Some Recent Developments in the Art of Receiver Technology: A Selected History on Receiver Innovations over the Last 100 Years

By **Brad Brannon**

Share on 🐦 ⓕ 🔗

Part 1: The Early Days

While there were many contributors to the early days of wireless, Guglielmo Marconi is one of the more prominent. While he is known for his wireless technology, many people are less familiar with the business he built around wireless technology at the turn of the 19th century. For about 20 years after the start of the 1900s, he built a critical business that launched the world of wireless toward what we have today.

Figure 1. Marconi demonstrating his technology.

While his commercialized technology was not the most up to date, it was good enough despite rapid technological changes because he figured out how to use the technology available to him to enable a new industry. Marconi set out to deploy a worldwide network capable of sending and relaying messages wirelessly at a time when the world was in turmoil at the end of colonialism and the wars and disasters that pockmarked the start of the 1900s, including the sinking of the RMS Titanic in April of 1912. The role that wireless played in both the rescue of survivors and the dissemination of the news of that accident reinforced the importance of this fledgling technology. The importance of wireless technology wasn't missed by either the public or the military, notably Joseph Daniels, who later became the secretary of the U.S. Navy. In the U.S. and elsewhere, leaders such as Daniels felt that the military should nationalize radio to ensure that they had access to it during wartime. It must be kept in mind that during this period, the only usable spectrum was below 200 kHz or so. At least for a while things moved this direction, but after World War I, the government's control of wireless weakened, but not before the formation of the government sanctioned monopoly that created the Radio Corporation of America (RCA).[1]

By our expectations, the radios of Marconi's days were quite primitive. The transmitters employed spark gap devices (only later did they employ mechanical alternators) to generate the RF, but on the receiving end, the systems were fully passive and consisted of an antenna, resonant LC tuner, and some sort of detector. These detectors will be covered shortly, but they were either mechanical, chemical, or organic. Some of these systems employed a battery simply to bias them, but not to provide any circuit gain as we might recognize today. The output from these systems was supplied to some sort of headset to convert the signal to audio, which was always very weak and just a simple click or buzz at best.

Because these systems provided no gain on the receiving end, range was determined by the amount of transmitted power, the quality of the receiver, the experience of the operator to adjust it, and, of course, atmospheric conditions. What Marconi realized was that given a reasonably predictable range, a network of stations could be built that could be utilized to reliably communicate information across both continents and oceans. This included installations both on land and at sea. Marconi set off to install his wireless stations across the globe and at sea, both on passenger ships and cargo ships. By installing systems on seafaring ships, he not only enabled them to communicate with their commercial interests on shore, but this also allowed Marconi to fill critical gaps in his network by providing relay and redundancy where needed.

One of the technologies that Marconi possessed was that of early vacuum tubes. John Ambrose Fleming, the recognized inventor of the vacuum tube, worked for the Marconi Corporation, but Fleming and Marconi's analysis at the time was that their existing technology was sufficient at detecting radio signals. Furthermore, they felt that the benefit of his discovery wasn't worth the additional expense or the batteries to run the valve tubes. Marconi already possessed several technologies that could detect a signal and didn't require the high power to run the filament and plate that a tube needed. Thus, they passed on this technology initially.
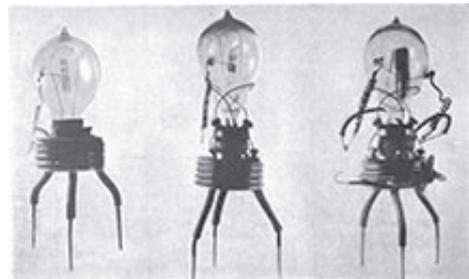
Figure 2. The first Fleming valve prototypes.

However, the so-called Father of Radio, Lee de Forest, took this technology and realized what the potential was. By insertion of a screen grid between the filament and the plate, he could not only rectify a signal, he could also control the amount of current in the plate. This enabled amplification. Even though there is evidence that he didn't understand how his Audion tube worked, he did realize the potential and did his best to capitalize on his invention both as a technology and as a value-added service similar to that created by Marconi. Through various business ventures, de Forest attempted to both manufacture and sell his vacuum tubes and to set up wireless networks like Marconi's. However, these ventures were doomed to failure not because of bad technology, but because de Forest's business partners were often less than honest, and often left him standing alone answering for the wrong doing of others. In the end, de Forest had to sell the rights to his own invention for others to profit from its capabilities.
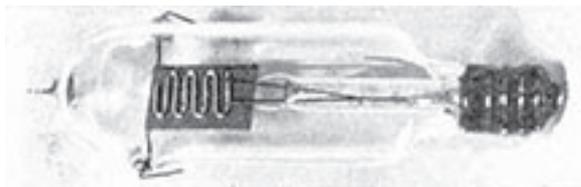
Figure 3. First audio triode de Forest tube.

One of those who early on realized the vacuum tube's possibilities was Edwin Armstrong. While he was still in high school, a family friend gave him one of de Forest's tubes to play with. Armstrong had already developed a reputation as an expert in wireless and, having built his own wireless station in the family home, quickly figured out how to utilize the device to develop a better receiver. While in college he continued to develop this technology, and developed the regenerative receiver that provided superior performance compared to the passive systems employed by all wireless stations of the day.

David Sarnoff was a senior figure in the American Marconi Corporation. His rise within the company was a direct result of a long-cultivated relationship with Marconi himself and a dedicated work ethic. Sarnoff began his career as an errand boy for AMC and by chance met Marconi on one of his visits to America. Sarnoff so impressed Marconi that he enabled him to rise to power within the company, with Sarnoff eventually taking senior leadership in both AMC and later in RCA. While visiting engineering labs around New York, he chanced across Armstrong. Armstrong's knowledge of wireless and the capabilities of his regenerative receiver helped forge a long-term professional and personal relationship between the two.

