concept.

An example of a little extra complexity added to compensate for real-world effects can be found in the so-called *Wien bridge*, which uses a parallel capacitor-resistor standard impedance to balance out an unknown series capacitor-resistor combination. (Figure below) All capacitors have some amount of internal resistance, be it literal or equivalent (in the form of dielectric heating losses) which tend to spoil their otherwise perfectly reactive natures. This internal resistance may be of interest to measure, and so the Wien bridge attempts to do so by providing a balancing impedance that isn't "pure" either:



Wein Bridge measures both capacitive C_x and resistive R_x components of “real” capacitor.
Being that there are two standard components to be adjusted (a resistor and a capacitor) this bridge will take a little more time to balance than the others we've seen so far. The combined effect of R_s and C_s is to alter the magnitude and phase angle until the bridge achieves a condition of balance. Once that balance is achieved, the settings of R_s and C_s can be read from their calibrated knobs, the parallel impedance of the two determined mathematically, and the unknown capacitance and resistance determined mathematically from the balance equation ($Z_1/Z_2 = Z_3/Z_4$).
It is assumed in the operation of the Wien bridge that the standard capacitor has negligible internal resistance, or at least that resistance is already known so that it can be factored into the balance equation. Wien bridges are useful for determining the values of “lossy” capacitor designs like electrolytics, where the internal resistance is relatively high. They are also used as frequency meters, because the balance of the bridge is frequency-dependent. When used in this fashion, the capacitors are made fixed (and usually of equal value) and the top two resistors are made variable and are adjusted by means of the same knob.
An interesting variation on this theme is found in the next bridge circuit, used to precisely measure inductances.



Maxwell-Wein bridge measures an inductor in terms of a capacitor standard.

This ingenious bridge circuit is known as the *Maxwell-Wien bridge* (sometimes known plainly as the *Maxwell bridge*), and is used to measure unknown inductances in terms of calibrated resistance and capacitance. (Figure above) Calibration-grade inductors are more difficult to manufacture than capacitors of similar precision, and so the use of a simple "symmetrical" inductance bridge is not always practical. Because the phase shifts of inductors and capacitors are exactly opposite each other, a capacitive impedance can balance out an inductive impedance if they are located in opposite legs of a bridge, as they are here.

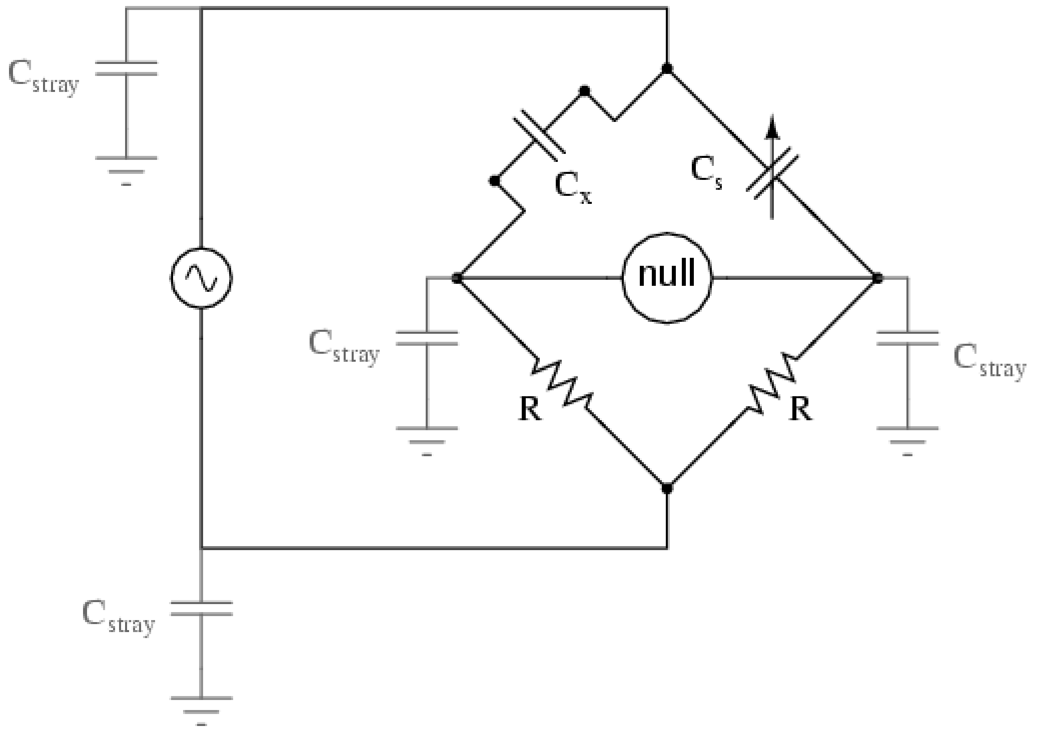
Another advantage of using a Maxwell bridge to measure inductance rather than a symmetrical inductance bridge is the elimination of measurement error due to mutual inductance between two inductors. Magnetic fields can be difficult to shield, and even a small amount of coupling between coils in a bridge can introduce substantial errors in certain conditions. With no second inductor to react with in the Maxwell bridge, this problem is eliminated.

For easiest operation, the standard capacitor (C_s) and the resistor in parallel with it (R_s) are made variable, and both must be adjusted to achieve balance. However, the bridge can be made to work if the capacitor is fixed (non-variable) and more than one resistor made variable (at least the resistor in parallel with the capacitor, and one of the other two). However, in the latter configuration it takes more trial-and-error adjustment to achieve balance, as the different variable resistors interact in balancing magnitude and phase.

Unlike the plain Wien bridge, the balance of the Maxwell-Wien bridge is independent of source frequency, and in some cases this bridge can be made to balance in the presence of mixed frequencies from the AC voltage source, the limiting factor being the inductor's stability over a wide frequency range.

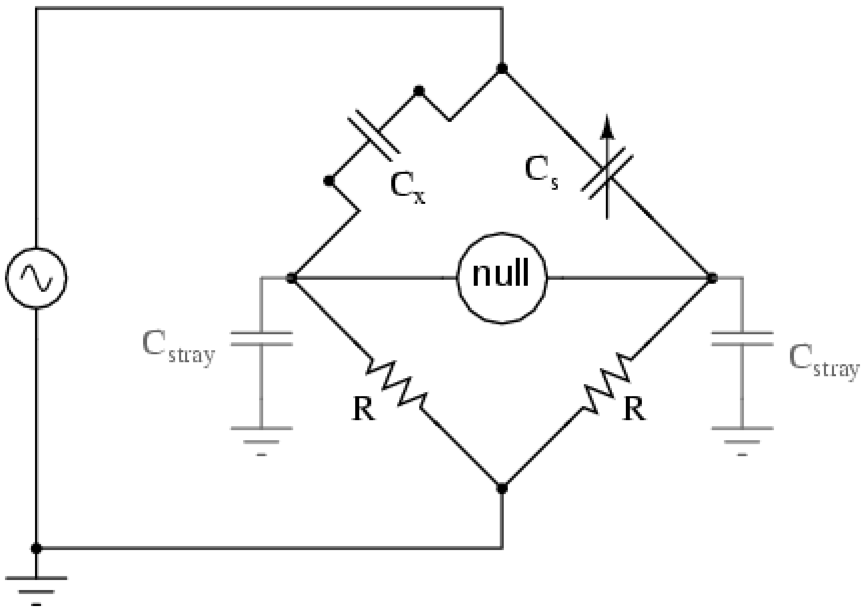
There are more variations beyond these designs, but a full discussion is not warranted here. General-purpose impedance bridge circuits are manufactured which can be switched into more than one configuration for maximum flexibility of use.

A potential problem in sensitive AC bridge circuits is that of stray capacitance between either end of the null detector unit and ground (earth) potential. Because capacitances can "conduct" alternating current by charging and discharging, they form stray current paths to the AC voltage source which may affect bridge balance: (Figure below)



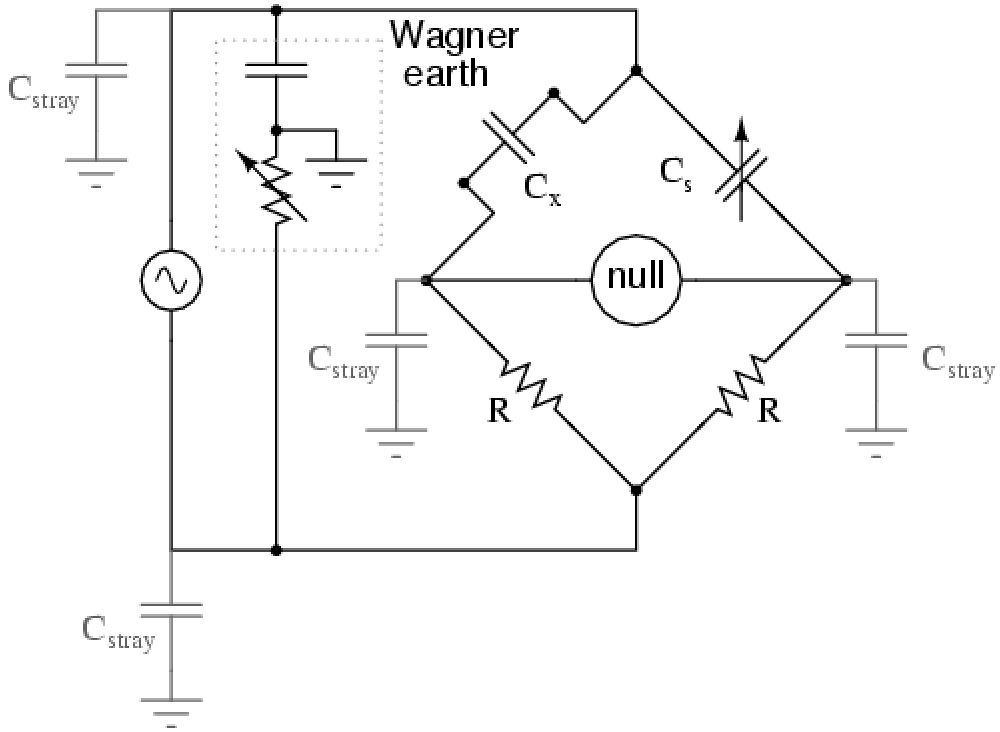
Stray capacitance to ground may introduce errors into the bridge.

While reed-type meters are imprecise, their operational principle is not. In lieu of mechanical resonance, we may substitute electrical resonance and design a frequency meter using an inductor and capacitor in the form of a tank circuit (parallel inductor and capacitor). One or both components are made adjustable, and a meter is placed in the circuit to indicate maximum amplitude of voltage across the two components. The adjustment knob(s) are calibrated to show resonant frequency for any given setting, and the frequency is read from them after the device has been adjusted for maximum indication on the meter. Essentially, this is a tunable filter circuit which is adjusted and then read in a manner similar to a bridge circuit (which must be balanced for a "null" condition and then read). The problem is worsened if the AC voltage source is firmly grounded at one end, the total stray impedance for leakage currents made far less and any leakage currents through these stray capacitances made greater as a result: (Figure below)



Stray capacitance errors are more severe if one side of the AC supply is grounded.

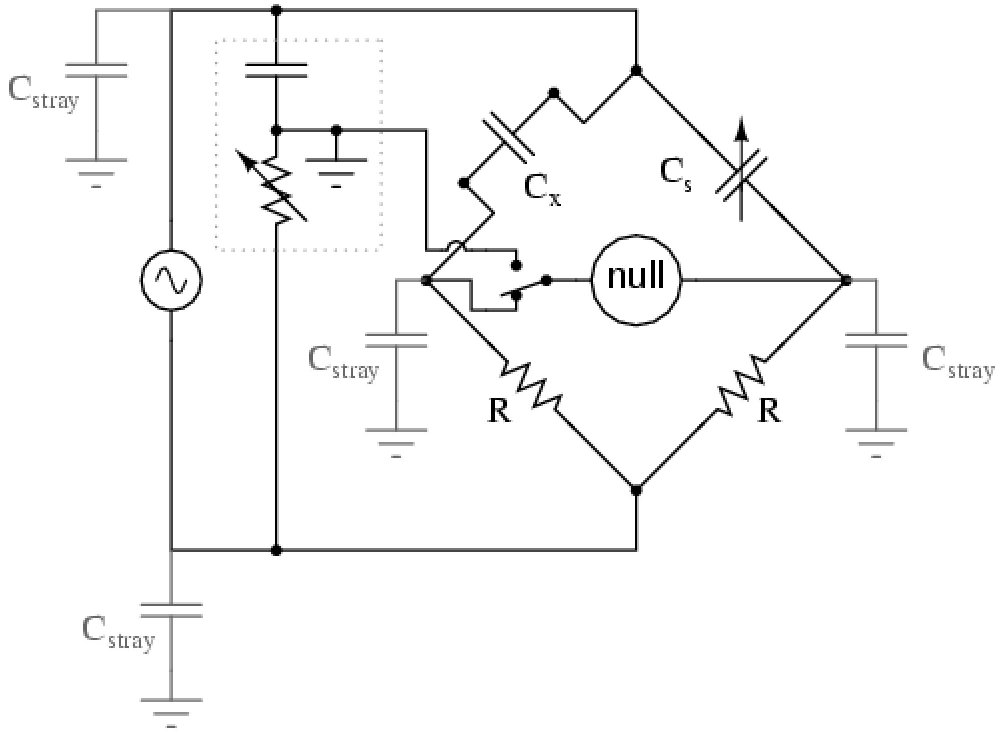
One way of greatly reducing this effect is to keep the null detector at ground potential, so there will be no AC voltage between it and the ground, and thus no current through stray capacitances. However, directly connecting the null detector to ground is not an option, as it would create a *direct* current path for stray currents, which would be worse than any capacitive path. Instead, a special voltage divider circuit called a *Wagner ground* or *Wagner earth* may be used to maintain the null detector at ground potential without the need for a direct connection to the null detector. (Figure below)



Wagner ground for AC supply minimizes the effects of stray capacitance to ground on the bridge.

