Radio receiver

From Wikipedia, the free encyclopedia

In radio communications, a radio receiver (commonly also called a radio) is an electronic device that receives radio waves and converts the information carried by them to a usable form. It is used with an antenna. The antenna intercepts radio waves (electromagnetic waves) and converts them to tiny alternating currents which are applied to the receiver, and the receiver extracts the desired information. The receiver uses electronic filters to separate the desired radio frequency signal from all the other signals picked up by the antenna, an electronic amplifier to increase the power of the signal for further processing, and finally recovers the desired information through demodulation.

The information produced by the receiver may be in the form of sound (an audio signal), images (a video signal) or digital data. A radio receiver may be a separate piece of electronic equipment, or an electronic circuit within another device. Devices that contain radio receivers include television sets, radar equipment, two-way radios, cell phones, wireless computer networks, GPS navigation devices, satellite dishes, radio telescopes, Bluetooth enabled devices, garage door openers, and baby monitors.

In consumer electronics, the terms radio and radio receiver are often used specifically for receivers designed to reproduce the audio (sound) signals transmitted by radio broadcasting stations, historically the first mass-market commercial radio application.

Contents

- 1 Types of radio receivers
- 2 Consumer audio receivers
  - 2.1 Hi-Fi / Home theater
  - 2.2 Portable radios
- 3 History
  - 3.1 Spark era
    - 3.1.1 Coherer receiver
    - 3.1.2 Other early detectors
    - 3.1.3 Tuning
    - 3.1.4 Inductive coupling
    - 3.1.5 Patent disputes
    - 3.1.6 Crystal radio receiver
    - 3.1.7 Heterodyne receiver and BFO
  - 3.2 Vacuum tube era
    - 3.2.1 The first vacuum tube receivers
    - 3.2.2 Regenerative (autodyne) receiver
    - 3.2.3 Superregenerative receiver
    - 3.2.4 Tuned radio frequency (TRF) receiver
    - 3.2.5 Neutrodyne receiver
    - 3.2.6 Reflex receiver
    - 3.2.7 Superheterodyne receiver
  - 3.3 Semiconductor era
    - 3.3.1 Digital technologies
    - 3.3.2 DSP technology
    - 3.3.3 PC controlled radio receivers
    - 3.3.4 Radio control software
    - 3.3.5 Software-defined radios
- 4 See also
- 5 References
6 Further reading

Types of radio receivers

Various types of radio receivers may include:

- Consumer audio and high fidelity audio receivers and AV receivers used by home stereo listeners and audio and home theatre system enthusiasts as well as audiophiles.
- Communications receivers, used as a component of a radio communication link, characterized by high stability and reliability of performance.
- Simple crystal radio receivers, also known as a crystal set, which operate using the power received from radio waves.
- Satellite television receivers, used to receive television programming from communication satellites in geosynchronous orbit.
- Specialized-use receivers such as telemetry receivers that allow the remote measurement and reporting of information.
- Measuring receivers or measurement receivers are calibrated laboratory-grade devices that are used to measure the signal strength of broadcasting stations, the electromagnetic interference radiation emitted by electrical products, as well as to calibrate RF attenuators and signal generators.
- Scanners are specialized receivers that can automatically scan two or more discrete frequencies, stopping when they find a signal on one of them and then continuing to scan other frequencies when the initial transmission ceases. They are mainly used for monitoring VHF and UHF radio systems.
- Internet radio devices

Consumer audio receivers

In the context of home audio systems, the term "receiver" often refers to a combination of a tuner, a preamplifier, and a power amplifier all on the same chassis. Audiophiles will refer to such a device as an integrated receiver, while a single chassis that implements only one of the three component functions is called a discrete component. Some audio purists still prefer three discrete units - tuner, preamplifier and power amplifier - but the integrated receiver has, for some years, been the mainstream choice for music listening. The first integrated stereo receiver was made by the Harman Kardon company, and came onto the market in 1958. It had undistinguished performance, but it represented a breakthrough to the "all in one" concept of a receiver, and rapidly improving designs gradually made the receiver the mainstay of the marketplace. Many radio receivers also include a loudspeaker.

Hi-Fi / Home theater

Today AV receivers are a common component in a high-fidelity or home-theatre system. The receiver is generally the nerve centre of a sophisticated home-theatre system providing selectable inputs for a number of different audio components like record players, CD players, tape decks and video components like VCRs, DVD players, video game consoles and television sets.

With the decline of gramophone record vinyl discs, modern receivers tend to omit inputs for phonograph turntables, which have separate requirements of their own. All other common audio/visual components can use any of the identical line-level inputs on the receiver for playback, regardless of how they are marked (the "name" on each input is mostly for the convenience of the user). For instance, a second CD player can be plugged into an "AUX" input, and will work the same as it will in the "CD" input jacks.

Some receivers can also provide digital signal processors (DSP) to give a more realistic auditory illusion of listening in a concert hall. Digital audio S/PDIF and USB connections are also common today. The home theater receiver, in the vocabulary of consumer electronics, comprises both the 'radio receiver' and other functions, such as control, sound processing, and power amplification. The standalone radio receiver is usually known in consumer electronics as a tuner.

Further reading

Types of radio receivers

Various types of radio receivers may include:

- Consumer audio and high fidelity audio receivers and AV receivers used by home stereo listeners and audio and home theatre system enthusiasts as well as audiophiles.
- Communications receivers, used as a component of a radio communication link, characterized by high stability and reliability of performance.
- Simple crystal radio receivers, also known as a crystal set, which operate using the power received from radio waves.
- Satellite television receivers, used to receive television programming from communication satellites in geosynchronous orbit.
- Specialized-use receivers such as telemetry receivers that allow the remote measurement and reporting of information.
- Measuring receivers or measurement receivers are calibrated laboratory-grade devices that are used to measure the signal strength of broadcasting stations, the electromagnetic interference radiation emitted by electrical products, as well as to calibrate RF attenuators and signal generators.
- Scanners are specialized receivers that can automatically scan two or more discrete frequencies, stopping when they find a signal on one of them and then continuing to scan other frequencies when the initial transmission ceases. They are mainly used for monitoring VHF and UHF radio systems.
- Internet radio devices

Consumer audio receivers

In the context of home audio systems, the term "receiver" often refers to a combination of a tuner, a preamplifier, and a power amplifier all on the same chassis. Audiophiles will refer to such a device as an integrated receiver, while a single chassis that implements only one of the three component functions is called a discrete component. Some audio purists still prefer three discrete units - tuner, preamplifier and power amplifier - but the integrated receiver has, for some years, been the mainstream choice for music listening. The first integrated stereo receiver was made by the Harman Kardon company, and came onto the market in 1958. It had undistinguished performance, but it represented a breakthrough to the "all in one" concept of a receiver, and rapidly improving designs gradually made the receiver the mainstay of the marketplace. Many radio receivers also include a loudspeaker.

Hi-Fi / Home theater

Today AV receivers are a common component in a high-fidelity or home-theatre system. The receiver is generally the nerve centre of a sophisticated home-theatre system providing selectable inputs for a number of different audio components like record players, CD players, tape decks and video components like VCRs, DVD players, video game consoles and television sets.

With the decline of gramophone record vinyl discs, modern receivers tend to omit inputs for phonograph turntables, which have separate requirements of their own. All other common audio/visual components can use any of the identical line-level inputs on the receiver for playback, regardless of how they are marked (the "name" on each input is mostly for the convenience of the user). For instance, a second CD player can be plugged into an "AUX" input, and will work the same as it will in the "CD" input jacks.

Some receivers can also provide digital signal processors (DSP) to give a more realistic auditory illusion of listening in a concert hall. Digital audio S/PDIF and USB connections are also common today. The home theater receiver, in the vocabulary of consumer electronics, comprises both the 'radio receiver' and other functions, such as control, sound processing, and power amplification. The standalone radio receiver is usually known in consumer electronics as a tuner.