When WWI came, Armstrong felt the call of duty to enlist. By this time, however, he had developed a reputation as a wireless expert and, instead of being assigned combat duties, was assigned the role of inspecting and installing radios for combat services throughout France. His duties allowed him access to equipment and labs, as well as various technologies that enabled him to continue his research activities on the side. During an air raid in early 1918, he made a series of discoveries that led him to synthesize the superheterodyne receiver. Throughout 1918 he developed his concept and, by November, met with a close group of friends to demonstrate a prototype of the superheterodyne radio. They were impressed and urged him to continue his developments. By the end of 1918 the war was ending and, prior to returning to the U.S., Armstrong filed for a French patent on December 30, 1918. On returning to the U.S., he spent a few weeks recovering from an illness that delayed his filing for a U.S. patent. Eventually, he filed a U.S. patent for the superheterodyne receiver on February 8, 1919.

While Marconi's vision for wireless focused only on commercial information carried by telegraphy between two parties, Sarnoff had a much broader vision—to send a signal to many parties. Sarnoff's vision was not shared broadly early on, but eventually others realized that this new technology offered a means by which news and entertainment could easily be delivered over great distances, including the rural reaches of America. To help drive his vision, Sarnoff and his team conceived of broadcasting the Dempsey vs. Carpentier boxing match on July 2, 1921. The success of this broadcast enabled others to see the potential of broadcast radio as we know it today.

However, the real challenge of the time was technical. Early radios were difficult to use and didn't function very well. This is where the story continues for Armstrong, Sarnoff, and the Radio Corporation of America. Through the relationship developed earlier and the patents that RCA had acquired, including that of the superheterodyne, radio technology had been simplified enough to make it both portable and easy enough for anyone to use. From a technology point of view, the superheterodyne architecture was key to this success and this largely remains true today.

Figure 4. Edwin Armstrong and his wife Marion on their honeymoon with the first portable radio.

Detectors

A radio must have some way of producing an output that conveys meaningful information. In the early days, this was a sympathetic spark created in the receiving loop antenna. It was quickly realized that a more sensitive way of converting the radiated energy into a meaningful signal was required. Early technology was quite limited and often leveraged a wide range of properties including chemical, mechanical, and electrical.

In the very beginning, one of the first detectors used was called a coherer detector and was based on the discoveries of a Frenchman named Édouard Branly. The coherer consisted of two metal plates closely spaced with a supply of metal filings. As the RF signal presented itself to the plates, the metal filings adhered to the plates, closing an electrical circuit. This worked quite well for the detection, but once the RF signal was removed, the filings tended to remain attached to the plates. To solve this problem, some sort of tapper was arranged to hit the side of the device to force the filings to dislodge. This crude detector was effective yet bulky to use and operate because of this. Despite this, it was in use as late as 1907.

Figure 5. Coherer device.
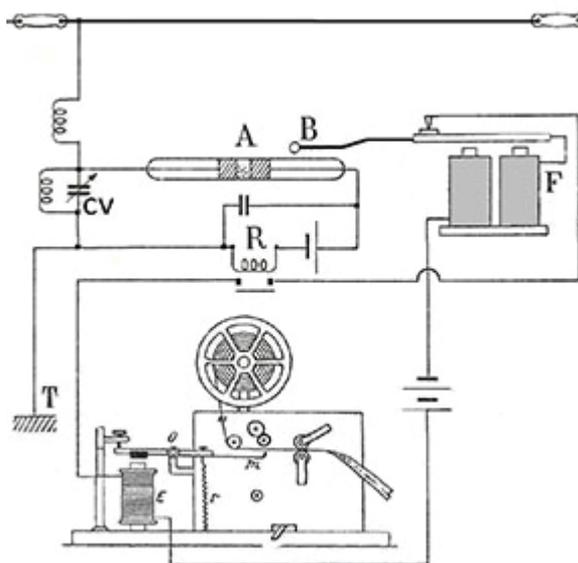
Figure 6. Coherer receiver schematic.[2]

A more practical solution was the electrolytic detector. This device consisted of a very fine platinum wire immersed in a solution of sulfuric or nitric acid. A battery was used to bias this circuit just to the point of electrolysis. This formed gas bubbles on the surface of the platinum wire, causing the current to drop. If the RF current was coupled into this circuit, it would modulate the electrolysis and cause the current to vary in relation to the strength of the coupled RF signal. This technique was developed by Fessenden and in common use from 1903 until 1913. One variation of this developed by de Forest was the responder, which consisted of two metal plates immersed in a solution of lead peroxide.
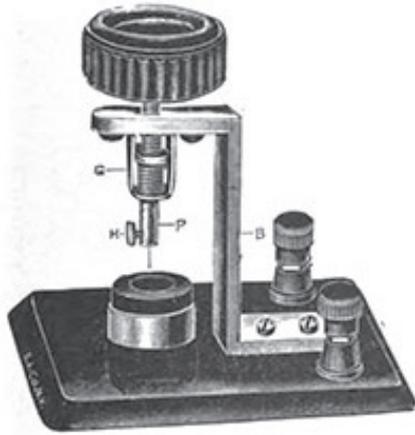
Figure 7. Electrolytic detector.
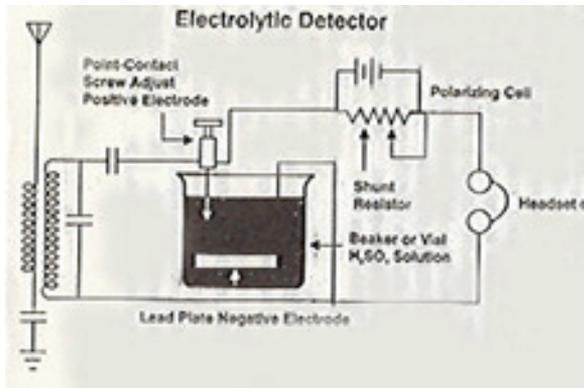
Figure 8. Electrolytic radio receiver.