The Wagner earth circuit is nothing more than a voltage divider, designed to have the voltage ratio and phase shift as each side of the bridge. Because the midpoint of the Wagner divider is directly grounded, any other divider circuit (including either side of the bridge) having the same voltage proportions and phases as the Wagner divider, and powered by the same AC voltage source, will be at ground potential as well. Thus, the Wagner earth divider forces the null detector to be at ground potential, without a direct connection between the detector and ground.

There is often a provision made in the null detector connection to confirm proper setting of the Wagner earth divider circuit: a two-position switch, (Figure below) so that one end of the null detector may be connected to either the bridge or the Wagner earth. When the null detector registers zero signal in both switch positions, the bridge is not only guaranteed to be balanced, but the null detector is also guaranteed to be at zero potential with respect to ground, thus eliminating any errors due to leakage currents through stray detector-to-ground capacitances:



Switch-up position allows adjustment of the Wagner ground.

REVIEW:

AC bridge circuits work on the same basic principle as DC bridge circuits: that a balanced ratio of impedances (rather than resistances) will result in a “balanced” condition as indicated by the null-detector device.

Null detectors for AC bridges may be sensitive electromechanical meter movements, oscilloscopes (CRT's), headphones (amplified or unamplified), or any other device capable of registering very small AC voltage levels. Like DC null detectors, its only required point of calibration accuracy is at zero.

AC bridge circuits can be of the “symmetrical” type where an unknown impedance is balanced by a standard impedance of similar type on the same side (top or bottom) of the bridge. Or, they can be “nonsymmetrical,” using parallel impedances to balance series impedances, or even capacitances balancing out inductances.

AC bridge circuits often have more than one adjustment, since both impedance magnitude *and* phase angle must be properly matched to balance.

Some impedance bridge circuits are frequency-sensitive while others are not. The frequency-sensitive types may be used as frequency measurement devices if all component values are accurately known.
A *Wagner earth* or *Wagner ground* is a voltage divider circuit added to AC bridges to help reduce errors due to stray capacitance coupling the null detector to ground.

Signal generator:

A **signal generator**, also known variously as a **test signal generator**, **function generator**, **tone generator**, **arbitrary waveform generator**, or **frequency generator** is an electronic device that generates repeating electronic signals (in either the analog or digital domains). They are generally used in designing, testing, troubleshooting, and repairing electronic or electroacoustic devices; though they often have artistic uses as well.

There are many different types of signal generators, with different purposes and applications (and at varying levels of expense); in general, no device is suitable for all possible applications.

Traditionally, signal generators have been embedded hardware units, but since the age of multimedia-PCs, flexible, programmable software tone generators have also been available.

Function generator:

A **function generator** is a piece of electronic test equipment or software used to generate electrical waveforms. These waveforms can be either repetitive, or single-shot in which case some kind of triggering source is required (internal or external).

function generator

Analog function generators usually generate a triangle waveform as the basis for all of its other outputs. The triangle is generated by repeatedly charging and discharging a capacitor from a constant current source. This produces a linearly ascending or descending voltage ramp. As the output voltage reaches upper and lower limits, the charging and discharging is reversed using a comparator, producing the linear triangle wave. By varying the current and the size of the capacitor, different frequencies may be obtained.

A 50% duty cycle square wave is easily obtained by noting whether the capacitor is being charged or discharged, which is reflected in the current switching comparator's output. Most function generators also contain a non-linear diode shaping circuit that can convert the triangle wave into a reasonably accurate sine wave. It does so by rounding off the hard corners of the triangle wave in a process similar to clipping in audio systems.

The type of output connector from the device depends on the frequency range of the generator. A typical function generator can provide frequencies up to 20 MHz and uses a BNC connector, usually requiring a 50 or 75 ohm termination. Specialised RF generators are capable of gigahertz frequencies and typically use N-type output connectors.

Function generators, like most signal generators, may also contain an attenuator, various means of modulating the output waveform, and often the ability to automatically and repetitively "sweep" the frequency of the output waveform (by means of a voltage-controlled oscillator) between two operator-determined limits. This capability makes it very easy to evaluate the frequency response of a given electronic circuit.

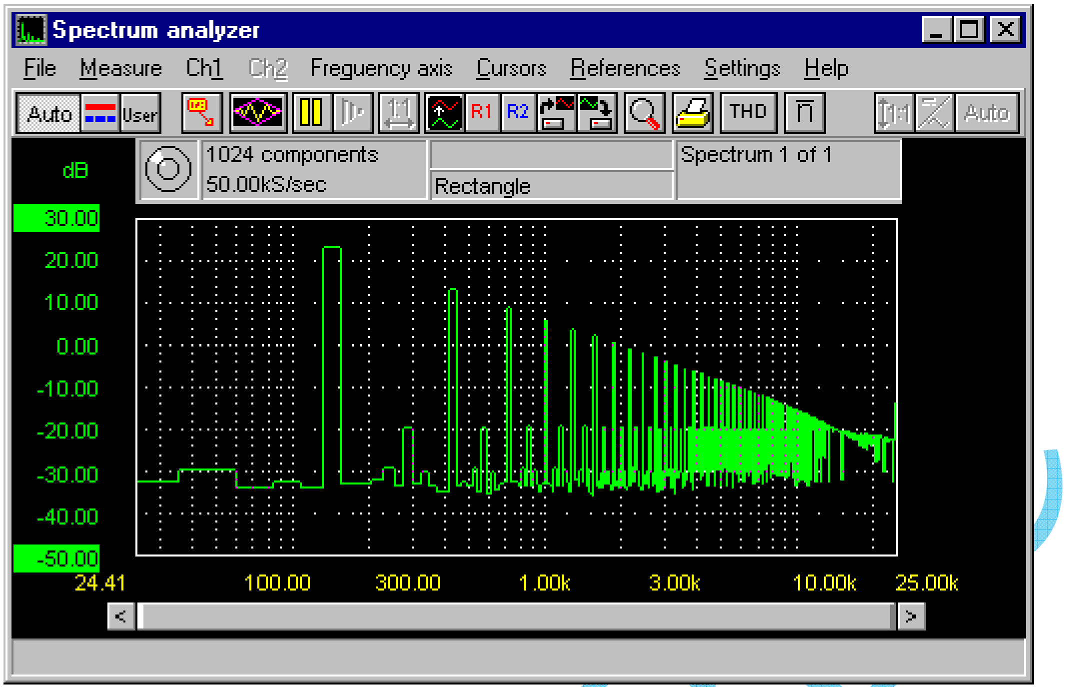
Some function generators can also generate white or pink noise.

More advanced function generators use Direct Digital Synthesis (DDS) to generate waveforms. Arbitrary waveform generators use DDS to generate any waveform that can be described by a table of amplitude values.

Spectrum analyzer:

The common way to examine electrical signals is in the time domain, using an oscilloscope. The time domain is used to determine amplitude, time and phase information, which is necessary to describe the behaviour of an electrical system.

When the time domain does not give the information you want from a signal, you can measure the frequency domain. For that purpose a spectrum analyzer is available. This instrument measures the frequency spectrum of a signal.



1024 components
50.00kS/sec
Rectangle
Spectrum 1 of 1

Key features of the spectrum analyzer are:
determine a spectrum of up to 16K long records
simultaneous use with the oscilloscope: you can see the signal both in the time domain and the frequency domain
intuitive adjustment of the horizontal and vertical offsets and gains using advanced graphical controls
five different window functions to reduce discontinuity errors in the Fast Fourier Transform
high precision Discrete Fourier Transform to eliminate discontinuity errors
linear and logarithmic vertical axes
frequency axis linear, logarithmic, in octaves and in thirds of octaves
interpolation to improve graph
averaging of measurements to reduce the influence of noise
measuring maximum values
Total Harmonic Distortion calculation up to 100 harmonics
Spectrum analyzer controlled frequency sweep on instruments with function generator
extended cursor measurements
all settings accessible through popup menus
storing and recalling reference signals
easy to use speedbutton toolbar for easy access of all functions
comment tags to mark specific events in the signal
colour hardcopy supported
explanatory comments with all controls

Logic analyzer:

A **logic analyzer** is an electronic instrument that displays signals in a digital circuit that are too fast to be observed and presents it to a user so that the user can more easily check correct operation of the digital system. They are typically used for capturing data in systems that have too many channels to be examined with an oscilloscope. Software running on the logic analyzer can convert the captured data into timing diagrams, protocol decodes, state machine traces, assembly language, or correlate assembly with source-level software.

Current analyzers are either mainframes, which consist of a chassis containing the display, controls, control computer, and multiple slots into which the actual data capturing hardware is installed, or standalone units which integrate everything into a single package, with options installed at the factory.

Operation:

A logic analyzer can trigger on a complicated sequence of digital events, and then capture a large amount of digital data from the system under test (SUT). The best logic analyzers behave like software debuggers by showing the flow of the computer program and decoding protocols to show messages and violations.

When logic analyzers first came into use, it was common to attach several hundred "clips" to a digital system. Later, specialized connectors came into use. The evolution of logic analyzer probe has led to a common footprint that multiple vendors support, which provides added freedom to end users. Introduced in April, 2002, connectorless technology (identified by several vendor specific trade names: Compression Probing; Soft Touch; D-Max) has become popular. These probes provide a durable, reliable mechanical and electrical connection between the probe and the circuit board with less than 0.5pF to 0.7 pF loading per signal.

Once the probes are connected, the user programs the analyzer with the names of each signal, and can group several signals into groups for easier manipulation. Next, a capture mode is chosen, either *timing* mode, where the input signals are sampled at regular intervals based on an internal or external clock source, or *state* mode, where one or more of the signals are defined as "clocks," and data is taken on the rising or falling edges of these clocks, optionally using other signals to qualify these clocks.

After the mode is chosen, a *trigger condition* must be set. A trigger condition can range from simple (such as triggering on a rising or falling edge of a single signal), to the very complex (such as configuring the analyzer to decode the higher levels of the TCP/IP stack and triggering on a certain HTTP packet).

At this point, the user sets the analyzer to "run" mode, either triggering once, or repeatedly triggering.

Once the data is captured, it can be displayed several ways, from the simple (showing waveforms or state listings) to the complex (showing decoded Ethernet protocol traffic). The analyzer can also operate in a "compare" mode, where it compares each captured data set to a previously recorded data set, and stopping triggering when this data set is either matched or not. This is useful for long-term empirical testing. Recent analyzers can even be set to email a copy of the test data to the engineer on a successful trigger.

Pulse generator: Pulse generators can either be internal circuits or pieces of electronic test equipment used to generate pulses.

Simple pulse generators usually allow control of the pulse repetition rate (frequency), pulse width, delay with respect to an internal or external trigger and the high- and low-voltage levels of the pulses. More-sophisticated pulse generators may allow control over the rise time and fall time of the pulses. Pulse generators may use digital techniques, analog techniques, or a combination of both techniques to form the output pulses. For example, the pulse repetition rate and duration may be digitally controlled but the pulse amplitude and rise and fall times may be determined by analog circuitry in the output stage of the pulse generator. With correct adjustment, pulse generators can also produce a 50% duty cycle square wave. Pulse generators are generally single-channel providing one frequency, delay, width and output. To produce multiple pulses, these simple pulse generators would have to be ganged in series or in parallel.

A new family of pulse generators can produce multiple-channels of independent widths and delays and independent outputs and polarities. Often called digital delay/pulse generators, the newest designs even offer differing repetition rates with each channel. These digital delay generators are useful in synchronizing, delaying, gating and triggering multiple devices usually with respect to one event.

Pulse generators are available for generating output pulses having widths (durations) ranging from minutes down to under 1 picosecond. In general, generators for pulses with widths over a few microseconds employ digital counters for timing these pulses, while widths between approximately 1 nanosecond and several microseconds are typically generated by analog techniques such as RC (resistor-capacitor) networks or switched delay lines. Pulse generators capable of generating pulses with widths under approximately 100 picoseconds are often termed "microwave pulsers", and typically generate these ultra-short pulses using Step recovery diode (SRD) or Nonlinear Transmission Line (NLTL) methods (see, for example, [1]). Step Recovery Diode pulse generators are inexpensive but typically require several volts of input drive level and have a moderately high level of random jitter (usually undesirable variation in the time at which successive pulses occur). NLTL-based pulse generators generally have lower jitter, but are more complex to manufacture, and are not suited for integration in low-cost monolithic ICs. A new class of microwave pulse generation architecture, the RACE (Rapid Automatic Cascode Exchange) pulse generation circuit[2], is implemented using low-cost monolithic IC technology and can produce pulses as short as 1 picosecond, and with a repetition rates exceeding 30 billion pulses per second. These pulsers are typically used in military communications applications, and low-power microwave transceiver ICs.