Further reading

Types of radio receivers

Various types of radio receivers may include:

- Consumer audio and high fidelity audio receivers and AV receivers used by home stereo listeners and audio and home theatre system enthusiasts as well as audiophiles.
- Communications receivers, used as a component of a radio communication link, characterized by high stability and reliability of performance.
- Simple crystal radio receivers, also known as a crystal set, which operate using the power received from radio waves.
- Satellite television receivers, used to receive television programming from communication satellites in geosynchronous orbit.
- Specialized-use receivers such as telemetry receivers that allow the remote measurement and reporting of information.
- Measuring receivers or measurement receivers are calibrated laboratory-grade devices that are used to measure the signal strength of broadcasting stations, the electromagnetic interference radiation emitted by electrical products, as well as to calibrate RF attenuators and signal generators.
- Scanners are specialized receivers that can automatically scan two or more discrete frequencies, stopping when they find a signal on one of them and then continuing to scan other frequencies when the initial transmission ceases. They are mainly used for monitoring VHF and UHF radio systems.
- Internet radio devices

Consumer audio receivers

In the context of home audio systems, the term "receiver" often refers to a combination of a tuner, a preamplifier, and a power amplifier all on the same chassis. Audiophiles will refer to such a device as an integrated receiver, while a single chassis that implements only one of the three component functions is called a discrete component. Some audio purists still prefer three discrete units - tuner, preamplifier and power amplifier - but the integrated receiver has, for some years, been the mainstream choice for music listening. The first integrated stereo receiver was made by the Harman Kardon company, and came onto the market in 1958. It had undistinguished performance, but it represented a breakthrough to the "all in one" concept of a receiver, and rapidly improving designs gradually made the receiver the mainstay of the marketplace. Many radio receivers also include a loudspeaker.

Hi-Fi / Home theater

Today AV receivers are a common component in a high-fidelity or home-theatre system. The receiver is generally the nerve centre of a sophisticated home-theatre system providing selectable inputs for a number of different audio components like record players, CD players, tape decks and video components like VCRs, DVD players, video game consoles and television sets.

With the decline of gramophone record vinyl discs, modern receivers tend to omit inputs for phonograph turntables, which have separate requirements of their own. All other common audio/visual components can use any of the identical line-level inputs on the receiver for playback, regardless of how they are marked (the "name" on each input is mostly for the convenience of the user). For instance, a second CD player can be plugged into an "AUX" input, and will work the same as it will in the "CD" input jacks.

Some receivers can also provide digital signal processors (DSP) to give a more realistic auditory illusion of listening in a concert hall. Digital audio S/PDIF and USB connections are also common today. The home theater receiver, in the vocabulary of consumer electronics, comprises both the 'radio receiver' and other functions, such as control, sound processing, and power amplification. The standalone radio receiver is usually known in consumer electronics as a tuner.
Some modern integrated receivers can send audio out to seven loudspeakers and an additional channel for a subwoofer and often include connections for headphones. Receivers vary greatly in price, and support stereophonic or surround sound. A high-quality receiver for dedicated audio-only listening (two channel stereo) can be relatively inexpensive. Excellent ones can be purchased for $300 or less. Because modern receivers are purely electronic devices with no moving parts unlike electromechanical devices like turntables and cassette decks, they tend to offer many years of trouble-free service. In recent years, the home theater in a box has become common, which often integrates a surround-capable receiver with a DVD player. The user simply connects it to a television, perhaps other components, and a set of loudspeakers.

**Portable radios**

Portable radios include simple transistor radios that are typically monoaural and receive the AM, FM, or short wave broadcast bands. FM, and often AM, radios are sometimes included as a feature of portable DVD/CD, MP3 CD, and USB key players, as well as cassette player/recorders.

AM/FM stereo car radios can be a separate dashboard mounted component or a feature of in car entertainment systems.

A boombox or boom-box, sometimes known as a ghetto blaster or a jambox, in parts of Europe as a "radio-cassette", is a name given to larger portable stereo systems capable of playing radio stations and recorded music, often at a high level of volume.

Self-powered portable radios, such as clockwork radios are used in developing nations or as part of an emergency preparedness kit.[3]

**History**

Radio waves were first identified in German physicist Heinrich Hertz's 1887 series of experiments to prove James Clerk Maxwell's electromagnetic theory. Hertz used spark-excited dipole antennas to generate the waves and micrometer spark gaps attached to dipole and loop antennas to detect them.[4][5][6] These primitive devices are more accurately described as radio wave sensors, not "receivers", as they could only detect radio waves within about 100 feet of the transmitter, and were not used to attempt communication but were instead used as part of an experiment to verify a theoretical proof.

**Spark era**

The earliest radio communication systems, used during the first three decades of radio, 1887-1917, called the wireless telegraphy or "spark" era, used spark gap transmitters which generated radio waves by discharging a capacitance through an electric spark.[8][9][10] Each spark produced a transient pulse of radio waves consisting of a sinusoidal wave which decreased rapidly exponentially to zero.[4][6] Spark transmitters produced strings of these damped waves, and could not generate the sinusoidal continuous waves which are modulated to carry sound in modern AM and FM transmission. Thus spark transmitters could not transmit audio (sound), and instead transmitted information by radiotelegraphy; the transmitter was switched on and off rapidly by the operator using a telegraph key, creating different length pulses of these damped radio waves ("dots" and "dashes") to spell out text messages in Morse code.[6][9]

Therefore, the first radio receivers did not have to demodulate the radio signal - extract an audio signal from it as modern receivers do - just detect the presence or absence of the radio signal, and produce a sound in the earphone during the "dots" and "dashes" of the Morse code.[6] The device which did this was called a "detector". Since there were no amplifying devices during this era, the sensitivity of the receiver mostly depended on the detector, and many different detector devices were tried. Radio receivers during the spark era (see diagram) consisted of these parts:
- An **antenna**, to intercept the radio waves and convert them to tiny radio frequency electric currents.
- A **resonant circuit** (tuned circuit), consisting of a capacitor connected to a coil of wire (an inductor), which acted as a bandpass filter to select the desired signal out of all the signals picked up by the antenna. Either the capacitor or inductor was adjustable to tune the receiver to the frequency of different transmitters. The earliest receivers, before 1897, did not have tuned circuits, they responded to all radio signals picked up by their antennas, so they had little frequency-discriminating ability and received any transmitter in their vicinity.[11] Most receivers used a pair of tuned circuits with their coils magnetically coupled, called a resonant transformer (oscillation transformer) or "loose coupler".
- A **detector**, which produced a pulse of DC current for each damped wave received.
- An indicating device such as an **earphone**, which converted the pulses of current into sound waves. The first receivers used an electric bell instead. Later receivers in commercial wireless systems used a Morse siphon recorder,[4] which consisted of an ink pen mounted on a needle swung by an electromagnet (a galvanometer) which drew a line on a moving paper tape. Each string of damped waves constituting a Morse "dot" or "dash" caused the needle to swing over, creating a displacement of the line, which could be read off the tape. With such an automated receiver a radio operator didn't have to continuously monitor the receiver.

The signal from the spark gap transmitter consisted of damped waves repeated at an audio frequency rate, from 120 to perhaps 4000 per second, so in the earphone the signal sounded like a musical tone or buzz, and the Morse code "dots" and "dashes" sounded like beeps.

The first person to use radio waves for **communication** was Guglielmo Marconi. Marconi invented little himself, but he was first to believe that radio could be a practical communication medium, and singlehandedly developed the first wireless telegraphy systems, transmitters and receivers, beginning in 1895,[12] mainly by improving technology invented by others.[9][13][14][15][16][17][18][19][20] Oliver Lodge and Alexander Popov were also experimenting with similar radio wave receiving apparatus at the same time in 1895,[14][18] but they are not known to have transmitted Morse code during this period,[9][12] just strings of random pulses. Therefore, Marconi is usually given credit for building the first radio receivers.