Marconi preferred another approach called the magnetic detector. These devices were affectionately referred to as a Maggie by their users. They worked by creating an endless loop of steel wire that was magnetized by permanent magnets while being circularly rotated. This magnetized portion of the wire was passed through a loop of wire connected to an antenna. The RF field in this coil would demagnetize the wire according to the received signal level present. The variations in the magnetic field in the wire was then picked up by another coil that was connected to an ear piece that provided an audible version of the RF signal. This approach was used by all Marconi installations until 1912, including on the RMS Titanic.

Figure 9. Magnetic detector as Marconi would have used.[3]
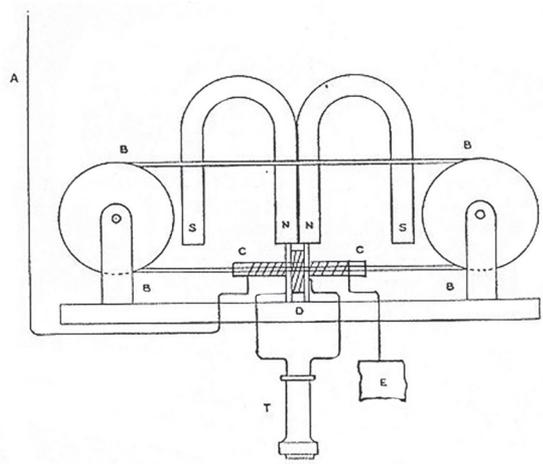
Figure 10. Schematic of a magnetic detector radio.

Another common type detector was the crystal detector, which remained popular until around 1925. This type of popular device was often referred to as a cat whisker and was basically an early semiconductor junction fashioned out of various types of mineral. Typical minerals included galena (PbS), iron pyrite (FeS2), molybdenite (MoS2), and carborundum (SiC). Small samples of these rocks were fashioned in a metal cup with a fine wire making a point contact on the rock. This contact could be moved and placed at various locations on the rock in search of the best operation. Crystal radios are still available today; the circuit is identical to that available 100 years ago, with the exception that a manufactured semiconductor diode replaces the cat whisker. One advantage of crystal detectors is that these devices provide more of a linear detection, which became important as AM broadcast began. This made voice communication possible, while earlier transmissions were only sent by Morse code.

Figure 11. Galena cat whisker detector.
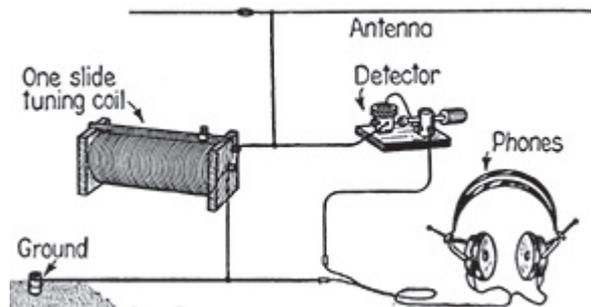
Figure 12. Typical crystal radio schematic.[4]

Another type of detector was created in 1904 by an engineer working for Marconi. John Ambrose Fleming discovered that by adding a plate to an Edison incandescent bulb he had created a rectifier or valve as it is often called. Marconi and Ambrose believed that their existing solution for

detection, typically a Maggie, was better than that offered by the Fleming valve and they temporarily discontinued their efforts to find a better solution until after 1912. Others, including de Forest, however, did see the immediate value and picked up where Fleming and Marconi left off by adding a screen grid between the filament and plate. This work was patented and published in 1906. While de Forest realized the value of his invention for improved radios, he was not able to capitalize on this, partly due to his business partner's misdeeds and partly due to various infringement cases against his patent.

Part 2: Receiver Architectures

One key point early radio pioneers like de Forest and Armstrong understood was that their successes were determined by a solid, reliable detector; in the early days, this was largely the wireless operator whose technical and auditory skills made it possible. However, as the industry grew, other aspects became important, such as including linearity and bandwidth.

In 1912, to address these issues, de Forest figured out regeneration and how a receiver could benefit from this technique. At nearly the same time, Armstrong made similar discoveries, and he noted that if energy was coupled from the plate circuit back into the screen tuner, significant amplification occurred as the amplifier response peaked prior to free oscillation. These discoveries set off a multidecade long patent dispute as each inventor claimed their invention came first.

Regardless, the key advantage of the regenerative receiver was that, in addition to the very high levels of gain achieved, the receiver facilitated connecting the output to a speaker as opposed to a small headphone with a weak audio output as had previously been used. Armstrong noted that, with this arrangement, he could easily copy Marconi's installation in Ireland from his New York lab, whereas Marconi typically required a relay station to achieve transatlantic coverage. After he was satisfied with his results, Armstrong invited Sarnoff to his lab to share his discoveries. With his regenerative setup, they spent the night DXing and received signals from the West Coast and into the Pacific with ease. This was a major enhancement for detector technology. The biggest challenge for the regenerative receivers was adjusting the feedback for proper operation; a challenging task even for an experienced operator. As early models of the regenerative and super-regenerative radios were put into productions, this challenge became apparent and required resolution before radio technology could be put into widespread usage.