Pulse generators are generally voltage sources, with true current pulse generators being available only from a few suppliers.

These pulses can then be injected into a device under test and used as a stimulus or clock signal or analyzed as they progress through the device, confirming the proper operation of the device or pinpointing a fault in the device. Pulse generators are also used to drive devices such as switches, lasers and optical components, modulators, intensifiers as well as resistive loads.The output of a pulse generator may also be used as the modulation signal for a signal generator.

Voltmeter:

A **voltmeter** is an instrument used for measuring the electrical potential difference between two points in an electric circuit. Analog voltmeters move a pointer across a scale in proportion to the voltage of the circuit; digital voltmeters give a numerical display of voltage by use of an Analog to digital converter.

Voltmeters are made in a wide range of styles. Instruments permanently mounted in a panel are used to monitor generators or other fixed apparatus. Small portable instruments, usually equipped with facilities to also measure current and resistance in the form of a multimeter, are standard test instruments used in electrical and electronics work. Any measurement that can be converted to a voltage can be displayed on a meter that is suitably calibrated; for example, pressure, temperature, flow or level ina chemical process plant.

General purpose analog voltmeters may have an accuracy of a few per cent of full scale, and are used with voltages from a fraction of a volt to several thousand volts. Digital meters can be made with high accuracy, typically better than 1%. Specially calibrated test instruments have higher accuracies, with laboratory instruments capable of measuring to accuracies of a few parts per million. Meters using amplifiers can measure tiny voltages of microvolts or less.

Digital voltmeters:

The first *digital* voltmeter was invented and produced by Andrew Kay of Non-Linear Systems (and later founder of Kaypro) in 1954.

Digital voltmeters usually employ an electronic circuit that acts as an integrator, linearly ramping output voltage when input voltage is constant (this can be easily realized with an opamp). The dual-slope integrator method applies a known reference voltage to the integrator for a fixed time to ramp the integrator's output voltage up, then the unknown voltage is applied to ramp it back down, and the time to ramp output voltage down to zero is recorded (realized in an ADC implementation). The unknown voltage being measured is the product of the voltage reference and the ramp-up time divided by the ramp-down time. The voltage reference must remain constant during the ramp-up time, which may be difficult due to supply voltage and temperature variations.

Digital voltmeters necessarily have input amplifiers and like vacuum tube voltmeters generally have a constant input resistance of 10 megohms regardless of set measurement range.

Potentiometer:

An important laboratory technique is measurement of voltage with a potentiometer in the null-balance method. The potentiometer's voltage divider is changed at the wiper until the null detector shows zero voltage between the two circuits.

where

V_t: Voltage across test points

V_k: Known voltage

R_e: Potentiometer resistance from one end terminal to the other end terminal

R_w: Potentiometer resistance from wiper to end terminal

There are many implementations for null detectors, including moving-coil galvanometers, nanovolt-sensitive integrated circuits, and simple audio circuits that click to indicate voltage difference. The null detector need only be sensitive to small voltage differences but does not need to be linear or accurate. The voltage divider can be made with high uniformity and accuracy, with calculable sources of error. While the method was originally used with manually-adjusted potentiometers, automatic and recording analog instruments are commonly made with the same principle of operation.

Multimeter

A **multimeter** or a **multitester**, also known as a **volt/ohm meter** or **VOM**, is an electronic measuring instrument that combines several functions in one unit. A standard multimeter may include features such as the ability to measure voltage, current and resistance. There are two categories of multimeters, **analog multimeters** and **digital multimeters** (often abbreviated **DMM**.)

A multimeter can be a hand-held device useful for basic fault finding and field service work or a bench instrument which can measure to a very high degree of accuracy. They can be used to troubleshoot electrical problems in a wide array of industrial and household devices such as batteries, motor controls, appliances, power supplies, and wiring systems.

Analog multimeter

A digital multimeter

Quantities measured:

Contemporary multimeters can measure many quantities. The common ones are:

Voltage in volts.
Current in amperes.
Resistance in ohms.

Additionally, multimeters may also measure:

Capacitance in farads.
Frequency in hertz
Duty cycle as a percentage.
Temperature in degrees Celsius or Fahrenheit.
Conductance in siemens.
Inductance in henrys
Audio signal levels in decibels.

Digital multimeters may also include circuits for:

Continuity that beeps when a circuit conducts.
Diodes and Transistors
Digital Multimeters (DMM):

Modern multimeters are often digital due their accuracy, durability and extra features.

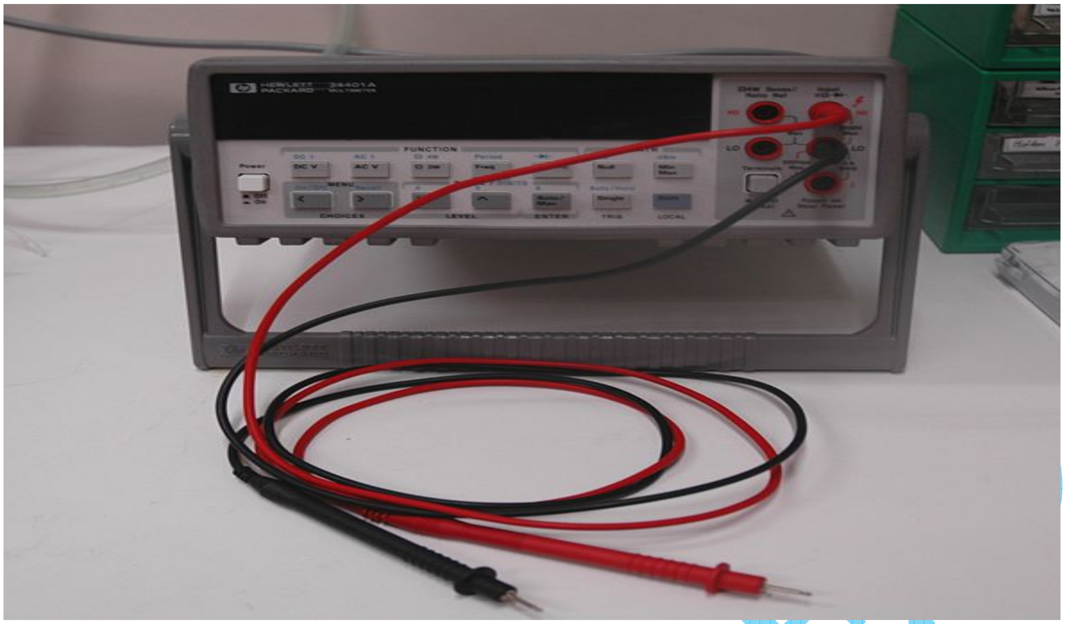
In a DMM the signal under test is converted to a voltage and an amplifier with an electronically controlled gain preconditions the signal.

A DMM displays the quantity measured as a number, which prevents parallax errors.

The inclusion of solid state electronics, from a control circuit to small embedded computers, has provided a wealth of convenience features in modern digital meters. Commonly available measurement enhancements include:

Auto-ranging, which selects the correct range for the quantity under test so that the most significant digits are shown. For example, a four-digit multimeter would automatically select an appropriate range to display 1.234 instead of 0.012, or overloading. Auto-ranging meters usually include a facility to 'freeze' the meter to a particular range, because a measurement that causes frequent range changes is distracting to the user.
Auto-polarity for direct-current readings, shows if the applied voltage is positive (agrees with meter lead labels) or negative (opposite polarity to meter leads).
Sample and hold, which will latch the most recent reading for examination after the instrument is removed from the circuit under test.
Current-limited tests for voltage drop across semiconductor junctions. While not a replacement for a transistor tester, this facilitates testing diodes and a variety of transistor types.
A **graphic representation** of the quantity under test, as a bar graph. This makes go/no-go testing easy, and also allows spotting of fast-moving trends.
A low-bandwidth **oscilloscope**.
Automotive circuit testers, including tests for automotive timing and dwell signals.[7]
Simple data acquisition features to record maximum and minimum readings over a given period, or to take a number of samples at fixed intervals.[8]

Modern meters may be interfaced with a personal computer by IrDA links, RS-232 connections, USB, or an instrument bus such as IEEE-488. The interface allows the computer to record measurements as they are made. Some DMM's can store measurements and upload them to a computer.



A bench-top multimeter

Resolution:

The resolution of a multimeter is often specified in "digits" of resolution. For example, the term 5½ digits refers to the number of digits displayed on the readout of a multimeter.

By convention, a half digit can display either a zero or a one, while a three-quarters digit can display a numeral higher than a one but not nine. Commonly, a three-quarters digit refers to a maximum count of 3 or 5. The fractional digit is always the most significant digit in the displayed value. A 5½ digit multimeter would have five full digits that display values from 0 to 9 and one half digit that could only display 0 or 1.[1] Such a meter could show positive or negative values from 0 to 199,999. A 3¾ digit meter can display a quantity from 0 to 3,999 or 5,999, depending on the manufacturer.

Resolution of analog multimeters is limited by the width of the scale pointer, vibration of the pointer, parallax observation errors, and the accuracy of printing of scales. Resistance measurements, in particular, are of low precision due to the typical resistance measurement circuit which compresses the scale at the higher resistance values. Mirrored scales and larger meter movements are used to improve resolution; two and a half to three digits equivalent resolution is usual (and may be adequate for the limited precision actually necessary for most measurements).

While a digital display can easily be extended in precision, the extra digits are of no value if not accompanied by care in the design and calibration of the analog portions of the multimeter. Meaningful high-resolution measurements require a good understanding of the instrument specifications, good control of the measurement conditions, and traceability of the calibration of the instrument.

Accuracy:

Digital multimeters generally take measurements with superior accuracy to their analog counterparts. Analog multimeters typically measure with three to five percent accuracy.Standard portable digital multimeters claim to be capable of taking measurements with an accuracy of 0.5% on DC voltage and current scales. Mainstream bench-top multimeters make claim es to have as great accuracy as ±0.01%. Laboratory grade instruments can have accuracies in the parts per million.

Manufacturers can provide calibration services so that new meters may be purchased with a certificate of calibration indicating the meter has been adjusted to standards traceable to the National Institute of Standards and Technology. Such manufacturers usually provide calibration services after sales, as well, so that older equipment may be recertified. Multimeters used for critical measurements may be part of a metrology program to assure calibration.

Sensitivity:

The current load, or how much current is drawn from the circuit being tested may affect a multimeter's accuracy. A small current draw usually will result in more precise measurements. With improper usage or too much current load, a multimeter may be damaged therefore rendering its measurements unreliable and substandard.

Meters with electronic amplifiers in them, such as all digital multimeters and transistorized analog meters, have a standardized input impedance usually considered high enough not to disturb the circuit tested. This is often one million ohms, or ten million ohms. The standard input impedance allows use of external probes to extend the direct-current measuring range up to tens of thousands of volts.

Analog multimeters of the moving pointer type draw current from the circuit under test to deflect the meter pointer. The impedance of the meter varies depending on the basic sensitivity of the meter movement and the range which is selected. For example, a meter with a 20,000 ohms/volt sensitivity will have an input resistance of two million ohms on the 100 volt range (100 V * 20,000 ohms/volt = 2,000,000 ohms). Low-sensitivity meters are useful for general purpose testing especially in power circuits, where source impedances are low compared to the meter impedance. Measurements in signal circuits generally require higher sensitivity so as not to load down the circuit under test with the meter impedance.[3]

The sensitivity of a meter is also a measure of the lowest voltage, current or resistance that can be measured with it. For general-purpose digital multimeters, a full-scale range of several hundred millivolts AC or DC is common, but the minimum full-scale current range may be several hundred milliamps. Since general-purpose mulitmeters have only two-wire resistance measurements, which do not compensate for the effect of the lead wire resistance, measurements below a few tens of ohms will be of low accuracy. The upper end of multimeter measurement ranges varies considerably by manufacturer;

generally measurements over 1000 volts, over 10 amperes, or over 100 megohms would require a specialized test instrument, as would accurate measurement of currents on the order of microamperes or less.

Oscilloscope:

An **oscilloscope** (sometimes abbreviated **CRO**, for cathode-ray oscilloscope, or commonly just **scope** or **O-scope**) is a type of electronic test equipment that allows signal voltages to be viewed, usually as a two-dimensional graph of one or more electrical potential differences (vertical axis) plotted as a function of time or of some other voltage (horizontal axis).