### Coherer receiver

The first radio receivers invented by Marconi, Oliver Lodge and Alexander Popov in 1894-5 used a primitive radio wave detector called a coherer, invented in 1890 by Edouard Branly and improved by Lodge and Marconi.[4][9][11][14][18][19][20][21] The coherer was a glass tube with metal electrodes at each end, with loose metal powder between the electrodes.[4][9][21] It initially had a high resistance. When a radio frequency voltage was applied to the electrodes, its resistance dropped and it conducted electricity. In the receiver the coherer was connected directly between the antenna and ground. In addition to the antenna, the coherer was connected in a DC circuit with a battery and relay. When the incoming radio wave reduced the resistance of the coherer, the current from the battery flowed through it, turning on the relay to ring a bell or make a mark on a paper tape in a siphon recorder. In order to restore the coherer to its previous nonconducting state to receive the next pulse of radio waves, it had to be tapped mechanically to disturb the metal particles.[4][9][18][22] This was done by a "decoherer", a clapper which struck the tube, operated by an electromagnet powered by the relay.

The coherer is an obscure antique device, and even today there is some uncertainty about the exact physical mechanism by which the various types worked.[4][13][22] However it can be seen that it was essentially a bistable device, a radio-wave-operated switch, and so it did not have the ability to rectify the radio wave to demodulate the later amplitude modulated (AM) radio transmissions that carried sound.[4][13]
In a long series of experiments Marconi found that by using an elevated wire monopole antenna instead of Hertz's dipole antennas he could transmit longer distances, beyond the curve of the Earth, demonstrating that radio was not just a laboratory curiosity but a commercially viable communication method. This culminated in his historic transatlantic wireless transmission on December 12, 1901 from Poldhu, Cornwall to St. John's, Newfoundland, a distance of 3500 km (2200 miles), which was received by a coherer.\[13\][17] However the usual range of coherer receivers even with the powerful transmitters of this era was limited to a few hundred miles.

The coherer remained the dominant detector used in early radio receivers for about 10 years,\[21\] until replaced by the crystal detector and electrolytic detector around 1907. In spite of much development work, it was a very crude unsatisfactory device.\[4\][9\] It was not very sensitive, and also responded to impulsive radio noise (RFI), such as nearby lights being switched on or off, as well as to the intended signal.\[9\][21\] Due to the cumbersome mechanical "tapping back" mechanism it was limited to a data rate of about 12-15 words per minute of Morse code, while a spark-gap transmitter could transmit Morse at up to 100 WPM with a paper tape machine.\[24\][25\]

Other early detectors

The coherer's poor performance motivated a great deal of research to find better radio wave detectors, and many were invented. Some strange devices were tried; researchers experimented with using frog legs and even a human brain from a cadaver as detectors.\[4\][26\]

By the first years of the 20th century, experiments in using amplitude modulation (AM) to transmit sound by radio (radiotelephony) were being made. So a second goal of detector research was to find detectors that could demodulate an AM signal, extracting the audio (sound) signal from the radio carrier wave. It was found by trial and error that this could be done by a detector that exhibited "asymmetrical conduction"; a device that conducted current in one direction but not in the other.\[27\] This rectified the alternating current radio signal, removing one side of the carrier cycles, leaving a pulsing DC current whose amplitude varied with the audio modulation signal. When applied to an earphone this would reproduce the transmitted sound.

Below are the detectors that saw wide use before vacuum tubes took over around 1920.\[28\][29\] All except the magnetic detector could rectify and therefore receive AM signals:

- **Magnetic detector** - Developed by Guglielmo Marconi in 1902 from a method invented by Ernest Rutherford and used by the Marconi Co. until it adopted the Audion vacuum tube around 1912, this was a mechanical device consisting of an endless band of iron wires which passed between two pulleys turned by a windup mechanism.\[30\][31\][32][33\] The iron wires passed through a coil of fine wire attached to the antenna, in a magnetic field created by two magnets. The hysteresis of the iron induced a pulse of current in a sensor coil each time a radio signal passed through the exciting coil. The magnetic detector was used on shipboard receivers due to its insensitivity to vibration. One was part of the wireless station of the RMS *Titanic* which was used to summon help during its famous 15 April 1912 sinking.\[34\]

- **Electrolytic detector ("liquid barretter")** - Invented in 1903 by Reginald Fessenden, this consisted of a thin silver-plated platinum wire enclosed in a glass rod, with the tip making contact with the surface of a cup of nitric acid.\[4\][35][36][37\] The
electrolytic action caused current to be conducted in only one direction. The detector was used until about 1910. Electrolytic detectors that Fessenden had installed on US Navy ships received the first AM radio broadcast on Christmas Eve, 1906, an evening of Christmas music transmitted by Fessenden using his new alternator transmitter.

- **Thermionic diode (Fleming valve)** - The first vacuum tube, invented in 1904 by John Ambrose Fleming, consisted of an evacuated glass bulb containing two electrodes: a cathode consisting of a hot wire filament similar to that in an incandescent light bulb, and a metal plate anode. Fleming, a consultant to Marconi, invented the valve as a more sensitive detector for transatlantic wireless reception. The filament was heated by a separate current through it and emitted electrons into the tube by thermionic emission, an effect which had been discovered by Thomas Edison. The radio signal was applied between the cathode and anode. When the anode was positive, a current of electrons flowed from the cathode to the anode, but when the anode was negative the electrons were repelled and no current flowed. The Fleming valve was used to a limited extent but was not popular because it was expensive, had limited filament life, and was not as sensitive as electrolytic or crystal detectors.

- **Crystal detector (cat's whisker detector)** - invented around 1904-1906 by Henry H. C. Dunwoody and Greenleaf Whittier Pickard, based on Karl Ferdinand Braun's 1874 discovery of "asymmetrical conduction" in crystals, these were the most successful and widely used detectors before the vacuum tube era and gave their name to the crystal radio receiver (below). One of the first semiconductor electronic devices, a crystal detector consisted of a pea-sized pebble of a crystalline semiconductor mineral such as galena (lead sulfide) whose surface was touched by a fine springy metal wire mounted on an adjustable arm. This functioned as a primitive diode which conducted electric current in only one direction. In addition to their use in crystal radios, carborundum crystal detectors were also used in some early vacuum tube radios because they were more sensitive than the vacuum tube grid-leak detector.

During the vacuum tube era, the term "detector" changed from meaning a radio wave detector to mean a demodulator, a device that could extract the audio modulation signal from a radio signal. That is its meaning today.

**Tuning**

The word "tuning" means adjusting the frequency of the receiver (its passband) to the frequency of the desired radio transmitter, to receive the desired radio transmission. Although in many modern radio receivers (such as television sets) this is done automatically by pressing a "channel up" or "channel down" button, some older radios still use the traditional method of turning a "tuning" knob until the station is heard in the speaker. However the very first receivers in the spark era could not be tuned at all.

The first receivers had no tuned circuit, the detector was connected directly between the antenna and ground. Due to the lack of any frequency selective components besides the antenna, the bandwidth of the receiver was equal to the broad bandwidth of the antenna. This was acceptable and even necessary because the first Hertzian spark transmitters also lacked a tuned
Marconi's inductively coupled coherer receiver from his controversial April 1900 "four circuit" patent no. 7,777. Due to the impulsive nature of the spark, they radiated a very "noisy" signal, the energy of the radio waves was spread over a very wide band of frequencies, a wide bandwidth.[44][45] To receive enough energy from this wideband signal the receiver had to have a wide bandwidth also.