World War I eventually drew the U.S. into the engagement and Armstrong received duty in France, where he was responsible for installing technology in the field. This afforded him the opportunity to continue his research where he conceived of the superheterodyne architecture in February of 1918 after working with colleagues in both France and Britain. Eventually, this architecture resolved many of the tedious adjustments required in prior architectures like the super-regenerative type without sacrificing performance.

Armstrong continued to develop the superheterodyne architecture throughout 1918, which solved many of the challenges of the regenerative and the super-regenerative receivers. This development enabled easy to operate radios consistent with those in production today. While a superheterodyne receiver is not strictly a detector, it does facilitate better, more consistent detection by including gain and additional selectivity, and by presenting a fixed IF regardless of the RF frequency being monitored. This allows the detector to be optimized without concern of degradation as a function of the desired RF frequency, which was a huge challenge of early radio and continues to challenge radio designers today, albeit at much higher frequencies even as we continue to explore new architectures including zero-IF and direct RF sampling.
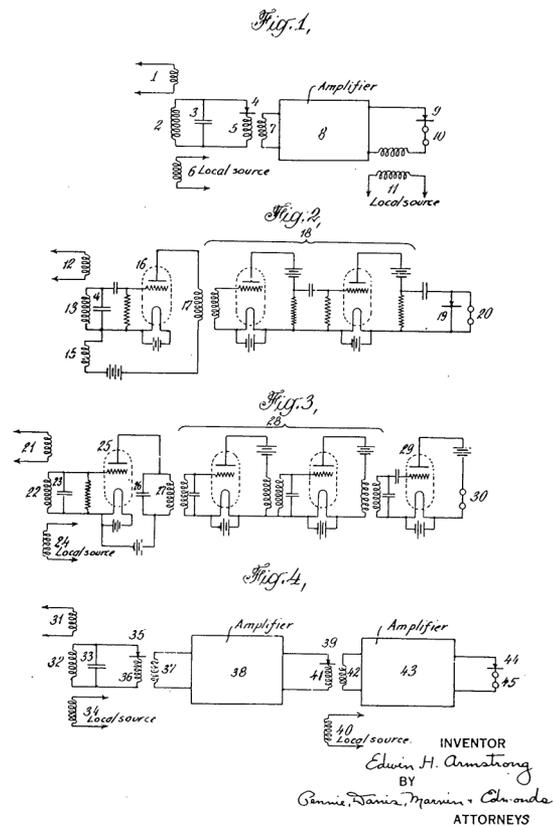
Figure 13. Superheterodyne patent figures.

These advantages have cemented the importance of heterodyne architectures, and this continues today. While the implementing technology has moved from tube to transistor to integrated circuit, the architecture remains key to many modern systems.

Outside of the shifting of technology types, little changed in radio architectures until the 1970s, with the advent of general-purpose DSPs and FPGAs. Detector functions moved from linear detector elements like diodes, discriminators, and PLLs to analog-to-digital converters followed by digital signal processing. This enabled significantly more capability not possible with older technology. While data converters followed by DSP can and do perform traditional AM and FM[5] demodulation, use of digital processing techniques enables complex digital demodulation used widely for digital television, HD Radio® in the United States, and DAB in Europe and other regions around the world.

In early digital systems, the IF stage was typically converted to a baseband signal with an I/Q demodulator and then digitized by dual low frequency ADCs, as shown in Figure 14. These early ADCs were relatively low bandwidth and therefore radios tended to be narrow-band systems. While these systems are workable for low bandwidth systems, they suffered from quadrature mismatches that caused image rejection issues that had to be corrected for by analog and later digital techniques. Because early systems were not highly integrated, it was difficult to maintain balance between I/Q, which resulted in image errors (quadrature). This was complicated due to
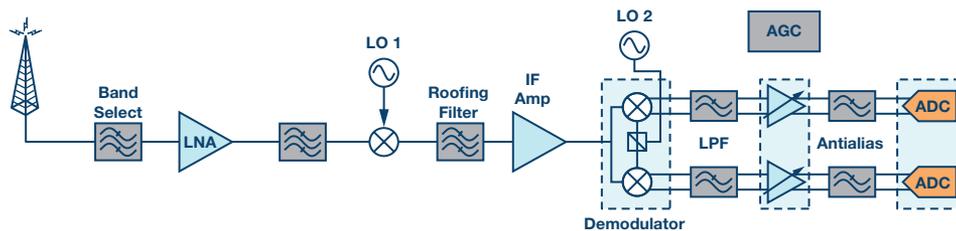
Figure 14. Dual conversion baseband sampling.
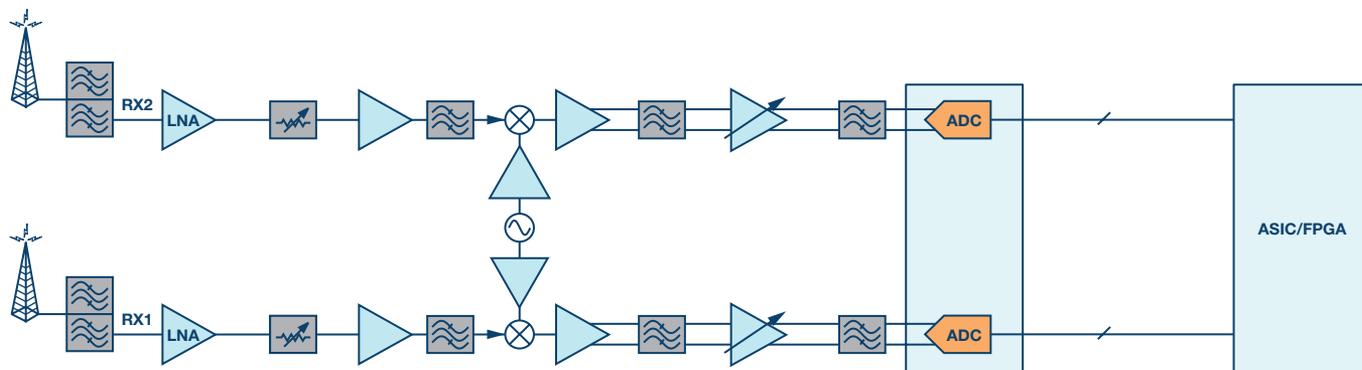
Figure 15. Typical IF sampling architecture.

shifts over time and temperature that had to be carefully accounted for. Even in highly integrated systems, I/Q balance is typically limited to 40 dB or worse image rejection without some sort of corrective algorithm.