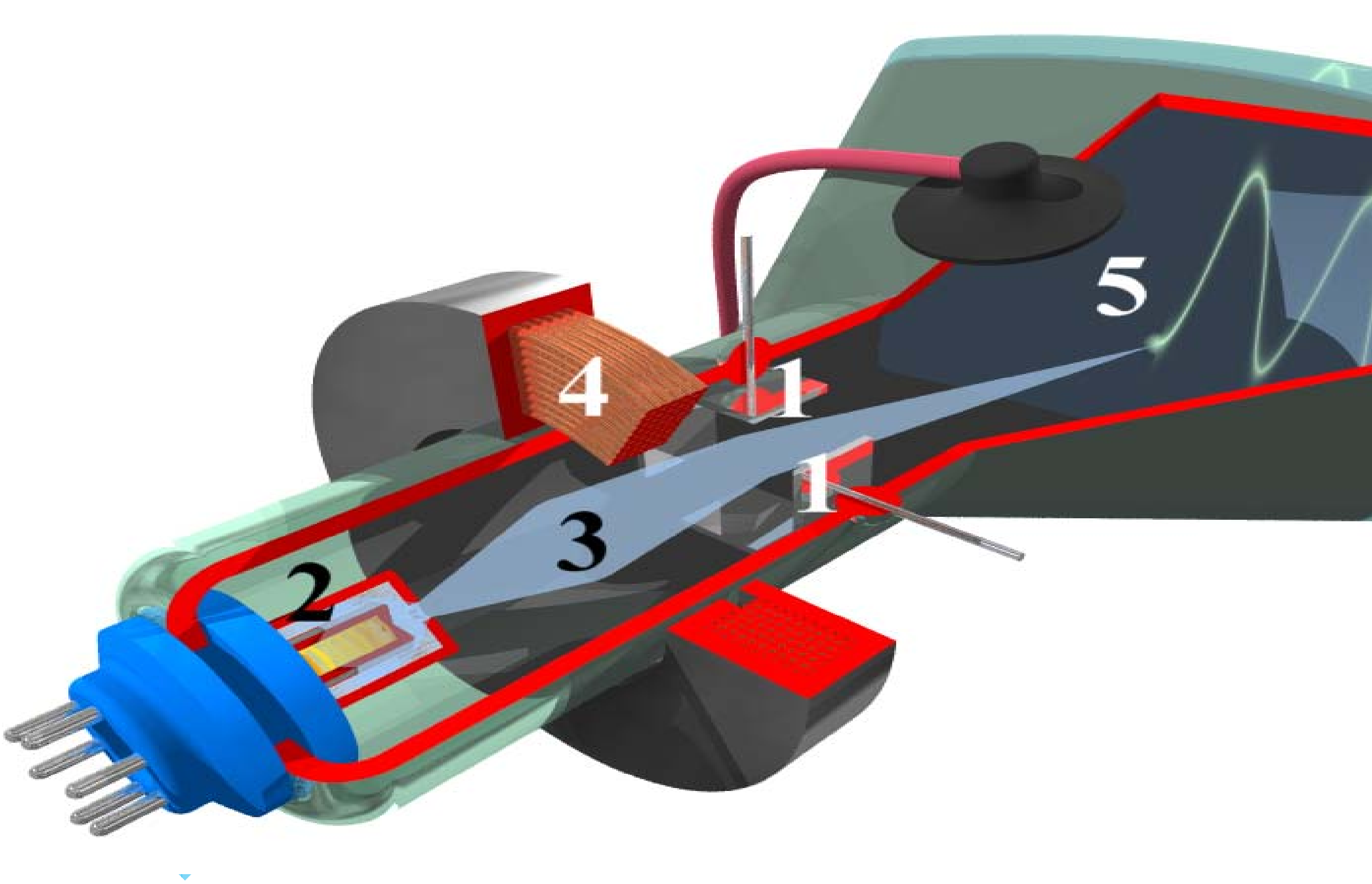


Illustration showing the interior of a cathode-ray tube for use in an oscilloscope. Numbers in the picture indicate: 1. Deflection voltage electrod; 2. Electron gun; 3. Electron beam; 4. Focusing coil; 5. Phosphor-coated inner side of the screen.

A Tektronix model 475A portable analogue oscilloscope, a very typical instrument of the late 1970s. This dual-trace, dual-sweep instrument had a horizontal bandwidth of 250 MHz, a maximum vertical sensitivity of 5 mV per division, and maximum (unmagnified) horizontal sweep speed of 10 ns per division. The vertical controls are on the left with Channel 1 above and Channel 2 below. The horizontal sweep controls are on the right with the Main Trigger above and the Delayed Trigger below. The CRT controls are below the screen. The metal loop to the lower right of the screen provided a calibration signal for voltage and current probes.

The oscilloscope is basically a graph-displaying device - it draws a graph of an electrical signal. In most applications the graph shows how signals change over time: the vertical (Y) axis represents voltage and the horizontal (X) axis represents time. The intensity or brightness of the display is sometimes called the Z axis. (See Figure 1.) This simple graph can tell you many things about a signal. Here are a few:

You can determine the time and voltage values of a signal.
You can calculate the frequency of an oscillating signal.
You can see the "moving parts" of a circuit represented by the signal.
You can tell if a malfunctioning component is distorting the signal.
You can find out how much of a signal is direct current (DC) or alternating current (AC).
You can tell how much of the signal is noise and whether the noise is changing with time.

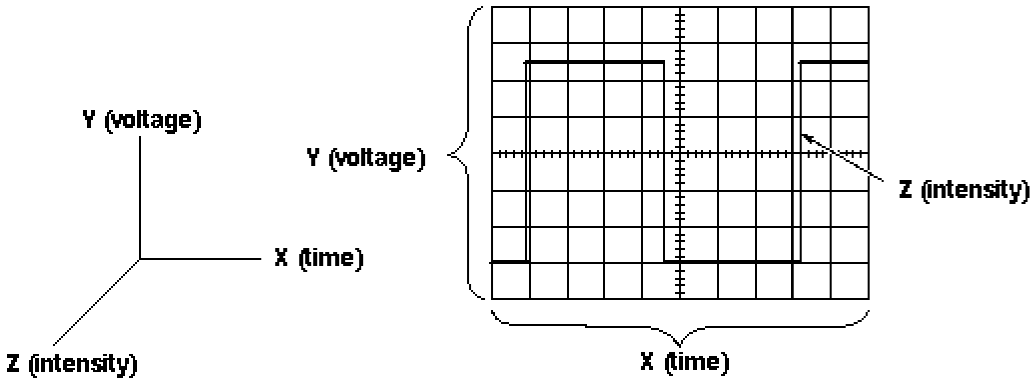


Figure 1: X, Y, and Z Components of a Displayed Waveform

An oscilloscope looks a lot like a small television set, except that it has a grid drawn on its screen and more controls than a television. The front panel of an oscilloscope normally has control sections divided into Vertical, Horizontal, and Trigger sections. There are also display controls and input connectors. See if you can locate these front panel sections in Figures 2 and 3 and on your oscilloscope.

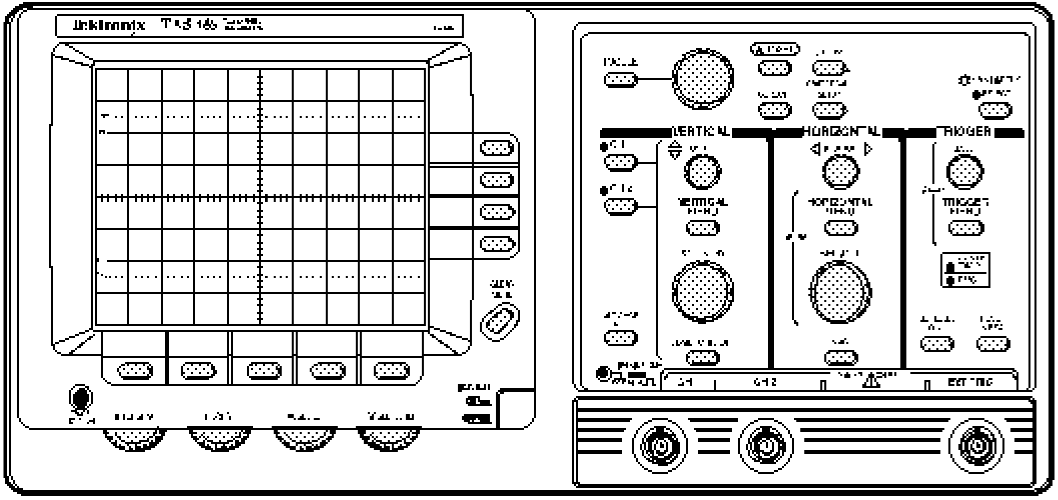


Figure 2: The TAS 465 Analog Oscilloscope Front Panel

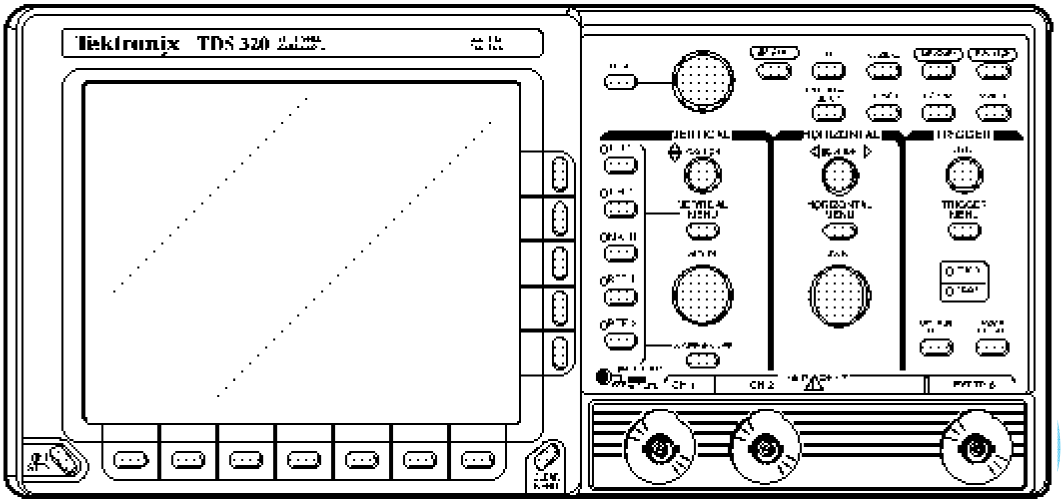
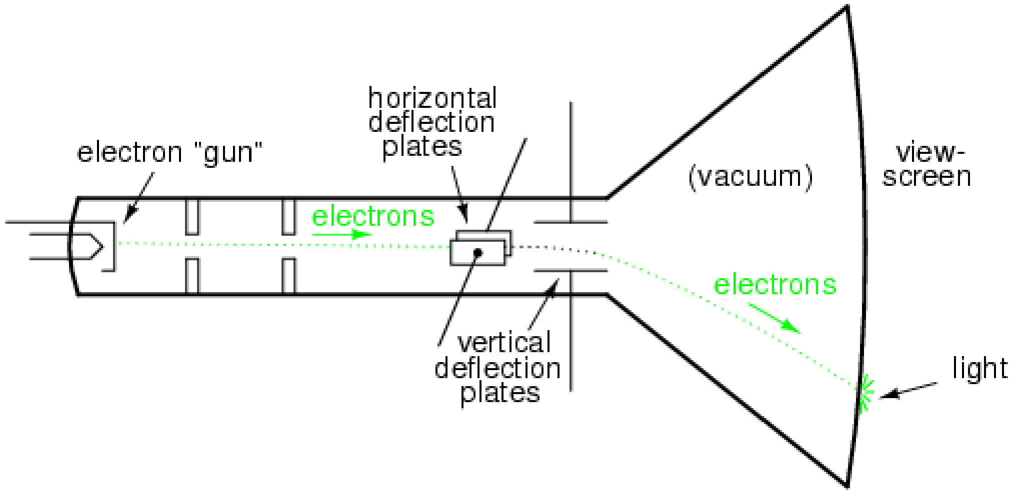


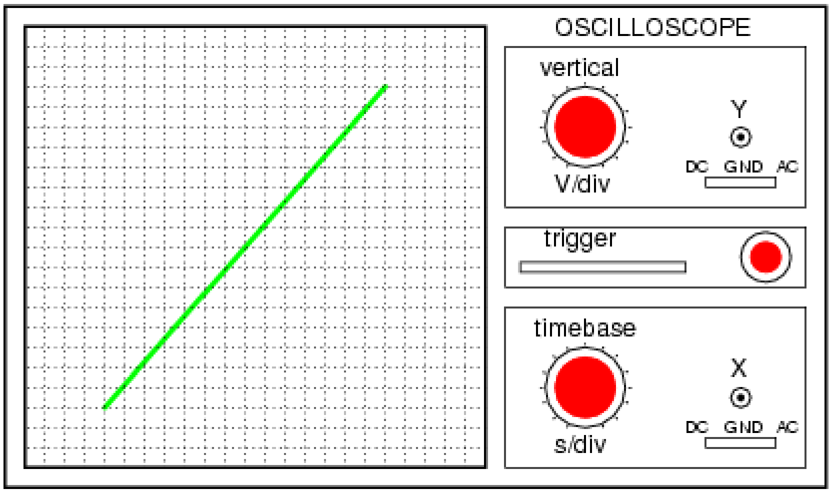
Figure 3: The TDS 320 Digital Oscilloscope Front Panel



Cathode ray tube (CRT) with vertical and horizontal deflection plates.