It was found that when more than one spark transmitter was radiating in a given area, their frequencies overlapped, so their signals interfered with each other, resulting in garbled reception.[10][43][46] It became clear that, if multiple transmitters were to operate simultaneously, some method of selective signaling was needed, to allow the receiver to select which transmitter's signal to receive.[46][47] As radio transmission and detection systems were being developed it was also theorized that the multiple wavelengths produced by a poorly tuned transmitter were overlapping each other causing the signal to "dampen", or die down, greatly reducing the power and range of transmission.[48] In 1892, William Crookes gave a lecture[49] on radio in which he suggested using resonance to reduce the bandwidth of transmitters and receivers. Different transmitters could then be "tuned" to transmit on different frequencies so they didn't interfere.[17][44][50] The receiver would also have a resonant circuit (tuned circuit), and could receive a particular transmission by "tuning" its resonant circuit to the same frequency as the transmitter, analogously to tuning a musical instrument to resonance with another. This is the system used in all modern radio.

Tuning was used in Hertz's original experiments[51] and practical application of tuning showed up in the early to mid 1890s in wireless systems not specifically designed for radio communication. Nikola Tesla's March 1893 lecture demonstrating the wireless transmission of power for lighting (mainly by what he thought was ground conduction[52]) included elements of tuning. The wireless lighting system consisted of a spark-excited grounded resonant transformer with a wire antenna which transmitted power across the room to another resonant transformer tuned to the frequency of the transmitter, which lighted a Geissler tube.[14][50] Use of tuning in free space "Hertzian waves" (radio) was explained and demonstrated in Oliver Lodge's 1894 lectures on Hertz's work.[53] At the time Lodge was demonstrating the physics and optical qualities of radio waves instead of attempting to build a communication system but he would go on to develop methods (patented in 1897) of tuning radio (what he called "syntony"), including using variable inductance to tune antennas.[54][55][56]

By 1897 the advantages of tuned systems had become clear, and Marconi and the other wireless researchers had incorporated tuned circuits, consisting of capacitors and inductors connected together, into their transmitters and receivers.[10][14][17][19][43][55] The tuned circuit acted like an electrical analog of a tuning fork. It had a high impedance at its resonant frequency, but a low impedance at all other frequencies. Connected between the antenna and the detector it served as a bandpass filter, passing the signal of the desired station to the detector, but routing all other signals to ground.[11] The frequency of the station received $f$ was determined by the capacitance $C$ and inductance $L$ in the tuned circuit:

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

**Inductive coupling**

In order to reject radio noise and interference from other transmitters near in frequency to the desired station, the bandpass filter (tuned circuit) in the receiver has to have a narrow bandwidth, allowing only a narrow band of frequencies through.[10][11] The form of bandpass filter that was used in the first receivers, which has continued to be used in receivers until recently, was the double-tuned inductively-coupled circuit, or resonant transformer (oscillation transformer or RF transformer).[10][14][17][19][55][57] The antenna and ground were connected to a coil of wire, which was magnetically coupled to a second coil with a capacitor across it, which was connected to the detector.[11] The RF alternating current from the antenna through the primary coil created a magnetic field which induced a current in the secondary coil which fed the detector. Both primary and secondary were tuned circuits;[43] the primary coil resonated with the capacitance of the antenna, while the secondary coil resonated with the capacitor across it. Both were adjusted to the same resonant frequency.

This circuit had two advantages.[11] One was that by using the correct turns ratio, the impedance of the antenna could be matched to the impedance of the receiver, to transfer maximum RF power to the receiver. Impedance matching was important to achieve maximum receiving range.
in the unamplified receivers of this era.[7][11] The coils usually had taps which could be selected by a multiposition switch. The second advantage was that due to "loose coupling" it had a much narrower bandwidth than a simple tuned circuit, and the bandwidth could be adjusted.[10][57] Unlike in an ordinary transformer, the two coils were "loosely coupled"; separated physically so not all the magnetic field from the primary passed through the secondary, reducing the mutual inductance. This gave the coupled tuned circuits much "sharper" tuning, a narrower bandwidth than a single tuned circuit. In the "Navy type" loose coupler (see picture), widely used with crystal receivers, the smaller secondary coil was mounted on a rack which could be slid in or out of the primary coil, to vary the mutual inductance between the coils.[10][58] When the operator encountered an interfering signal at a nearby frequency, the secondary could be slid further out of the primary, reducing the coupling, which narrowed the bandwidth, rejecting the interfering signal. A disadvantage was that all three adjustments in the loose coupler - primary tuning, secondary tuning, and coupling - were interactive; changing one changed the others. So tuning in a new station was a process of successive adjustments.

Selectivity became more important as spark transmitters were replaced by continuous wave transmitters which transmitted on a narrow band of frequencies, and broadcasting led to a proliferation of closely spaced radio stations crowding the radio spectrum. [11] Resonant transformers continued to be used as the bandpass filter in vacuum tube radios, and new forms such as the variometer were invented.[58][59] Another advantage of the double-tuned transformer for AM reception was that when properly adjusted it had a "flat top" frequency response curve as opposed to the "peaked" response of a single tuned circuit.[60] This allowed it to pass the sidebands of AM modulation on either side of the carrier with little distortion, unlike a single tuned circuit which attenuated the higher audio frequencies. Until recently the bandpass filters in the superheterodyne circuit used in all modern receivers were made with resonant transformers, called IF transformers.

Patent disputes

Marconi's initial radio system had relatively poor tuning limiting its range and adding to interference.[61] To overcome this drawback he developed a four circuit system with tuned coils in "symphony" at both the transmitters and receivers.[61] His 1900 British #7,777 (four sevens) patent for tuning filed in April 1900 and granted a year later opened the door to patents disputes since it infringed on the Syntonic patents of Oliver Lodge, first filed in May 1897, as well as patents filed by Ferdinand Braun.[61] Marconi was able to obtain patents in the UK and France but the US version of his tuned four circuit patent, filed in November 1900, was initially rejected based on it being anticipated by Lodge's tuning system, and refiled versions were rejected because of the prior patents by Braun, and Lodge.[62] A further clarification and re-submission was rejected because it infringed on parts of two prior patents Tesla had obtained for his wireless power transmission system.[63] Marconi's lawyers managed to get a resubmitted patent reconsidered by another examiner who initially rejected it due to a pre-existing John Stone Stone tuning patent, but it was finally approved it in June 1904 based on it having a unique system of variable inductance tuning that was different from Stone[64][65] who tuned by varying the length of the antenna.[62] When Lodge's Syntonic patent was extended in 1911 for another 7 years the Marconi Company agreed to settle that patent dispute, purchasing Lodge's radio company with its patent in 1912, giving them the priority patent they needed.[66][67] Other patent disputes would crop up over the years including a 1943 US Supreme Court ruling on the Marconi Companies ability to sue the US government over patent infringement during World War I. The Court rejected the Marconi Companies suit saying they could not sue for patent infringement when their own patents did not seem to have priority over the patents of Lodge, Stone, and Tesla.[14][50]

Crystal radio receiver

Although it was invented in 1904 in the wireless telegraphy era, the crystal radio receiver could also rectify AM transmissions and served as a bridge to the broadcast era. In addition to being the main type used in commercial stations during the wireless telegraphy era, it was the first receiver to be used widely by the public.[68] During the first two decades of the 20th century, as
Prior to 1920 the crystal receiver was the main type used in wireless telegraphy stations, and sophisticated models were made, like this Marconi Type 106 from 1915.

Family listening to the first broadcasts around 1920 with a crystal receiver. The mother and father have to share an earphone.

After vacuum tube receivers appeared around 1920, the crystal set became a simple cheap alternative radio used by youth and the poor.

Simple crystal radio. The capacitance of the wire antenna connected to the coil serves as the capacitor in the tuned circuit.