By the mid-'90s, converter technology began to improve enough that base-band I/Q sampling could be replaced with IF sampling. This had several advantages. First, the demodulator and baseband converter pair could be eliminated and replaced with a single ADC saving power and board space. More importantly, the errors associated with analog I/Q extraction could be eliminated. Of course, complex data was still required for DSP processing, but it could easily be extracted digitally by using digital downconverters (DDCs) like the AD6624 that provide perfect quadrature that doesn't drift over time or temperature.

Initially these IF sampling converters were narrow band, but in the late '90s, wideband IF sampling converters became available, including devices like the AD9042 and AD6645. These new devices could sample IF frequencies as high as 200 MHz and provide a signal bandwidth of up to 35 MHz. This became interesting enough that many high performance receivers began adopting IF sampling to both simplify the radio and to improve performance. One of the many advantages of this technique is that one receiver signal path could process multiple RF carriers.[6] This had the effect of allowing one radio to replace many analog narrow-band radios and greatly reduce the cost of ownership in many telecommunications applications. Any application that processed multiple independent (or dependent) RF signals could benefit from this type of architecture, which allowed for a reduction in cost, size, and complexity. Individual RF carriers are easily sorted out in the digital data stream where they could be independently processed as required. Each signal could be modulated differently with unique information or the signal bandwidth could be widened to increase the data throughput. Integrated mixer technology, including the ADRF6612 and ADRF6655, continues to move IF sampling heterodyne radios forward by providing highly integrated and low cost solutions when combined with new IF sampling converters like the AD9684 and AD9694. These new ADCs include digital downconverters (DDCs) that not only digitally filter unneeded spectrum, but that also digitally extract out the I/Q components.

Side by Side: Then and Now

Armstrong's patent[7] states that "It is well known that all detectors rapidly lose their sensitiveness as the strength of the received signal is decreased, and that when the strength of the high frequency oscillation falls below a certain point, the response of a detector becomes so feeble that it is impossible to receive signals." Armstrong claimed that as amplitude fell or as frequency increased, detector sensitivity was reduced. He and others sought a method to extend the usefulness of radio to higher frequencies and to improve overall performance.

Based on earlier work with tubes such as the Audion tube and regeneration, Armstrong realized that he could convert the incoming frequency to one that worked more efficiently with the detectors available. Furthermore, gain could be applied to increase not only the RF signal level, but the audio signal level provided to the user.

Figure 16 shows one of the patent's diagrams, which "illustrates in detail the utilization of [Armstrong's] method using a tuned amplifier system wherein 21 is the source of the incoming oscillations (signals), and a vacuum tube rectifying system 22-23-25 converts the combined oscillations of the incoming and those from the separate heterodyne 24 (local oscillator). The circuit 26-27 is tuned to the converted combination of the two oscillations (desired mixer product). A multitube high frequency amplifier 28 amplifies the resulting energy heterodyned and detected by the vacuum tube system 29 and indicated by the telephones 30."[7] By using this method, Armstrong was able to take the RF energy and shift the frequency to one that could easily and efficiently be detected as well as provide sufficient amplification for a comfortable audio level. In his patent, he goes on to show that multiple stages of heterodyning can be applied, which has the advantage of providing additional selectivity and higher levels of gain without the concern of uncontrolled feedback causing oscillation—a problem that plagued earlier radio architectures such as the regenerative receiver.
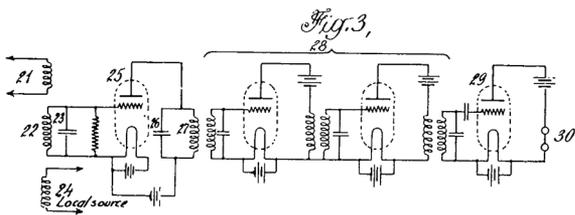
Figure 16. Armstrong's superhet diagram.

The following figures help to better compare tube technology to a contemporary implementation and show how modern designs remain similar to the original design proposed 100 years ago.

Figure 17 shows both circuits placed side by side. The first tube stage, according to Armstrong's patent, consists of a vacuum tube rectifying system. This first stage combines the mixing of the desired signal with the LO by taking advantage of the rectifying properties of the tube to produce the typical mixing products. Armstrong suggested 10 MHz (shown in Figure 18) as an RF because this was beyond what detectors of his time could respond to and because it represented a technical challenge to him during the period when he was developing the superheterodyne receiver. Modern receivers typically include at least one RF amplifier prior to the mixer to provide lower noise and better sensitivity as is shown in the lower signal chain. These devices are typically very low noise FET designs optimized for the frequency range of operation. The only fundamental difference between what Armstrong originally patented and modern designs is the separate RF amplifier placed before the mixer. By WW2, it was common to find tube type designs with front-end amplifiers equivalent to today's FET front ends.