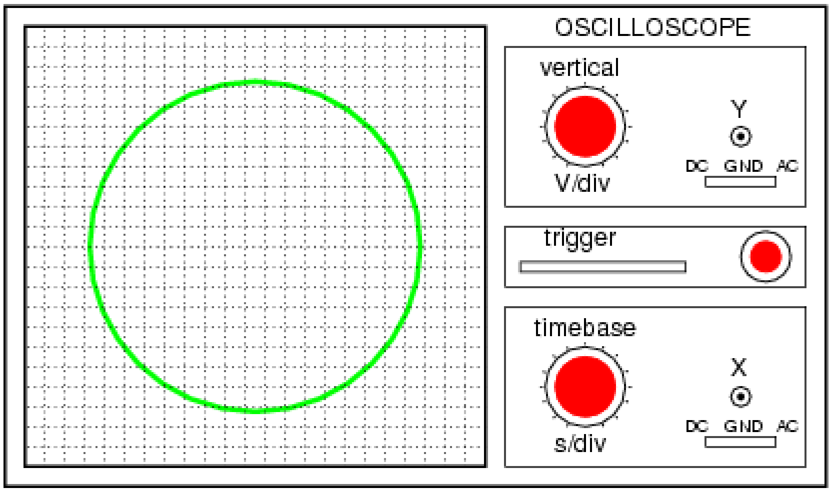
If we allow one AC signal to deflect the beam up and down (connect that AC voltage source to the “vertical” deflection plates) and another AC signal to deflect the beam left and right (using the other pair of deflection plates), patterns will be produced on the screen of the CRT indicative of the *ratio* of these two AC frequencies. These patterns are called *Lissajous figures* and are a common means of comparative frequency measurement in electronics.

If the two frequencies are the same, we will obtain a simple figure on the screen of the CRT, the shape of that figure being dependent upon the phase shift between the two AC signals. Here is a sampling of Lissajous figures for two sine-wave signals of equal frequency, shown as they would appear on the face of an oscilloscope (an AC voltage-measuring instrument using a CRT as its “movement”). The first picture is of the Lissajous figure formed by two AC voltages perfectly in phase with each other: (Figure below)



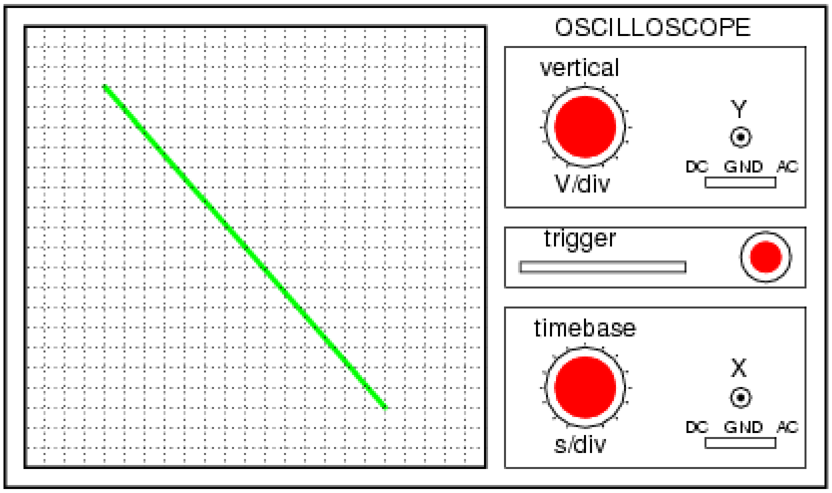
Lissajous figure: same frequency, zero degrees phase shift.

If the two AC voltages are not in phase with each other, a straight line will not be formed. Rather, the Lissajous figure will take on the appearance of an oval, becoming perfectly circular if the phase shift is exactly 90° between the two signals, and if their amplitudes are equal: (Figure below)



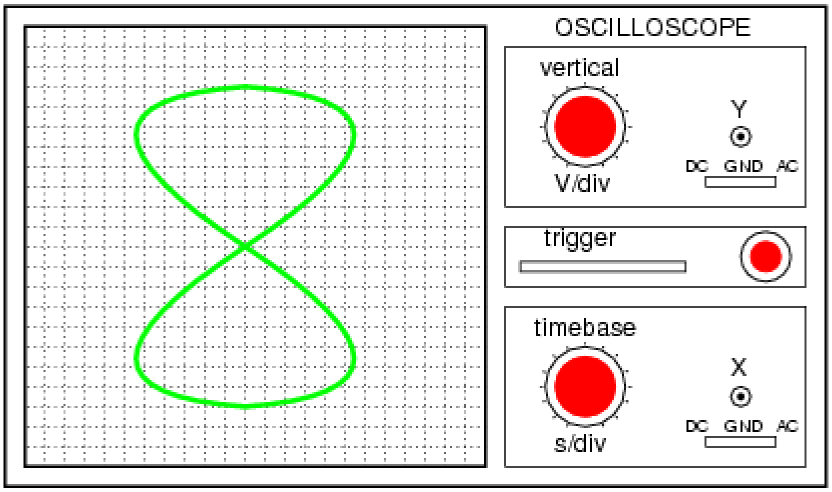
Lissajous figure: same frequency, 90 or 270 degrees phase shift.

Finally, if the two AC signals are directly opposing one another in phase (180° shift), we will end up with a line again, only this time it will be oriented in the opposite direction: (Figure below)



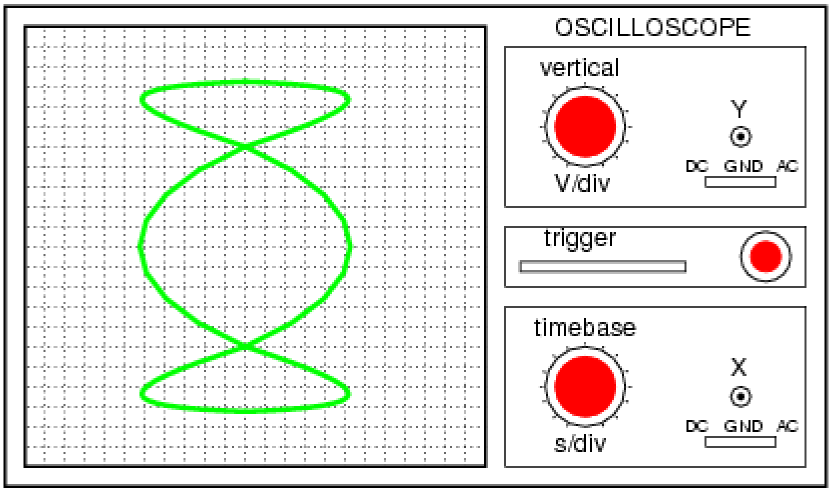
Lissajous figure: same frequency, 180 degrees phase shift.

When we are faced with signal frequencies that are not the same, Lissajous figures get quite a bit more complex. Consider the following examples and their given vertical/horizontal frequency ratios: (Figure below)



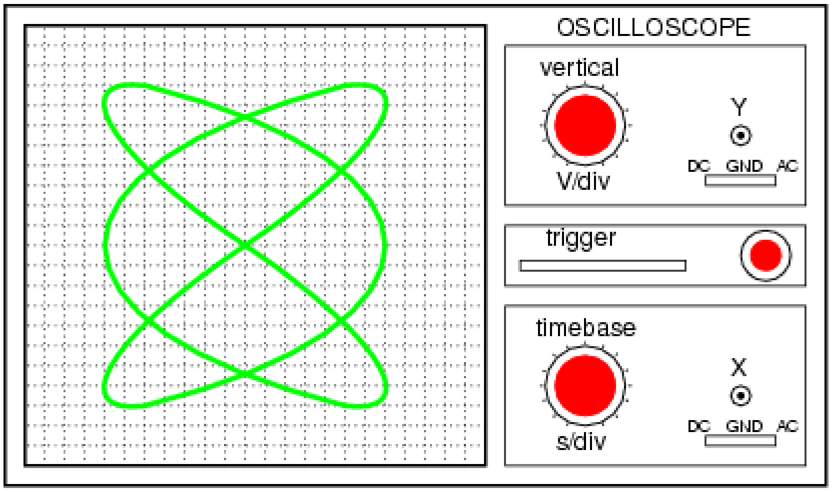
Lissajous figure: Horizontal frequency is twice that of vertical.

The more complex the ratio between horizontal and vertical frequencies, the more complex the Lissajous figure. Consider the following illustration of a 3:1 frequency ratio between horizontal and vertical: (Figure below)



Lissajous figure: Horizontal frequency is three times that of vertical.

. . . and a 3:2 frequency ratio (horizontal = 3, vertical = 2) in Figure below.



Lissajous figure: Horizontal/Vertical frequency ratio is 3:2

Lissajous figure: Horizontal/vertical frequency ratio is 3:2.

In cases where the frequencies of the two AC signals are not exactly a simple ratio of each other (but close), the Lissajous figure will appear to “move,” slowly changing orientation as the phase angle between the two waveforms rolls between 0° and 180°. If the two frequencies are locked in an exact integer ratio between each other, the Lissajous figure will be stable on the viewscreen of the CRT.

The physics of Lissajous figures limits their usefulness as a frequency-comparison technique to cases where the frequency ratios are simple integer values (1:1, 1:2, 1:3, 2:3, 3:4, etc.). Despite this limitation, Lissajous figures are a popular means of frequency comparison wherever an accessible frequency standard (signal generator) exists.

REVIEW:

Some frequency meters work on the principle of mechanical resonance, indicating frequency by relative oscillation among a set of uniquely tuned “reeds” shaken at the measured frequency.

Other frequency meters use electric resonant circuits (LC tank circuits, usually) to indicate frequency. One or both components is made to be adjustable, with an accurately calibrated adjustment knob, and a sensitive meter is read for maximum voltage or current at the point of resonance.

Frequency can be measured in a comparative fashion, as is the case when using a CRT to generate *Lissajous figures*. Reference frequency signals can be made with a high degree of accuracy by oscillator circuits using quartz crystals as resonant devices. For ultra precision, atomic clock signal standards (based on the resonant frequencies of individual atoms) can be used.

What Can You Do With It?

Oscilloscopes are used by everyone from television repair technicians to physicists. They are indispensable for anyone designing or repairing electronic equipment.

The usefulness of an oscilloscope is not limited to the world of electronics. With the proper *transducer*, an oscilloscope can measure all kinds of phenomena. A transducer is a device that creates an electrical signal in response to physical stimuli, such as sound, mechanical stress, pressure, light, or heat. For example, a microphone is a transducer.

An automotive engineer uses an oscilloscope to measure engine vibrations. A medical researcher uses an oscilloscope to measure brain waves. The possibilities are endless.

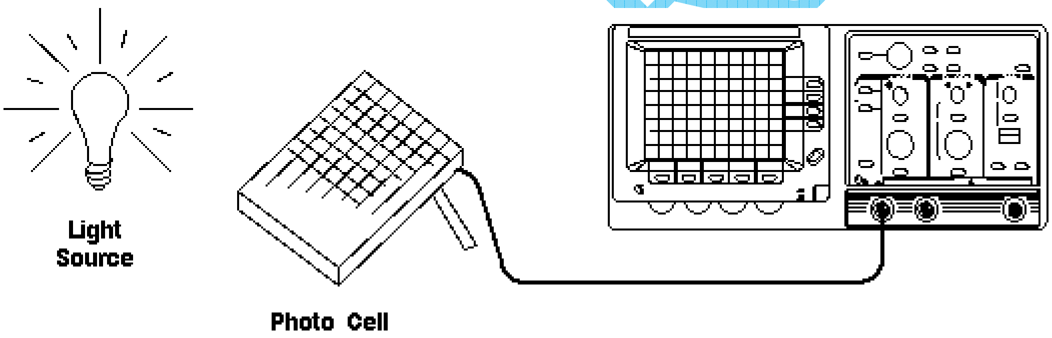


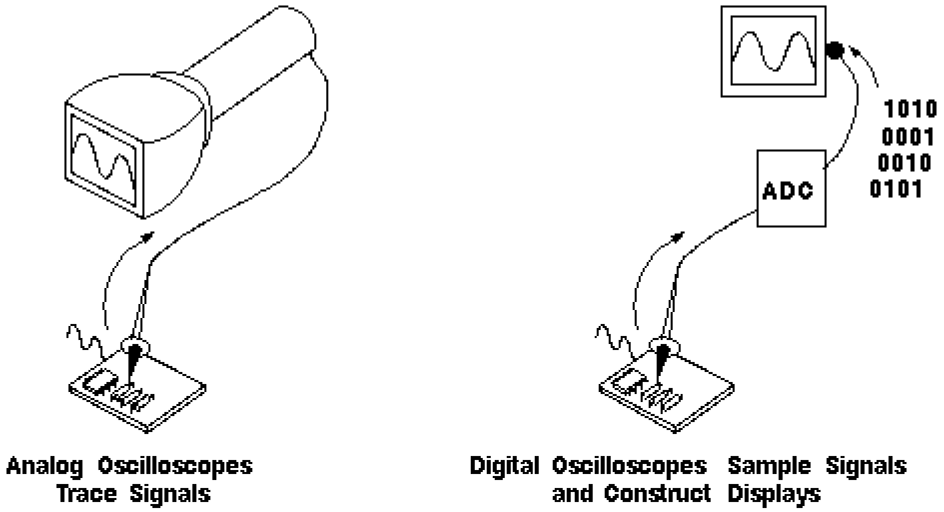
Figure 4: Scientific Data Gathered by an Oscilloscope

Analog and Digital

Electronic equipment can be divided into two types: analog and digital. *Analog* equipment works with continuously variable voltages, while *digital* equipment works with discrete binary numbers that may represent voltage samples. For example, a conventional phonograph turntable is an analog device; a compact disc player is a digital device.