Typical "loose coupler" crystal radio circuit

Radio stations began to transmit in AM voice (radiotelephony) instead of radiotelegraphy, radio listening became a popular hobby, and the crystal was the simplest, cheapest detector. The millions of people who purchased or homemade these inexpensive reliable receivers created the mass listening audience for the first radio broadcasts, which began around 1920.\[69\] In the 1920s the crystal receiver was superseded by vacuum tube receivers and became obsolete. However it continued to be used by youth and the poor until World War 2.\[68\] Today crystal radios, the simplest type of radio receivers, are constructed by students as educational science projects.

The crystal radio used a crystal detector called a cat's whisker detector, invented by Harrison H. C. Dunwoody and Greenleaf Whittier Pickard in 1904, to rectify the radio signal to extract the audio from the radio frequency carrier.\[11\][31][70] It consisted of a mineral crystal, usually galena (PbS, lead sulfide) which was lightly touched by a fine springy wire (the "cat whisker") on an adjustable arm.\[31\][71] The resulting crude semiconductor junction functioned as a Schottky barrier diode, it only conducted current in one direction. Only particular sites on the crystal surface functioned as rectifying junctions, and the junction could be disrupted by the slightest vibration. So a usable site was found by trial and error before each use; the operator would drag the cat's whisker across the crystal until the radio began functioning. Frederick Seitz, a later semiconductor researcher, wrote:

Such variability, bordering on what seemed the mystical, plagued the early history of crystal detectors and caused many of the vacuum tube experts of a later generation to regard the art of crystal rectification as being close to disreputable.\[72\]

Like other receivers of this era the crystal radio was unamplified and ran off the power of the radio waves received from the radio station, so it had to be listened to with earphones; it could not drive a loudspeaker.\[11\][71] It required a long wire antenna, and its sensitivity depended on how large the antenna was. During the wireless era it was used in commercial and military longwave stations with huge antennas to receive long distance radiotelegraphy traffic, even including transatlantic traffic.\[73\][74] However, when used to receive broadcast stations a typical home crystal set had a more limited range of about 25 miles.\[75\] In sophisticated crystal radios the "loose coupler" inductively-coupled tuned circuit was used to increase the Q. However it still had poor selectivity compared to modern receivers.\[71\]

**Heterodyne receiver and BFO**

Beginning around 1905 continuous wave (CW) transmitters began to replace spark transmitters for radiotelegraphy because they had much greater range. The first continuous wave transmitters were the Poulsen arc invented in 1904 and the Alexanderson alternator developed 1906-1910, which were replaced by vacuum tube transmitters beginning around 1920.\[6\]
The continuous wave radiotelegraphy signals produced by these transmitters required a different method of reception.\cite{76,77} The radiotelegraphy signals produced by spark gap transmitters consisted of strings of damped waves repeating at an audio rate, so the "dots" and "dashes" of Morse code were audible as a tone or buzz in the receivers' earphones. However the new continuous wave radiotelegraph signals simply consisted of pulses of unmodulated carrier (sine waves). These were inaudible in the receiver headphones. To receive this new modulation type, the receiver had to produce some kind of tone during the pulses of carrier.

The first crude device that did this was the "ticker" or "tikker", invented in 1908 by Valdemar Poulsen.\cite{28,76,78} This was a vibrating interrupter with a capacitor at the tuner output which served as a rudimentary modulator, interrupting the carrier at an audio rate, thus producing a buzz in the earphone when the carrier was present.\cite{79} A similar device was the "tone wheel" invented by Rudolph Goldschmidt, a wheel spun by a motor with contacts spaced around its circumference, which made contact with a stationary brush.

In 1901 Reginald Fessenden had invented a better means of accomplishing this.\cite{76,78,80,81} In his heterodyne receiver an unmodulated sine wave radio signal at a frequency \( f_0 \) offset from the incoming radio wave carrier \( f_C \) was applied to a rectifying detector such as a crystal detector or electrolytic detector, along with the radio signal from the antenna. In the detector the two signals mixed, creating two new heterodyne (beat) frequencies at the sum \( f_C + f_0 \) and the difference \( f_C - f_0 \) between these frequencies. By choosing \( f_0 \) correctly the lower heterodyne \( f_C - f_0 \) was in the audio frequency range, so it was audible as a tone in the earphone whenever the carrier was present. Thus the "dots" and "dashes" of Morse code were audible as musical "beeps". A major attraction of this method during this pre-amplification period was that the heterodyne receiver actually amplified the signal somewhat, the detector had "mixer gain".\cite{78}

The receiver was ahead of its time, because when it was invented there was no oscillator capable of producing the radio frequency sine wave \( f_0 \) with the required stability.\cite{82} Fessenden first used his large radio frequency alternator,\cite{79} but this wasn't practical for ordinary receivers. The heterodyne receiver remained a laboratory curiosity until a cheap compact source of continuous waves appeared, the vacuum tube electronic oscillator\cite{78} invented by Edwin Armstrong and Alexander Meissner in 1913.\cite{28,83} After this it became the standard method of receiving CW radiotelegraphy. The heterodyne oscillator is the ancestor of the beat frequency oscillator (BFO) which is used to receive radiotelegraphy in communications receivers today. The heterodyne oscillator had to be retuned each time the receiver was tuned to a new station, but in modern superheterodyne receivers the BFO signal beats with the fixed intermediate frequency, so the beat frequency oscillator can be a fixed frequency.

Armstrong later used Fessenden's heterodyne principle in his superheterodyne receiver (below).\cite{78,79}

**Vacuum tube era**

The Audion (triode) vacuum tube invented by Lee De Forest in 1906 was the first practical amplifying device and revolutionized radio.\cite{38} Vacuum tube transmitters replaced spark transmitters and made possible four new types of modulation: continuous wave (CW) radiotelegraphy, amplitude modulation (AM) around 1915 which could carry audio (sound), frequency modulation (FM) around 1938 which had much improved audio quality, and single sideband (SSB).

The amplifying vacuum tube used energy from a battery or electrical outlet to increase the power of the radio signal, so vacuum tube receivers could be more sensitive and have a greater reception range than the previous unamplified receivers. The increased audio output power also allowed them to drive loudspeakers instead of earphones, permitting more than one person to listen. The first loudspeakers were produced around 1915. These changes caused radio listening to evolve explosively from a solitary hobby to a popular social and family pastime. The development of amplitude modulation (AM) and vacuum tube transmitters during World War I, and the availability of cheap receiving tubes after the war, set the stage for the start of AM broadcasting, which sprang up spontaneously around 1920.
Unlike today, when almost all radios use a variation of the superheterodyne design, during the 1920s vacuum tube radios used a variety of competing circuits.

A vacuum tube receiver required several power supplies at different voltages, which in early radios were supplied by separate batteries. These were the standard batteries used in early radios:[84][85][86]

- "A" battery - This supplied current to heat the filaments of the tubes, which consumed the bulk of the power in early radios. The first tubes used 6V at several amperes, so lead-acid automobile batteries were often used, as they could be recharged. Later tubes used 3V or 1.5V from dry cell batteries.
- "B" battery - this supplied the plate (anode) voltage for the tubes, including the audio output power to the earphone or loudspeaker. These were rectangular multicell carbon-zinc batteries. They were made in multiples of 22.5 volts: 22.5, 45, 67.5, and 90 volts, and often had taps to give different voltages.
- "C" battery - a few radios required a third voltage of about 4V to bias the grid of the tubes negative.

By 1930 adequate rectifier tubes were developed, and the expensive batteries were replaced by a transformer power supply that worked off the house current.[84][85]

Vacuum tubes were bulky, expensive, had a limited lifetime, consumed a large amount of power and produced a lot of waste heat, so the number of tubes a receiver could economically have was a limiting factor. Therefore, a goal of tube receiver design was to get the most performance out of a limited number of tubes. The major radio receiver designs, listed below, were invented during the vacuum tube era.