He suggested that this incoming RF signal could be combined with an LO of perhaps 10.1 MHz to produce a new tone at 0.1 MHz during the first stage. We recognize this as the sum and difference products of a typical mixer, as shown in Figure 19. In the tube schematic in Figure 18, the LO was coupled directly into the input circuitry where the nonlinear behavior of the tube produced these products. One challenge this original design would have posed would have been unintended radiation of the LO by

way of the direct coupling to the antenna. Contemporary designs are less susceptible to this radiation, although not completely, because, as shown in Figure 19, the LO is coupled into a mixer isolated from the input by the front-end amplifier. One improvement Armstrong proposed was that Amplifier 1 could also be used as the local oscillator in addition to the detector by taking advantage of feedback from the plate into the grid circuit similar to what he and de Forest had accomplished with the regenerative receiver. This would have created a compact front-end function. In today's circuitry, the mixer, local oscillator, and RF and IF amplifiers are often included on a single IC. These devices are widely available for many different applications from consumer to industrial needs.

For both the tube and monolithic front ends, the mixing process produces sum and differences between the RF and LO. In Armstrong's case, this meant 0.1 MHz and 20.1 MHz. In addition, it is common to have both RF and LO leakage to the output as well. The unwanted terms created by the mixer must be filtered out in order to receive the desired signal. Since the bandwidth of the detector was limited, Armstrong focused in the difference term, 100 kHz. It is likely that his 2-stage IF amplifier provided some filtering of the other terms in addition to the resonant L-C structures he included. A contemporary IF amplifier will include some sort of IF filter as well. Figure 19 shows a basic LC filter, but, often, some form of high Q filter is used. Narrow-band radios often use quartz or ceramic filters for the IF stage; wider band designs often take advantage of SAW or BAW depending on the requirements. Often, this filter is referred to as a roofing filter and is used to protect following stages from strong out-of-band signals.