Oscilloscopes also come in analog and digital types. An analog oscilloscope works by directly applying a voltage being measured to an electron beam moving across the oscilloscope screen. The voltage deflects the beam up and down proportionally, tracing the waveform on the screen. This gives an immediate picture of the waveform.

In contrast, a digital oscilloscope samples the waveform and uses an analog-to-digital converter (or ADC) to convert the voltage being measured into digital information. It then uses this digital information to reconstruct the waveform on the screen.



Analog Oscilloscopes Trace Signals

Digital Oscilloscopes Sample Signals and Construct Displays

Figure 5: Digital and Analog Oscilloscopes Display Waveforms

For many applications either an analog or digital oscilloscope will do. However, each type does possess some unique characteristics making it more or less suitable for specific tasks.

People often prefer analog oscilloscopes when it is important to display rapidly varying signals in "real time" (or as they occur).

Digital oscilloscopes allow you to capture and view events that may happen only once. They can process the digital waveform data or send the data to a computer for processing. Also, they can store the digital waveform data for later viewing and printing.

How Does an Oscilloscope Work?

To better understand the oscilloscope controls, you need to know a little more about how oscilloscopes display a signal. Analog oscilloscopes work somewhat differently than digital oscilloscopes. However, several of the internal systems are similar. Analog oscilloscopes are somewhat simpler in concept and are described first, followed by a description of digital oscilloscopes.

Analog Oscilloscopes

When you connect an oscilloscope probe to a circuit, the voltage signal travels through the probe to the vertical system of the oscilloscope. Figure 6 is a simple block diagram that shows how an analog oscilloscope displays a measured signal.

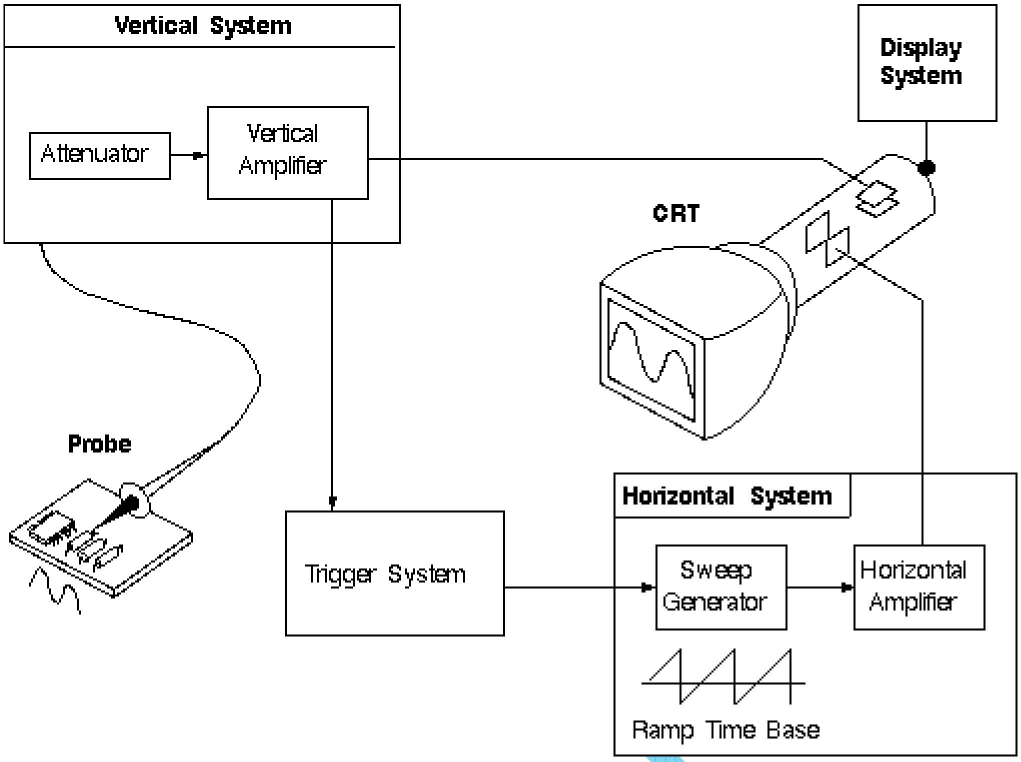


Figure 6: Analog Oscilloscope Block Diagram

Depending on how you set the vertical scale (volts/div control), an *attenuator* reduces the signal voltage or an *amplifier* increases the signal voltage.

Next, the signal travels directly to the vertical deflection plates of the cathode ray tube (CRT). Voltage applied to these deflection plates causes a glowing dot to move. (An electron beam hitting phosphor inside the CRT creates the glowing dot.) A positive voltage causes the dot to move up while a negative voltage causes the dot to move down.

The signal also travels to the trigger system to start or trigger a "horizontal sweep." Horizontal sweep is a term referring to the action of the horizontal system causing the glowing dot to move across the screen. Triggering the horizontal system causes the horizontal time base to move the glowing dot across the screen from left to right within a specific time interval. Many sweeps in rapid sequence cause the movement of the glowing dot to blend into a solid line. At higher speeds, the dot may sweep across the screen up to 500,000 times each second.

Together, the horizontal sweeping action and the vertical deflection action traces a graph of the signal on the screen. The trigger is necessary to stabilize a repeating signal. It ensures that the sweep begins at the same point of a repeating signal, resulting in a clear picture as shown in Figure 7.

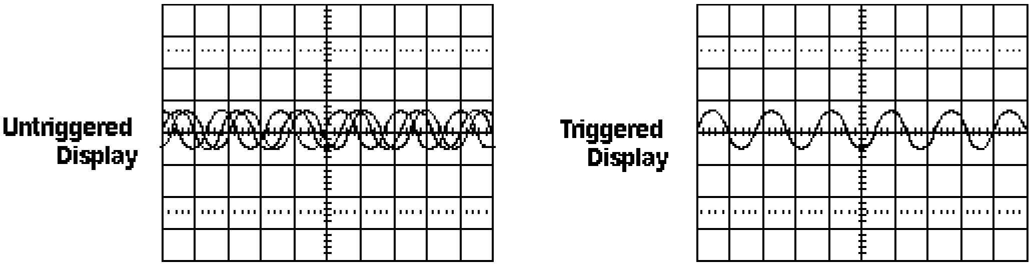


Figure 7: Triggering Stabilizes a Repeating Waveform

In conclusion, to use an analog oscilloscope, you need to adjust three basic settings to accommodate an incoming signal:

The attenuation or amplification of the signal. Use the volts/div control to adjust the amplitude of the signal before it is applied to the vertical deflection plates.
The time base. Use the sec/div control to set the amount of time per division represented horizontally across the screen.
The triggering of the oscilloscope. Use the trigger level to stabilize a repeating signal, as well as triggering on a single event.
Also, adjusting the focus and intensity controls enables you to create a sharp, visible display.

Digital Oscilloscopes

Some of the systems that make up digital oscilloscopes are the same as those in analog oscilloscopes; however, digital oscilloscopes contain additional data processing systems. (See Figure 8.) With the added systems, the digital oscilloscope collects data for the entire waveform and then displays it.

When you attach a digital oscilloscope probe to a circuit, the vertical system adjusts the amplitude of the signal, just as in the analog oscilloscope.

Next, the analog-to-digital converter (ADC) in the acquisition system samples the signal at discrete points in time and converts the signal's voltage at these points to digital values called *sample points*. The horizontal system's sample clock determines how often the ADC takes a sample. The rate at which the clock "ticks" is called the sample rate and is measured in samples per second.

The sample points from the ADC are stored in memory as *waveform points*. More than one sample point may make up one waveform point.

Together, the waveform points make up one waveform *record*. The number of waveform points used to make a waveform record is called the *record length*. The trigger system determines the start and stop points of the record. The display receives these record points after being stored in memory.

Depending on the capabilities of your oscilloscope, additional processing of the sample points may take place, enhancing the display. Pretrigger may be available, allowing you to see events before the trigger point.

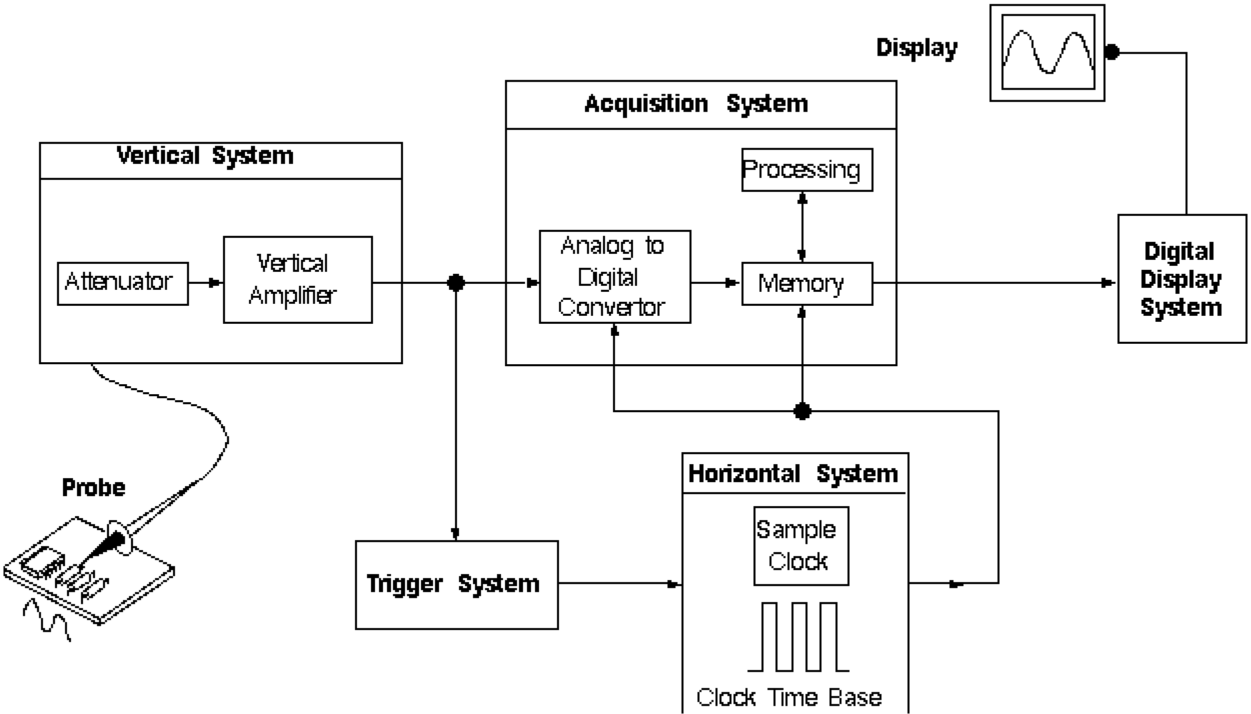


Figure 8: Digital Oscilloscope Block Diagram

Fundamentally, with a digital oscilloscope as with an analog oscilloscope, you need to adjust the vertical, horizontal, and trigger settings to take a measurement.

Sampling Methods

The sampling method tells the digital oscilloscope how to collect sample points. For slowly changing signals, a digital oscilloscope easily collects more than enough sample points to construct an accurate picture. However, for faster signals, (how fast depends on the oscilloscope's maximum sample rate) the oscilloscope cannot collect enough samples. The digital oscilloscope can do two things:

It can collect a few sample points of the signal in a single pass (in *real-time* sampling mode) and then use *interpolation*. Interpolation is a processing technique to estimate what the waveform looks like based on a few points.

It can build a picture of the waveform over time, as long as the signal repeats itself (*equivalent-time* sampling mode).

Real-Time Sampling with Interpolation

Digital oscilloscopes use real-time sampling as the standard sampling method. In real-time sampling, the oscilloscope collects as many samples as it can as the signal occurs. (See Figure 9.) For single-shot or transient signals you must use real time sampling.

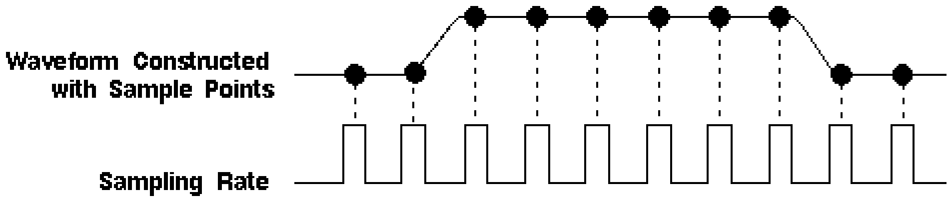


Figure 9: Real-time Sampling

Digital oscilloscopes use interpolation to display signals that are so fast that the oscilloscope can only collect a few sample points. Interpolation "connects the dots."

Linear interpolation simply connects sample points with straight lines. Sine interpolation (or *sin x over x* interpolation) connects sample points with curves. (See Figure 10.) *Sin x over x* interpolation is a mathematical process similar to the "oversampling" used in compact disc players. With sine interpolation, points are calculated to fill in the time between the real samples. Using this process, a signal that is sampled only a few times in each cycle can be accurately displayed or, in the case of the compact disc player, accurately played back.