A defect in many early vacuum tube receivers was that the amplifying stages could oscillate, act as an oscillator, producing unwanted radio frequency alternating currents.[92][93] These parasitic oscillations mixed with the carrier of the radio signal in the detector tube, producing audible beat notes (heterodynes); annoying whistles, moans, and howls in the speaker. This was due to feedback in the amplifiers; one major feedback path was the capacitance between the plate and grid in early triodes.[92][93] This was solved by the Neutrodyne circuit, and later the development of the tetrode and pentode around 1930.

Edwin Armstrong is one of the most important figures in radio receiver history, and during this period invented technology which continues to dominate radio communication.[79][99] He was the first to give a correct explanation of how De Forest's triode tube worked. He invented the feedback oscillator, regenerative receiver, the superregenerative receiver, the superheterodyne receiver, and modern frequency modulation (FM).
The first vacuum tube receivers

The first amplifying vacuum tube, the Audion, a crude triode, was invented in 1906 by Lee De Forest as a more sensitive detector for radio receivers, by adding a third electrode to the thermionic diode detector, the Fleming valve.[38][59][84][85] It was not widely used until its amplifying ability was recognized around 1912.[38] The first tube receivers, invented by De Forest and built by hobbyists until the mid 1920s, used a single Audion which functioned as a grid-leak detector which both rectified and amplified the radio signal.[59][92][96] The grid-leak detector circuit was also used in regenerative, TRF, and early superheterodyne receivers (below) until the 1930s.

To give enough output power to drive a loudspeaker, 2 or 3 additional Audion stages were needed for audio amplification.[59] Many early hobbyists could only afford a single tube receiver, and listened to the radio with earphones, so early tube amplifiers and speakers were sold as add-ons.

In addition to very low gain of about 5 and a short lifetime of about 100 hours, the primitive Audion had erratic characteristics because it was incompletely evacuated, some residual air was left in the tube by De Forest, who believed that ionization was key to its operation.[97][98] This made it a more sensitive detector[97] but also caused its electrical characteristics to vary during use.[59][94] As the tube heated up, gas released from the metal elements would change the pressure in the tube, changing the plate current and other characteristics, so it required periodic bias adjustments to keep it at the correct operating point. Each Audion stage usually had a rheostat to adjust the filament current, and often a potentiometer or multiposition switch to control the plate voltage. The filament rheostat was also used as a volume control. The many controls made multitube Audion receivers nightmarishly complicated to operate.

By 1914, Harold Arnold at Western Electric and Irving Langmuir at GE had realized that the residual gas in the tube that caused these problems was not necessary; the Audion could operate on electron conduction alone.[94][97][98] They were able to evacuate tubes to a lower pressure of $10^{-9}$ atm, producing the first "hard vacuum" triodes. These more stable tubes did not require bias adjustments and allowed radios to have fewer controls and be more user-friendly.[94] During World War I civilian radio use was prohibited, but by 1920 these tubes came on the market and large scale production of vacuum tube radios began. The "soft" incompletely-evacuated tubes were used as detectors through the 1920s then became obsolete.

Regenerative (autodyne) receiver

The regenerative receiver, invented by Edwin Armstrong[99] in 1913 when he was a 23-year-old college student,[100] was used very widely until the late 1920s particularly by hobbyists who could only afford a single-tube radio. Today transistor versions of the circuit are still used in a few inexpensive applications like walkie-talkies. In the regenerative receiver the gain (amplification) of a vacuum tube or transistor is increased by using regeneration (positive feedback); some of the energy from the tube's output
circuit is fed back into the input circuit with a feedback loop. The early vacuum tubes had very low gain (around 5). Regeneration could not only increase the gain of the tube enormously, by a factor of 15,000 or more, it also increased the Q factor of the tuned circuit, decreasing (sharpening) the bandwidth of the receiver by the same factor, improving selectivity greatly. The receiver had a control to adjust the feedback. The tube also acted as a grid-leak detector to rectify the AM signal.

Another advantage of the circuit was that the tube could be made to oscillate, and thus a single tube could serve as both a beat frequency oscillator and a detector, functioning as a heterodyne receiver to make CW radiotelegraphy transmissions audible. This mode was called an autodyne receiver. To receive radiotelegraphy, the feedback was increased until the tube oscillated, then the oscillation frequency was tuned to one side of the transmitted signal. The incoming radio carrier signal and local oscillation signal mixed in the tube and produced an audible heterodyne (beat) tone at the difference between the frequencies.

A widely used design was the Armstrong circuit, in which a "tickler" coil in the plate circuit was coupled to the tuning coil in the grid circuit, to provide the feedback. The feedback was controlled by a variable resistor, or alternately by moving the two windings physically closer together to increase loop gain, or apart to reduce it. This was done by an adjustable air core transformer called a variometer (variocoupler). Regenerative detectors were sometimes also used in TRF and superheterodyne receivers.

One problem with the regenerative circuit was that when used with large amounts of regeneration the selectivity (Q) of the tuned circuit could be too sharp, attenuating the AM sidebands, thus distorting the audio modulation. This was usually the limiting factor on the amount of feedback that could be employed.

A more serious drawback was that it could act as an inadvertent radio transmitter, producing interference (RFI) in nearby receivers. In AM reception, to get the most sensitivity the tube was operated very close to instability and could easily break into oscillation (and in CW reception did oscillate), and the resulting radio signal was radiated by its wire antenna. In nearby receivers, the regenerative's signal would beat with the signal of the station being received in the detector, creating annoying heterodynes, (beats), howls and whistles. Early regeneratives which oscillated easily were called "blopers", and were made illegal in Europe. One preventative measure was to use a stage of RF amplification before the regenerative detector, to isolate it from the antenna. But by the mid 1920s "regens" were no longer sold by the major radio manufacturers.

[11][92][101][102][103][104][105]
Superregenerative receiver

This was a receiver invented by Edwin Armstrong in 1922 which used regeneration in a more sophisticated way, to give greater gain.\[106\][107][108][109] It was used in a few shortwave receivers in the 1930s, and is used today in a few cheap high frequency applications such as walkie-talkies and garage door openers.

In the regenerative receiver the loop gain of the feedback loop was less than one, so the tube (or other amplifying device) did not oscillate but was close to oscillation, giving large gain.\[106\] In the superregenerative receiver, the loop gain was made equal to one, so the amplifying device actually began to oscillate, but the oscillations were interrupted periodically.\[93\][110] This allowed a single tube to produce gains of over \(10^6\).

Tuned radio frequency (TRF) receiver

The tuned radio frequency (TRF) receiver, invented in 1916 by Ernst Alexanderson, improved both sensitivity and selectivity by using several stages of amplification before the detector, each with a tuned circuit, all tuned to the frequency of the station.\[11][93][110][111][112] It was very popular in quality radios during the 1920s until the superheterodyne replaced it in the 1930s. The TRF receiver consisted of these parts:

- One or more tuned radio frequency amplifier stages, each consisting of an amplifying tube or transistor and a tuned circuit. In vacuum tube radios these consisted of a tube amplifier followed by an air core interstage coupling transformer with a capacitor across one winding.
- A detector stage, in tube radios usually a triode grid-leak detector.
- One or more audio amplifier stages

A major problem of early TRF receivers was that they were complicated to tune, because each resonant circuit had to be adjusted to the frequency of the station before the radio would work.\[11][93] In later TRF receivers the tuning capacitors were linked together mechanically ("ganged") on a common shaft so they could be adjusted with one knob, but in early receivers the frequencies of the tuned circuits could not be made to "track" well enough to allow this, and each tuned circuit had its own tuning knob.\[110][113] Therefore, the knobs had to be turned simultaneously. For this reason most TRF sets had no more than three tuned RF stages.\[92][104]
A second problem was that the multiple radio frequency stages, all tuned to the same frequency, were prone to oscillate, and the parasitic oscillations mixed with the radio station's carrier in the detector, producing audible heterodynes (beat notes), whistles and moans, in the speaker. This was solved by the invention of the Neutrodyne circuit (below) and the development of the tetrode later around 1930, and better shielding between stages.