With a well filtered and strong IF signal, Armstrong could now easily detect what were once weak RF signals outside of the bandwidth of his detector. Now at an IF, they easily matched what detectors were capable of. In the case of the tube, these signals were rectified and then amplified so that they could drive a speaker directly, at least for amplitude modulated signals. In contemporary receivers, an analog-to-digital converter samples the analog IF and produces a digital equivalent, which is then processed digitally (including demodulation). In the case of an audio application, it can then be converted back to analog with a digital-to-analog converter to drive a speaker if necessary.
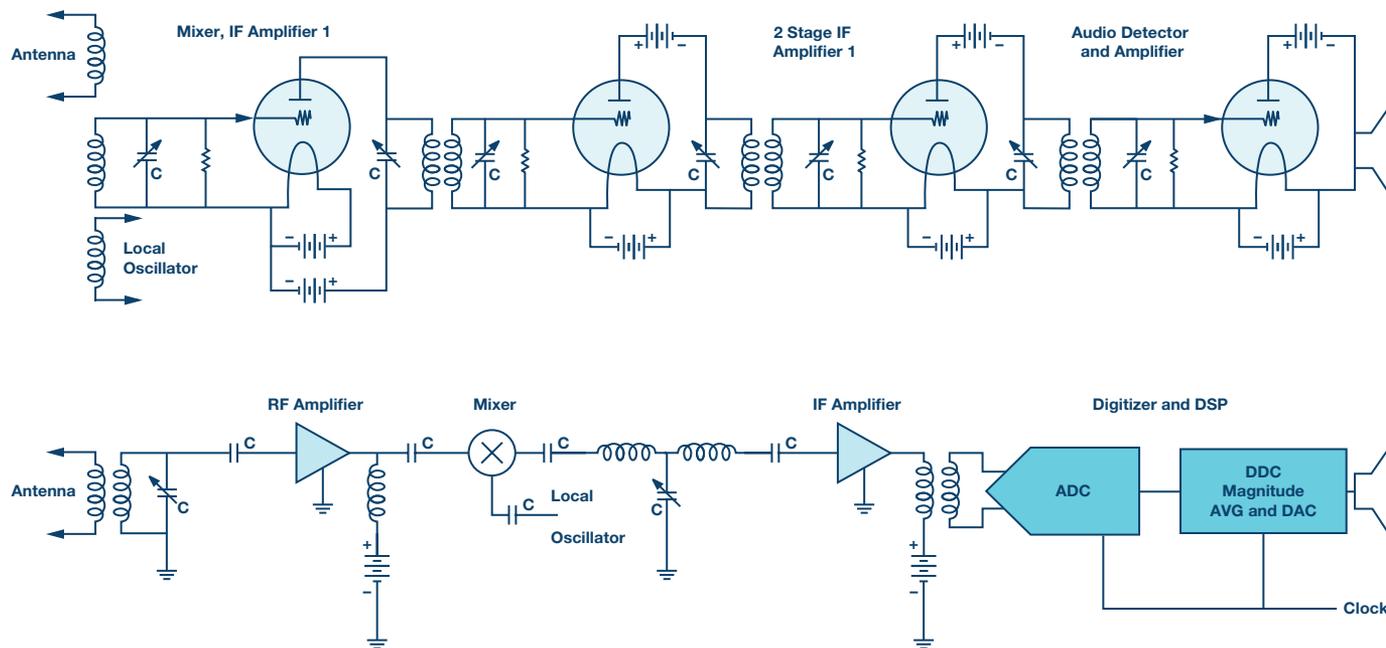
Figure 17. Tube vs. contemporary superhet.
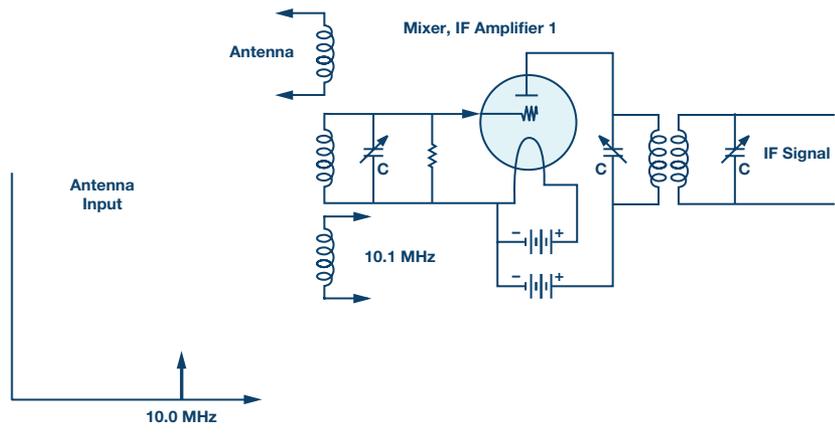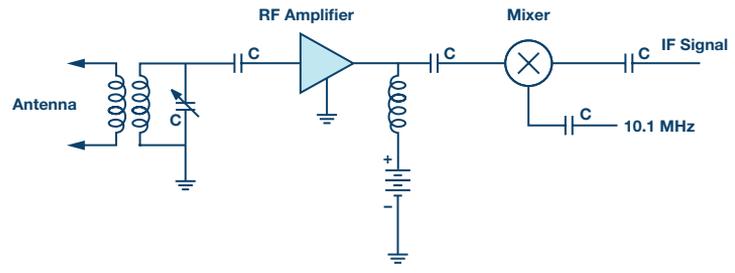
Figure 18. (a) Tube front end, (b) front end.
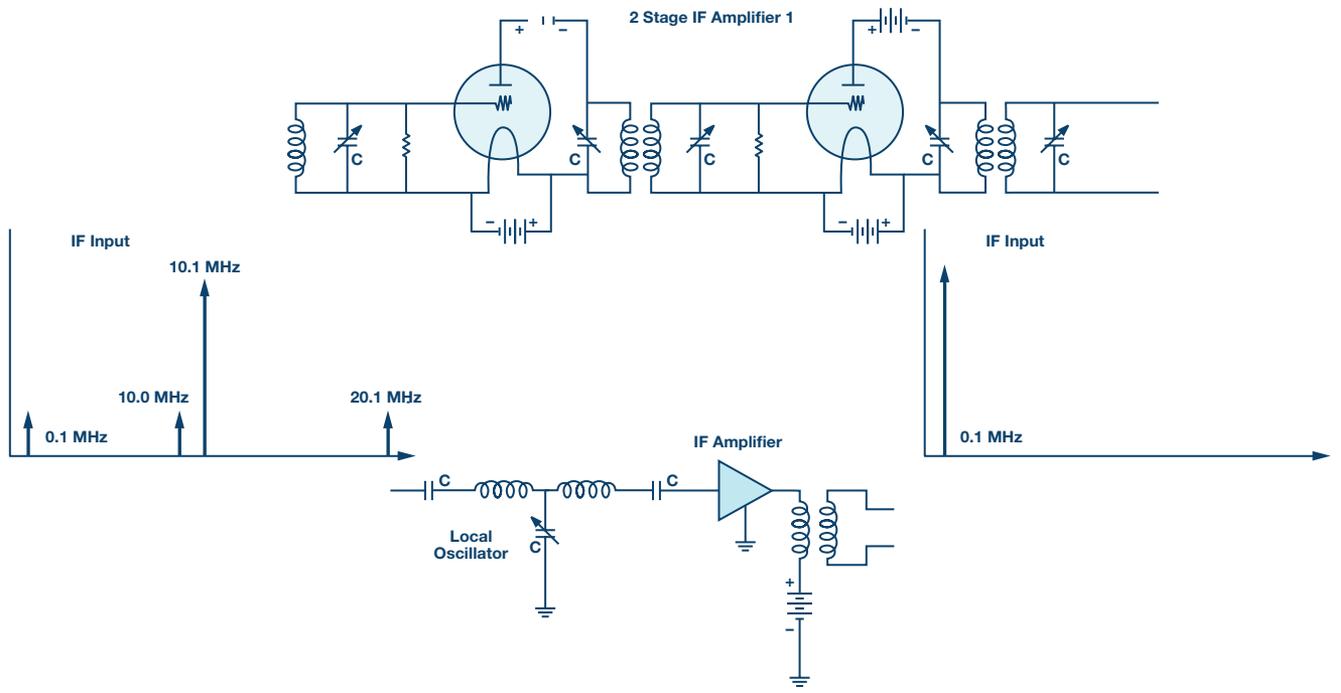
Figure 19. IF amplifier stage.

Audio Detector and Amplifier

Digitizer and DSP
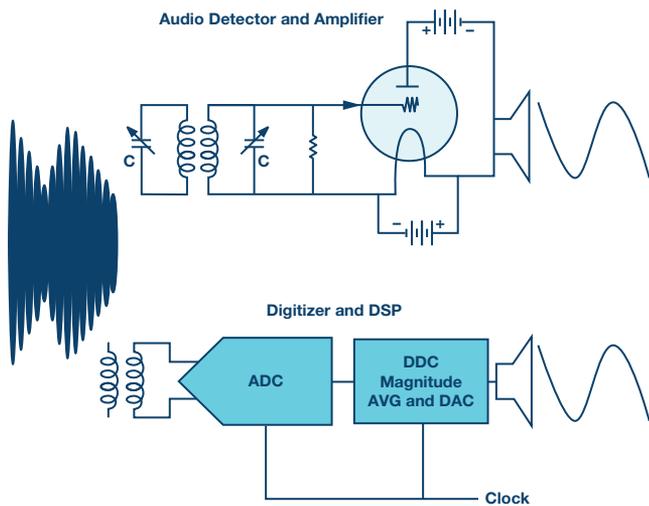
Figure 20. Detector.