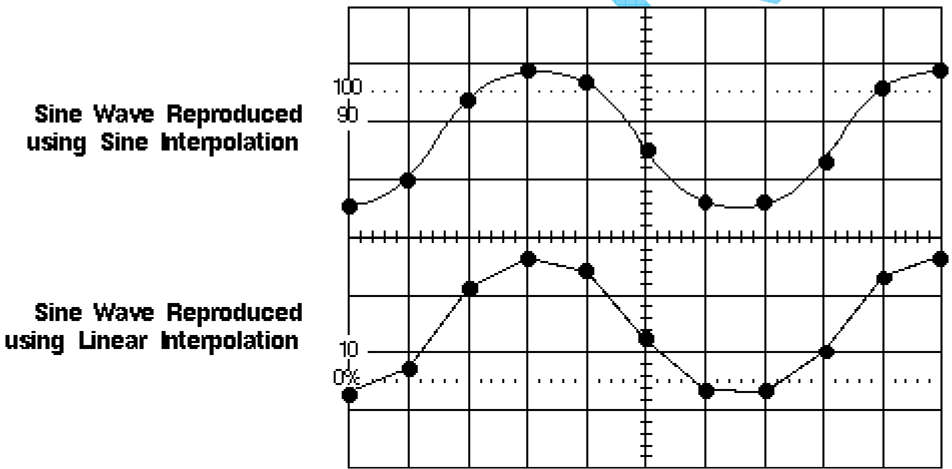


Figure 10: Linear and Sine Interpolation

Equivalent-Time Sampling

Some digital oscilloscopes can use equivalent-time sampling to capture very fast repeating signals. Equivalent-time sampling constructs a picture of a repetitive signal by capturing a little bit of information from each repetition. (See Figure 11.) You see the waveform slowly build up like a string of lights going on one-by-one. With *sequential sampling* the points appear from left to right in sequence; with *random sampling* the points appear randomly along the waveform.

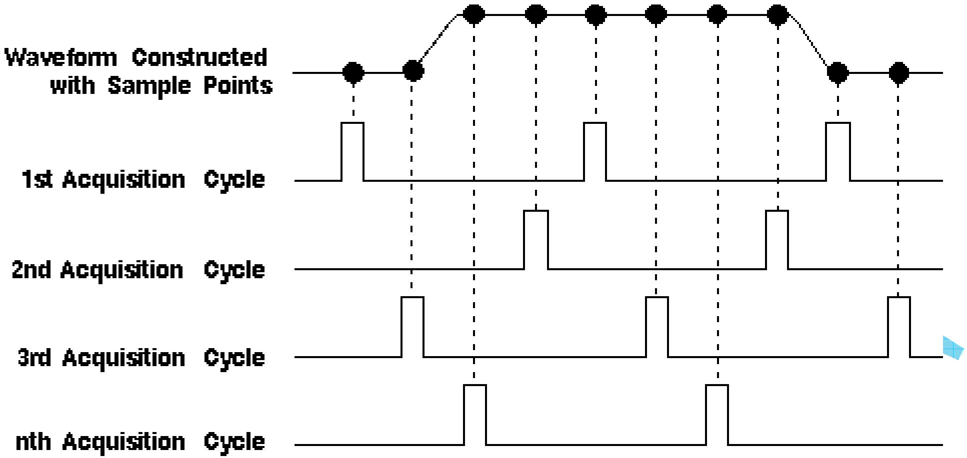


Figure 11: Equivalent-time Sampling

Cathode-ray oscilloscope (CRO):

The earliest and simplest type of oscilloscope consisted of a cathode ray tube, a vertical amplifier, a timebase, a horizontal amplifier and a power supply. These are now called 'analogue' scopes to distinguish them from the 'digital' scopes that became common in the 1990s and 2000s.

Before the introduction of the CRO in its current form, the cathode ray tube had already been in use as a measuring device. The cathode ray tube is an evacuated glass envelope, similar to that in a black-and-white television set, with its flat face covered in a phosphorescent material (the phosphor). The screen is typically less than 20 cm in diameter, much smaller than the one in a television set.

In the neck of the tube is an electron gun, which is a heated metal plate with a wire mesh (the grid) in front of it. A small grid potential is used to block electrons from being accelerated when the electron beam needs to be turned off, as during sweep retrace or when no trigger events occur. A potential difference of at least several hundred volts is applied to make the heated plate (the cathode) negatively charged relative to the deflection plates. For higher bandwidth oscilloscopes where the trace may move more rapidly across the phosphor target, a positive post-deflection acceleration voltage of over 10,000 volts is often used, increasing the energy (speed) of the electrons that strike the phosphor. The kinetic energy of the electrons is converted by the phosphor into visible light at the point of impact. When switched on, a CRT normally displays a single bright dot in the center of the screen, but the dot can be moved about electrostatically or magnetically. The CRT in an oscilloscope uses electrostatic deflection.

Between the electron gun and the screen are two opposed pairs of metal plates called the deflection plates. The vertical amplifier generates a potential difference across one pair of plates, giving rise to a vertical electric field through which the electron beam passes. When the plate potentials are the same, the beam is not deflected.

When the top plate is positive with respect to the bottom plate, the beam is deflected upwards; when the field is reversed, the beam is deflected downwards. The horizontal amplifier does a similar job with the other pair of deflection plates, causing the beam to move left or right. This deflection system is called electrostatic deflection, and is different from the electromagnetic deflection system used in television tubes. In comparison to magnetic deflection, electrostatic deflection can more readily follow random changes in potential, but is limited to small deflection angles.

The timebase is an electronic circuit that generates a ramp voltage. This is a voltage that changes continuously and linearly with time. When it reaches a predefined value the ramp is reset, with the voltage reestablishing its initial value. When a trigger event is recognized the reset is released, allowing the ramp to increase again. The timebase voltage usually drives the horizontal amplifier. Its effect is to sweep the electron beam at constant speed from left to right across the screen, then quickly return the beam to the left in time to begin the next sweep. The timebase can be adjusted to match the sweep time to the period of the signal.

Meanwhile, the vertical amplifier is driven by an external voltage (the vertical input) that is taken from the circuit or experiment that is being measured. The amplifier has a very high input impedance, typically one megohm, so that it draws only a tiny current from the signal source. The amplifier drives the vertical deflection plates with a voltage that is proportional to the vertical input. Because the electrons have already been accelerated by hundreds of volt, this amplifier also has to deliver almost hundred volts, and this with a very high bandwidth. The gain of the vertical amplifier can be adjusted to suit the amplitude of the input voltage. A positive input voltage bends the electron beam upwards, and a negative voltage bends it downwards, so that the vertical deflection of the dot shows the value of the input. The response of this system is much faster than that of mechanical measuring devices such as the multimeter, where the inertia of the pointer slows down its response to the input.

When all these components work together, the result is a bright trace on the screen that represents a graph of voltage against time. Voltage is on the vertical axis, and time on the horizontal.

Observing high speed signals, especially non-repetitive signals, with a conventional CRO is difficult, due to non-stable or changing triggering threshold which makes it hard to "freeze" the waveform on the screen. This often requires the room to be darkened or a special viewing hood to be placed over the face of the display tube. To aid in viewing such signals, special oscilloscopes have borrowed from night vision technology, employing a microchannel plate in the tube face to amplify faint light signals.

Although a CRO allows one to view a signal, in its basic form it has no means of recording that signal on paper for the purpose of documentation. Therefore, special oscilloscope cameras were developed to photograph the screen directly. Early cameras used roll or plate film, while in the 1970s Polaroid® instant cameras became popular.

The vertical amplifier and timebase controls are calibrated to show the vertical distance on the screen that corresponds to a given voltage difference, and the horizontal distance that corresponds to a given time interval.

The power supply is an important component of the scope. It provides low voltages to power the cathode heater in the tube, and the vertical and horizontal amplifiers. High voltages are needed to drive the electrostatic deflection plates. These voltages must be very stable. Any variations will cause errors in the position and brightness of the trace.

Later analogue oscilloscopes added digital processing to the standard design. The same basic architecture - cathode ray tube, vertical and horizontal amplifiers - was retained, but the electron beam was controlled by digital circuitry that could display graphics and text mixed with the analogue waveforms. The extra features that this system provides include:

on-screen display of amplifier and timebase settings;
voltage cursors - adjustable horizontal lines with voltage display;
time cursors - adjustable vertical lines with time display;
on-screen menus for trigger settings and other functions.

Dual beam oscilloscope

A **dual beam oscilloscope** was a type of oscilloscope once used to compare one signal with another. There were two beams produced in a special type of CRT. Unlike an ordinary "dual-trace" oscilloscope (which time-shared a single electron beam, thus losing about 50% of each signal), a dual beam oscilloscope simultaneously produced two separate electron beams, capturing the entirety of both signals.

Two pairs of vertical plates deflected the beams. Vertical plates for channel A had no effect on channel B beam. Similarly for channel B, separate vertical plates existed which deflected the beam B only.

On some scopes the time base, horizontal plates and horizontal amplifier were common to both beams; on more elaborate scopes like the Tektronix 556 there were two independent time bases and two sets of horizontal plates and horizontal amplifiers. Thus one could look at a very fast signal on one beam and a slow signal on another beam.

Most multichannel 'scopes do not actually have multiple electron beams. Instead, they display only one dot at a time, but switch the dot between one channel and the other either on alternate sweeps (ALT mode) or many times per sweep (CHOP mode). Very few actual *dual beam* oscilloscopes were built.

With the advent of digital signal capture, true dual beam oscilloscopes became obsolete, as it was then possible to display two truly simultaneous signals from memory using either the ALT or CHOP display technique, or even possibly a raster display mode.

Analogue storage oscilloscope

An extra feature available on some analogue scopes is called 'storage'. This feature allows the trace pattern that normally decays in a fraction of a second to remain on the screen for several minutes or longer. An electrical circuit can then be deliberately activated to store and erase the trace on the screen.

The storage is accomplished using the principle of secondary emission. When the ordinary writing electron beam passes a point on the phosphor surface, not only does it momentarily cause the phosphor to illuminate, but the kinetic energy of the electron beam knocks other electrons loose from the phosphor surface. This can leave a net positive charge. Storage oscilloscopes then provide one or more secondary electron guns (called the "flood guns") that provide a steady flood of low-energy electrons traveling towards the phosphor screen. The electrons from the flood guns are more strongly drawn to the areas of the phosphor screen where the writing gun has left a net positive charge; in this way, the electrons from the flood guns re-illuminate the phosphor in these positively-charged areas of the phosphor screen.

If the energy of the flood gun electrons is properly balanced, each impinging flood gun electron knocks out one secondary electron from the phosphor screen, thus preserving the net positive charge in the illuminated areas of the phosphor screen. In this way, the image originally written by the writing gun can be maintained for a long time. Eventually, small imbalances in the secondary emission ratio cause the entire screen to "fade positive" (light up) or cause the originally-written trace to "fade negative" (extinguish). It is these imbalances that limit the ultimate storage time possible.

Some oscilloscopes used a strictly binary (on/off) form of storage known as "bistable storage". Others permitted a constant series of short, incomplete erasure cycles which created the impression of a phosphor with "variable persistence". Certain oscilloscopes also allowed the partial or complete shutdown of the flood guns, allowing the preservation (albeit invisibly) of the latent stored image for later viewing. (Fading positive or fading negative only occurs when the flood guns are "on"; with the flood guns off, only leakage of the charges on the phosphor screen degrades the stored image.)

Digital storage oscilloscope

A digital storage oscilloscope manufactured by Agilent Technologies

The **digital storage oscilloscope**, or DSO for short, is now the preferred type for most industrial applications, although simple analogue CROs are still used by hobbyists. It replaces the unreliable storage method used in analogue storage scopes with digital memory, which can store data as long as required without degradation. It also allows complex processing of the signal by high-speed digital signal processing circuits.

The vertical input, instead of driving the vertical amplifier, is digitised by an analog to digital converter to create a data set that is stored in the memory of a microprocessor. The data set is processed and then sent to the display, which in early DSOs was a cathode ray tube, but is now more likely to be an LCD flat panel. DSOs with color LCD displays are common. The data set can be sent over a LAN or a WAN for processing or archiving. The screen image can be directly recorded on paper by means of an attached printer or plotter, without the need for an oscilloscope camera. The scope's own signal analysis software can extract many useful time-domain features (e.g. rise time, pulse width, amplitude), frequency spectra, histograms and statistics, persistence maps, and a large number of parameters meaningful to engineers in specialized fields such as telecommunications, disk drive analysis and power electronics.

Digital oscilloscopes are limited principally by the performance of the analogue input circuitry and the sampling frequency. In general, the sampling frequency should be at least the Nyquist rate, double the frequency of the highest-frequency component of the observed signal, otherwise aliasing may occur.

Digital storage also makes possible another unique type of oscilloscope, the equivalent-time sample scope. Instead of taking consecutive samples after the trigger event, only one sample is taken. However, the oscilloscope is able to vary its timebase to precisely time its sample, thus building up the picture of the signal over the subsequent repeats of the signal. This requires that either a clock or repeating pattern be provided. This type of scope is frequently used for very high speed communication because it allows for a very high "sample rate" and low amplitude noise compared to traditional real-time scopes.