Today the TRF design is used in a few integrated (IC) receiver chips. From the standpoint of modern receivers the disadvantage of the TRF is that the gain and bandwidth of the tuned RF stages are not constant but vary as the receiver is tuned to different frequencies. Since the bandwidth of a filter with a given Q is proportional to the frequency, as the receiver is tuned to higher frequencies its bandwidth increases.

Neutrodyne receiver

The Neutrodyne receiver, invented in 1922 by Louis Hazeltine, was a TRF receiver with a "neutralizing" circuit added to each radio amplification stage to cancel the feedback to prevent the oscillations which caused the annoying whistles in the TRF. In the neutralizing circuit a capacitor fed a feedback current from the plate circuit to the grid circuit which was 180° out of phase with the feedback which caused the oscillation, canceling it. The Neutrodyne was popular until the advent of cheap tetrode tubes around 1930.

Reflex receiver

The reflex receiver, invented in 1914 by Wilhelm Schloemilch and Otto von Bronk, and rediscovered and extended to multiple tubes in 1917 by Marius Latour and William H. Priess, was a design used in some inexpensive radios of the 1920s which enjoyed a resurgence in small portable tube radios of the 1930s and again in a few of the first transistor radios in the 1950s. It is another example of an ingenious circuit invented to get the most out of a limited number of active devices. In the reflex receiver the RF signal from the tuned circuit is passed through one or more amplifying tubes or transistors, demodulated in a detector, then the resulting audio signal is passed through the same amplifier stages for audio amplification. The separate radio and audio signals present simultaneously in the amplifier do not interfere with each other since they are at different frequencies, allowing the amplifying tubes to do "double duty". In addition to single tube reflex receivers, some TRF and superheterodyne receivers had several stages "reflexed". Reflex radios were prone to a defect called "play-through" which meant that the volume of audio did not go to zero when the volume control was turned down.

Superheterodyne receiver

The superheterodyne, invented in 1918 during World War I by Edwin Armstrong, when he was in the Signal Corps, is the design used in almost all modern receivers, except a few specialized applications. It is a more complicated design than the other receivers above, and when it was invented required 6 - 9 vacuum tubes, putting it beyond the budget of most consumers, so it was initially used mainly in commercial and military communication stations. However, by the 1930s the "superhet" had replaced all the other receiver types above.

In the superheterodyne, the "heterodyne" technique invented by Reginald Fessenden is used to shift the frequency of the radio signal down to a lower "intermediate frequency" (IF), before it is processed. An unmodulated sinusoidal signal \( f_{LO} \) generated by a local oscillator (LO) is mixed with the radio signal and processed to produce an audio signal.
The first superheterodyne receiver built at Armstrong's Signal Corps laboratory in Paris during World War I. It is constructed in two sections, the mixer and local oscillator (left) and three IF amplification stages and a detector stage (right). The intermediate frequency was 75 kHz.

During the 1940s the vacuum tube superheterodyne receiver was refined into a cheap-to-manufacture form called the "All American Five" because it only required 5 tubes, which was used in almost all broadcast radios until the end of the tube era in the 1970s.

Modern transistor superheterodyne clock radio

How a superheterodyne works. The incoming radio signal (RF) from the antenna consists of a carrier frequency (dark blue) with sidebands (light blue) on either side containing the modulation. Mixing this with the local oscillator signal (LO), creates a heterodyne intermediate frequency (IF) at the difference between these frequencies. This is bandpass filtered in the IF amplifier, and the demodulator extracts the modulation (M). Because the Image frequency when mixed with the LO also creates a signal at the IF, the

incoming radio signal $f_{RF}$ from the antenna in a nonlinear vacuum tube or transistor called the "mixer". The result at the output of the mixer is a heterodyne or beat frequency at the difference between these two frequencies: $f_{IF} = f_{RF} - f_{LO}$. This lower frequency is called the intermediate frequency (IF). The IF also contains the modulation (sidebands) that was present in the original RF signal. This signal is amplified and bandpass filtered as in the TRF receiver, then is demodulated in a detector, producing an audio signal.

The superheterodyne design is superior to other receivers for these reasons:

- At the high frequencies used for radio transmission, signal processing circuitry often performs poorly. Amplifying devices have little gain, and are prone to instability and parasitic oscillation, as was seen in the TRF receiver. By shifting the signal down to a lower intermediate frequency, the amplification, filtering and detection can be done at a more convenient frequency where the electronics works better. Armstrong invented the superheterodyne to receive high frequencies around 5 MHz, which were above the 1 MHz frequency range of most vacuum tubes of the time. Today active devices still perform poorly at the microwave frequencies which modern telecommunication links use.

- The different frequencies of different stations are all converted to the same frequency, the IF, for filtering, so the bandwidth and gain of the receiver is constant over its frequency range. To tune the receiver to a different frequency, only the frequency of the local oscillator $f_{LO}$ needs to be changed. The rest of the receiver after the mixer operates at a fixed frequency, the IF. All the other receivers above require that the bandpass filter (tuned circuit) be adjustable to different frequencies. Since the bandwidth of the filter is proportional to the frequency, the bandwidth of the receiver increases as it is tuned to higher frequencies. It is also easier to build tunable oscillators than tunable electronic filters. In an analog television set, for example, if the superheterodyne design was not used, all the complicated filters that separate out the luminance carrier, chroma carrier and the audio subcarrier from the television signal would have to be made adjustable, and retuned each time the channel was changed.

- The total amplification of the receiver is divided between three amplifiers at different frequencies; the RF, IF, and audio amplifier. This reduces problems with feedback and parasitic oscillations that are encountered in receivers where most of the amplifier stages operate at the same frequency, as in the TRF receiver.

- The most important advantage is that better selectivity can be achieved by doing the filtering at the lower intermediate frequency. One of the most important parameters of a receiver is its bandwidth, the band of frequencies it accepts. In order to reject nearby interfering stations or noise, a narrow bandwidth is required. In all known filtering techniques, the bandwidth of the filter increases in proportion with the frequency, so by performing the filtering at the lower IF, rather than the frequency of the original radio signal $f_{RF}$, a narrower bandwidth can be achieved. Modern FM and television broadcasting, cellphones and other communications services, with their narrow channel widths, would be impossible without the superheterodyne.
The RF filter on the front end of the receiver is needed to prevent interference from any radio signals at the image frequency. When the incoming radio signal $f_{RF}$ is mixed with

the local oscillator signal $f_{LO}$, two new frequencies (heterodynes) are created \(^{[129]}\) one at the sum $f_{RF} + f_{LO}$, and one at the difference $f_{RF} - f_{LO}$ of the frequencies. Because of this, without an input filter the receiver can receive incoming RF signals at two different frequencies, \(^{[116][126][129][130]}\) one above the LO frequency: $f_{RF1} = f_{LO} + f_{IF}$, and one below the LO frequency: $f_{RF2} = f_{LO} - f_{IF}$. The receiver can be designed to receive on either of these two frequencies; the former is called low-side injection because the LO signal frequency is below the received frequency, while the latter is high-side injection. \(^{[129]}\) Whichsoever is chosen, the other frequency is called the image frequency. If the receiver is receiving a signal on one of these frequencies, for example $f_{RF1}$, any other radio station or radio noise on the other frequency $f_{RF2}$ will be received also, interfering with the desired signal. Therefore, the superheterodyne requires a bandpass filter on the input to reject this image frequency. \(^{[126]}\) This image rejection filter does not need great selectivity, but as the receiver is tuned to different frequencies it must "track" in tandem with the LO.