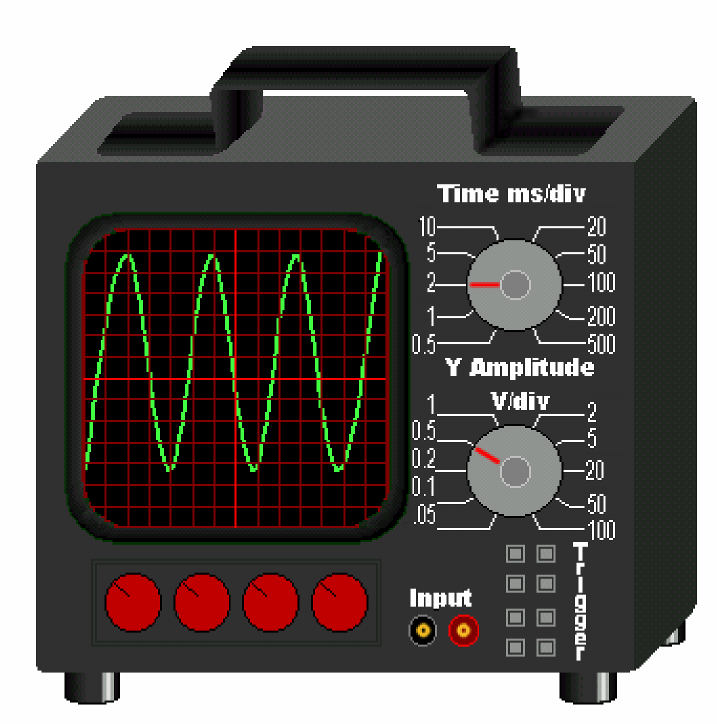
To sum this up: Advantages over the analogue oscilloscope:

Brighter and bigger display with color to distinguish multiple traces
Equivalent time sampling and Average across consecutive samples or scans lead to higher resolution down to µV
Peak detection
Pre-trigger
Easy pan and zoom across multiple stored traces allows beginners to work without a trigger
This needs a fast reaction of the display (some scopes have 1 s delay)
The knobs have to be large and turn smoothly
Also slow traces like the temperature variation across a day can be recorded
The memory of the oscilloscope can be arranged not only as a one-dimensional list but also as a two-dimensional array to simulate a phosphorus screen. The digital technique allows a quantitative analysis (E.g. Eye diagram)
Allows for automation, though most models lock the access to their software

A disadvantage of digital oscilloscopes is the limited refresh rate of the screen. On an analog oscilloscope, the user can get an intuitive sense of the trigger rate simply by looking at the steadiness of the CRT trace. For a digital oscilloscope, the screen looks exactly the same for any signal rate which exceeds the screen's refresh rate. Additionally, it is sometimes hard to spot "glitches" or other rare phenomena on the black-and-white screens of standard digital oscilloscopes; the slight persistence of CRT phosphors on analog scopes makes glitches visible even if many subsequent triggers overwrite them. Both of these difficulties have been overcome recently by "digital phosphor oscilloscopes", which store data at a very high refresh rate and display it with variable intensity, to simulate the trace persistence of a CRT scope.

Features:

Description



Exterior

A typical oscilloscope is a box with a display screen, numerous input connectors, and control knobs and buttons on the front panel. To aid measurement, a grid called the *graticule* is drawn on the face of the screen. Each square in the graticule is known as a *division*.

Inputs

The signal to be measured is fed to one of the input connectors, which is usually a coaxial connector such as a BNC or N type. If the signal source has its own coaxial connector, then a simple coaxial cable is used; otherwise, a specialised cable called a 'scope probe', supplied with the oscilloscope, is used. General-purpose oscilloscopes have a standardised input resistance of 1 megaohm in parallel with a capacitance of around 20 picofarads. This allows the use of standard oscilloscope probes. Scopes for use with very high frequencies may have 50-ohm inputs, which must be either connected directly to a 50-ohm signal source or used with Z_0 or active probes. It is used for measuring voltage.

The trace

In its simplest mode, the oscilloscope repeatedly draws a horizontal line called the *trace* across the middle of the screen from left to right. One of the controls, the *timebase control*, sets the speed at which the line is drawn, and is calibrated in seconds per division. If the input voltage departs from zero, the trace is deflected either upwards or downwards. Another control, the *vertical control*, sets the scale of the vertical deflection, and is calibrated in volts per division. The resulting trace is a graph of voltage against time, with the more distant past on the left and the more recent past on the right.

If the input signal is periodic, then a nearly stable trace can be obtained just by setting the timebase to match the frequency of the input signal. For example, if the input signal is a 50 Hz sine wave, then its period is 20 ms, so the timebase should be adjusted so that the time between successive horizontal sweeps is 20 ms. This mode is called *continual sweep*. Unfortunately, an oscilloscope's timebase is not perfectly accurate, and the frequency of most input signals are not perfectly stable, so the trace will drift across the screen making measurements difficult.

Trigger

An old Tektronix oscilloscope.

To provide a more stable trace, modern oscilloscopes have a function called the *saddle*. When using *saddling*, the scope will pause each time the sweep reaches the extreme right side of the screen. The scope then waits for a specified event before drawing the next trace. The trigger event is usually the input waveform reaching some user-specified threshold voltage in the specified direction (going positive or going negative).

The effect is to resynchronize the timebase to the input signal, preventing horizontal drift of the trace. In this way, triggering allows the display of periodic signals such as sine waves and square waves. Trigger circuits also allow the display of nonperiodic signals such as single pulses or pulses that don't recur at a fixed rate.

Types of trigger include:

external trigger, a pulse from an internal source connected to a dedicated input on the scope.
edge trigger, an edge-detector that generates a pulse when the input signal crosses a specified threshold voltage in a specified direction.
video trigger, a circuit that extracts synchronizing pulses from video formats such as PAL and NTSC and triggers the timebase on every line, a specified line, every field, or every frame. This circuit is typically found in a waveform monitor device.
delayed trigger, which waits a specified time after an edge trigger before starting the sweep. No trigger circuit acts instantaneously, so there is always a certain delay, but a trigger delay circuit extends this delay to a known and adjustable interval. In this way, the operator can examine a particular pulse in a long train of pulses.

X-Y mode

Most modern oscilloscopes have several inputs for voltages, and thus can be used to plot one varying voltage versus another. This is especially useful for graphing I-V curves (current versus voltage characteristics) for components such as diodes, as well as Lissajous patterns. Lissajous figures are an example of how an oscilloscope can be used to track phase differences between multiple input signals. This is very frequently used in broadcast engineering to plot the left and right stereophonic channels, to ensure that the stereo generator is calibrated properly.

Other features

Some oscilloscopes have *cursors*, which are lines that can be moved about the screen to measure the time interval between two points, or the difference between two voltages.

Oscilloscopes may have two or more input *channels*, allowing them to display more than one input signal on the screen. Usually the oscilloscope has a separate set of vertical controls for each channel, but only one triggering system and timebase.

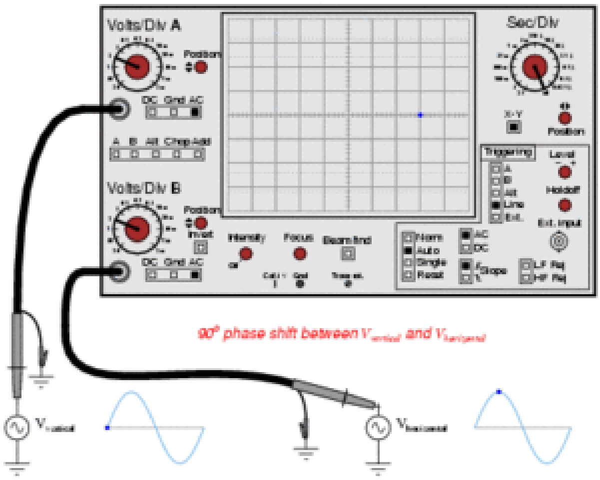
Sometimes the event that the user wants to see may only happen occasionally. To catch these events, some oscilloscopes, known as "storage scopes", preserve the most recent sweep on the screen. This was originally achieved by using a special CRT, a "storage tube", which would retain the image of even a very brief event for a long time.

Some digital oscilloscopes can sweep at speeds as slow as once per hour, emulating a strip chart recorder. That is, the signal scrolls across the screen from right to left. Most oscilloscopes with this facility switch from a sweep to a strip-chart mode at about one sweep per ten seconds. This is because otherwise, the scope looks broken: it's collecting data, but the dot cannot be seen.

Oscilloscopes were originally analog devices. In more recent times digital signal sampling is more often used for all but the simplest models.

Many oscilloscopes have different plug-in modules for different purposes, e.g., high-sensitivity amplifiers of relatively narrow bandwidth, differential amplifiers, amplifiers with 4 or more channels, sampling plugins for repetitive signals of very high frequency, and special-purpose plugins.

[Examples of use



Lissajous figures on an oscilloscope, with 90 degrees phase difference between x and y inputs.

One of the most frequent uses of scopes is troubleshooting malfunctioning electronic equipment. One of the advantages of a scope is that it can graphically show signals: where a voltmeter may show a totally unexpected voltage, a scope may reveal that the circuit is oscillating. In other cases the precise shape of a pulse is important.

In a piece of electronic equipment, for example, the connections between stages (e.g. electronic mixers, electronic oscillators, amplifiers) may be 'probed' for the expected signal, using the scope as a simple signal tracer. If the expected signal is absent or incorrect, some preceding stage of the electronics is not operating correctly. Since most failures occur because of a single faulty component, each measurement can prove that half of the stages of a complex piece of equipment either work, or probably did not cause the fault.

Once the faulty stage is found, further probing can usually tell a skilled technician exactly which component has failed. Once the component is replaced, the unit can be restored to service, or at least the next fault can be isolated.

Another use is to check newly designed circuitry. Very often a newly designed circuit will misbehave because of design errors, bad voltage levels, electrical noise etc. Digital electronics usually operate from a clock, so a dual-trace scope which shows both the clock signal and a test signal dependent upon the clock is useful. "Storage scopes" are helpful for "capturing" rare electronic events that cause defective operation.

Another use is for software engineers who must program electronics. Often a scope is the only way to see if the software is running the electronics properly.

Analog recording:

Analog (or *analogue*) recording is a technique used to store audio or video signals for later playback. The first successful demonstration of analog recording for audio was by Thomas Alva Edison. The first analogs of moving pictures were those of the Lumiere Brothers.

The modern examples of the analog audio recording are:

Gramophone record (aka phonograph record, vinyl, etc).
Wire recording
Magnetic tape, magnetic tape sound recording

The earliest forms of video recording used analog technology initially. John Logie Baird developed a system in the 1920s for the storage of video signals on conventional phonograph records, which he called Phonovision. In the 1930s, he further developed the Intermediate Film Technique, which provided for an analog method of temporary video storage by using cine film.

The analog recording method stores signals as a continual wave in/on the media, rather than the discrete numbers used in digital recording. The wave is stored as a physical texture on a phonograph record, or a fluctuation in the field strength of a magnetic recording.

A perceived drawback of many analog recordings was noise of the media, or of the equipment, and of production equipment limitations. Repeat playing of a gramophone record introduces wear that made the original recording more difficult to hear over the noise level. Careful removal of dirt is helpful; as is careful handling.

Magnetic tape:

Magnetic tape is a medium for magnetic recording generally consisting of a thin magnetizable coating on a long and narrow strip of plastic. Nearly all recording tape is of this type, whether used for recording audio or video or for computer data storage. It was originally developed in Germany, based on the concept of magnetic wire recording.

Devices that record and playback audio and video using magnetic tape are generally called tape recorders and video tape recorders respectively. A device that stores computer data on magnetic tape can be called a tape drive, a tape unit, or a streamer.

Magnetic tape revolutionized the broadcast and recording industries. In an age when all radio (and later television) was live, it allowed programming to be prerecorded. In a time when gramophone records were recorded in one take, it allowed recordings to be created in multiple stages and easily mixed and edited with a minimal loss in quality between generations. It is also one of the key enabling technologies in the development of modern computers. Magnetic tape allowed massive amounts of data to be stored in computers for long periods of time and rapidly accessed when needed.

As of 2007, many other technologies exist that can perform the functions of magnetic tape. In many cases these technologies are replacing tape. Despite this, innovation in the technology continues and tape is still widely used.

Magnetic storage media can be classified as either sequential access memory or random access memory although in some cases the distinction is not perfectly clear. In the case of magnetic wire, the read/write head only covers a very small part of the recording surface at any given time. Accessing different parts of the wire involves winding the wire forward or backward until the point of interest is found. The time to access this point depends on how far away it is from the starting point. The case of ferrite-core memory is the opposite. Every core location is immediately accessible at any given time.

Hard disks and modern linear serpentine tape drives do not precisely fit into either category. Both have many parallel tracks across the width of the media and the read/write heads take time to switch between tracks and to scan within tracks. Different spots on the storage media take different amounts of time to access. For a hard disk this time is typically less than 10 ms, but tapes might take as much as 100 s.