**Double-conversion superheterodyne:** In choosing the intermediate frequency (IF) used in a superhet, there is a tradeoff between image rejection and selectivity. \(^{[126][129]}\) Since the separation between the desired frequency and the image is $2f_{IF}$, the higher the IF, the farther apart these two frequencies are, and the better the image rejection. However, the lower the IF, the narrower the bandwidth the IF filter can achieve, and the better the selectivity. In order to achieve both good image rejection and selectivity, many modern superhet receivers use two intermediate frequencies; this is called a dual-conversion or double-conversion superheterodyne. The incoming RF signal is first mixed with one local oscillator signal in the first mixer to convert it to a high IF frequency, to allow efficient filtering out of the image frequency, then this first IF is mixed with a second local oscillator signal in a second mixer to convert it to a low IF frequency for good bandpass filtering. Some receivers even use triple-conversion.

**Terminology used with superheterodyne receivers:**

- **RF front end** - refers to all the components of the receiver up to and including the mixer; all the parts that process the signal at the original incoming radio frequency.
- **Converter** - in many superhet circuits the same tube or transistor functions as both the local oscillator and the mixer. This is called a converter.
- **IF strip** - The IF filter and amplifier usually consists of several tuned amplifier stages as in the TRF receiver. This is called the IF strip.
- **First detector, second detector** - the mixer tube or transistor is sometimes called the "first detector", while the demodulator that extracts the modulation from the IF signal is called the "second detector". In a dual-conversion superhet there are two mixers, so the demodulator is called the "third detector".

To receive radiotelegraph (CW) or single-sideband modulation (SSB), a beat frequency oscillator (BFO) is used.

**Semiconductor era**

The invention of the transistor in 1947 revolutionized radio technology, making truly portable receivers possible, beginning with transistor radios in the late 1950s. Although portable vacuum tube radios were made, tubes were bulky and inefficient, consuming large amounts of power and requiring several large batteries to produce the filament and plate voltage. Transistors did not require a heated filament, reducing power consumption, and were smaller and much less fragile than vacuum tubes.

The development of integrated circuits (ICs) in the 1970s created another revolution, allowing an entire radio receiver to be put on a chip. ICs reversed the economics of vacuum tube radios; since the marginal cost of adding additional amplifying devices (transistors) to the chip was essentially zero, the size and cost of the receiver was dependent not on the active components as with vacuum tubes, but on the passive components; inductors and capacitors, which could not be integrated easily on the chip. \(^{[4]}\) As a result, the current trend in receivers is to use digital circuitry on the chip to do functions that were formerly done by analog circuits which require passive components. In a digital receiver the IF signal is sampled and digitized, and the bandpass filtering
and detection functions are performed by digital signal processing (DSP) on the chip. Another benefit of DSP is that the properties of the receiver; channel frequency, bandwidth, gain, etc. can be dynamically changed by software to react to changes in the environment; these systems are known as software-defined radios or cognitive radio.

In modern wireless technology, radio transmitters and receivers are embedded in portable digital devices such as laptops, cell phones, GPS receivers, Bluetooth headsets, wireless routers and work automatically, in the background, to keep the device in touch with other devices, without wires. They transmit binary digital data by modulation methods such as frequency shift keying (FSK) allowing portable digital devices to communicate via wireless networks.

Digital technologies

Many of the functions performed by analogue electronics can be performed by software instead. The benefit is that software is not affected by temperature, physical variables, electronic noise and manufacturing defects. For really high-performance receivers, such as satellite communications receivers and military/naval receivers, two-stage ("double conversion") and even three-stage ("triple conversion") superheterodyne processing is frequently used. Single-conversion receivers are rather simple-minded in their nature.

DSP technology

DSP technology, short for digital signal processing, is the use of digital means to process signals and is coming into wide use in modern shortwave receivers. It is the basis of many areas of modern technology including cell phones, CD players, video recorders and computers. A digital signal is essentially a stream or sequence of numbers that relay a message through some sort of medium such as a wire. The primary benefit of DSP hardware in shortwave receivers is the ability to tailor the bandwidth of the receiver to current reception conditions and to the type of signal being listened to. A typical analog only receiver may have a limited number of fixed bandwidths, or only one, but a DSP receiver may have 40 or more individually selectable filters.

PC controlled radio receivers

"PC radios", or radios that are designed to be controlled by a standard PC are controlled by specialized PC software using a serial port connected to the radio. A "PC radio" may not have a front-panel at all, and may be designed exclusively for computer control, which reduces cost.

Some PC radios have the great advantage of being field upgradable by the owner. New versions of the DSP firmware can be downloaded from the manufacturer's web site and uploaded into the flash memory of the radio. The manufacturer can then in effect add new features to the radio over time, such as adding new filters, DSP noise reduction, or simply to correct bugs.

A full-featured radio control program allows for scanning and a host of other functions and, in particular, integration of databases in real-time, like a "TV-Guide" type capability. This is particularly helpful in locating all transmissions on all frequencies of a particular broadcaster, at any given time. Some control software designers have even integrated Google Earth to the shortwave databases, so it is possible to "fly" to a given transmitter site location with a click of a mouse. In many cases the user is able to see the transmitting antennas where the signal is originating from.

Radio control software

The field of software control of PC radios has grown rapidly in the last several years, with developers making a number of advances. Since the Graphical User Interface or GUI interface PC to the radio has unlimited flexibility, any number of new features can be added by the software designer. Features that can be found in advanced control software programs today include a band table, GUI controls corresponding to traditional radio controls, local time clock and a UTC clock, signal strength meter, an ILG database for shortwave listening with lookup capability, scanning capability, text-to-speech interface, and integrated Conference Server.
Software-defined radios

The next level in radio / software integration are so-called pure "software defined radios". The distinction here is that all filtering, modulation and signal manipulation is done in software, usually by a PC soundcard or by a dedicated piece of DSP hardware. There may be a minimal RF front-end or traditional radio that supplies an IF to the SDR. SDR's can go far beyond the usual demodulation capability of typical, and even high-end DSP shortwave radios. They can for example, record large swaths of the radio spectrum to a hard drive for "playback" at a later date. The same SDR that one minute is demodulating a simple AM broadcast may also be able to decode an HDTV broadcast in the next. A well known open-source project called GNU Radio is dedicated to evolving a high-performance SDR. All the source code for this SDR is freely downloadable and modifiable by anyone.

See also

- Car audio
- Crystal radio receiver
- Demodulation
- Dielectric wireless receiver
- Digital Audio Broadcast (DAB)
- Direct conversion receiver
- Minimum detectable signal
- Neutrodyne
- Radar
- Radar warning receiver
- Radiogram (furniture)
- Receiver (information theory)
- Regenerative radio receiver
- Selectivity (electronic)
- Sensitivity (electronics)
- Noise (electronics)
- Distortion
- Satnav
- Superheterodyne receiver
- Table radio
- Telecommunication
- Television receive-only
- Transistor radio
- Tuned radio frequency receiver
- Tuner (radio)

References

Radio receiver - Wikipedia

   (https://archive.org/stream/radiosconquestof00mcnirich/page/25-p
64. MARCONI WIRELESS TELEGRAPH CO. OF AMERICA v. UNITED STATES. UNITED STATES v. MARCONI WIRELESS TELEGRAPH CO. OF AMERICA. 320 U.S. 1 (63 S.Ct. 1393, 87 L.Ed. 1731) (https://www.law.cornell.edu/supremecourt/text/320/1#fn3)
83. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 267-270
84. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 341-344
89. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 267-270
90. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 341-344
91. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 267-270
92. McNicol, Donald (1946) *Radio's Conquest of Space*, p. 341-344
Further reading

- McNicol, Donald (1946) Radio's Conquest of Space, p. 283-284 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i
- McNicol, Donald (1946) Radio's Conquest of Space, p. 272-278 (https://archive.org/stream/radiosconquestof00mcnirich/page/2i