While both tube and transistor versions of these radios can achieve a similar result, contemporary designs have a range of advantages. Notably, modern designs are much smaller and the power requirements are greatly reduced. While portable tube radios existed from the beginning, transistors enabled pocket size radios. Integrated circuits enabled single-chip radios for a wide range of applications from short-range wireless like the ADF7021 through high performance offered by way of AD9371. In many cases, this includes both the receiver and the transmitter.

Since monolithic radios typically utilize analog-to-digital and digital-to-analog converters, these readily facilitate complex modulation. Tube type radios historically were limited to basic modulation types such as AM and FM. When data converters are added to the radio, as is typically done on monolithic radios, new forms of modulation can be introduced by way of digital techniques including spread spectrum and OFMD, which are at the heart of most modern communications (digital TV, HD Radio, DAB, cell phones) that we rely on daily.

As radio technology continues to evolve, more advancements will come that may enable radio architectures or provide functions not currently possible. Today we have a wide selection of IF sampling superheterodyne and zero-IF architectures in highly integrated forms. Other architectures on the horizon include direct RF sampling where the signal is directly converted to digital without analog downconversion. As radio technology continues to evolve, the number of available options will grow. However, it is likely that some form of heterodyne will be with us for some time to come.

Conclusion

In the 100 years of the superheterodyne radio, little has changed in the architecture except the implementation technology. We have witnessed many changes through the years in the medium on which radios have been constructed as we've seen technology migrate from tubes to transistors to monolithic integrated circuits. These changes have enabled possibilities that were only a daydream to the early radio pioneers and on which our daily lives are so closely tied to.
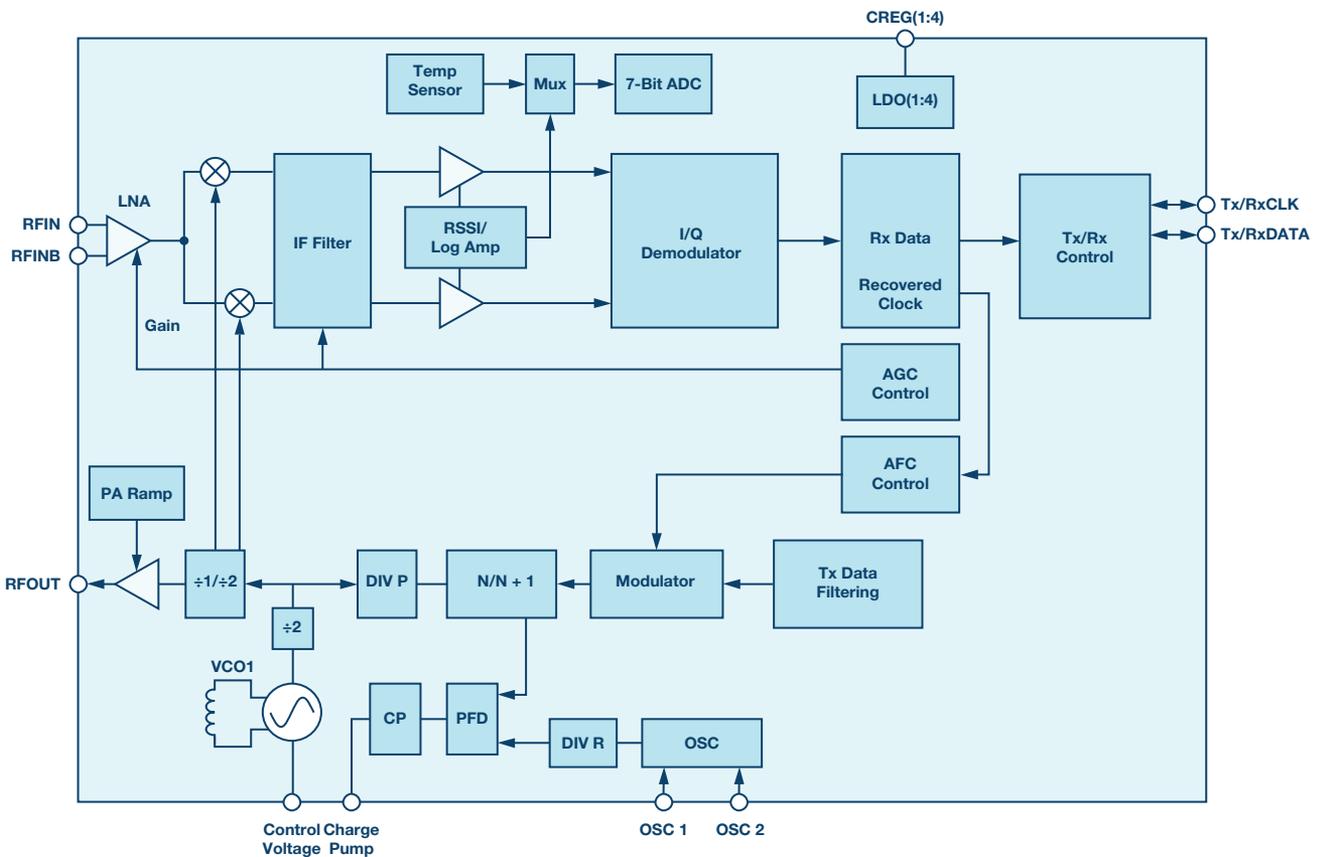
Figure 21. ADF7021 short-range wireless (simplified).
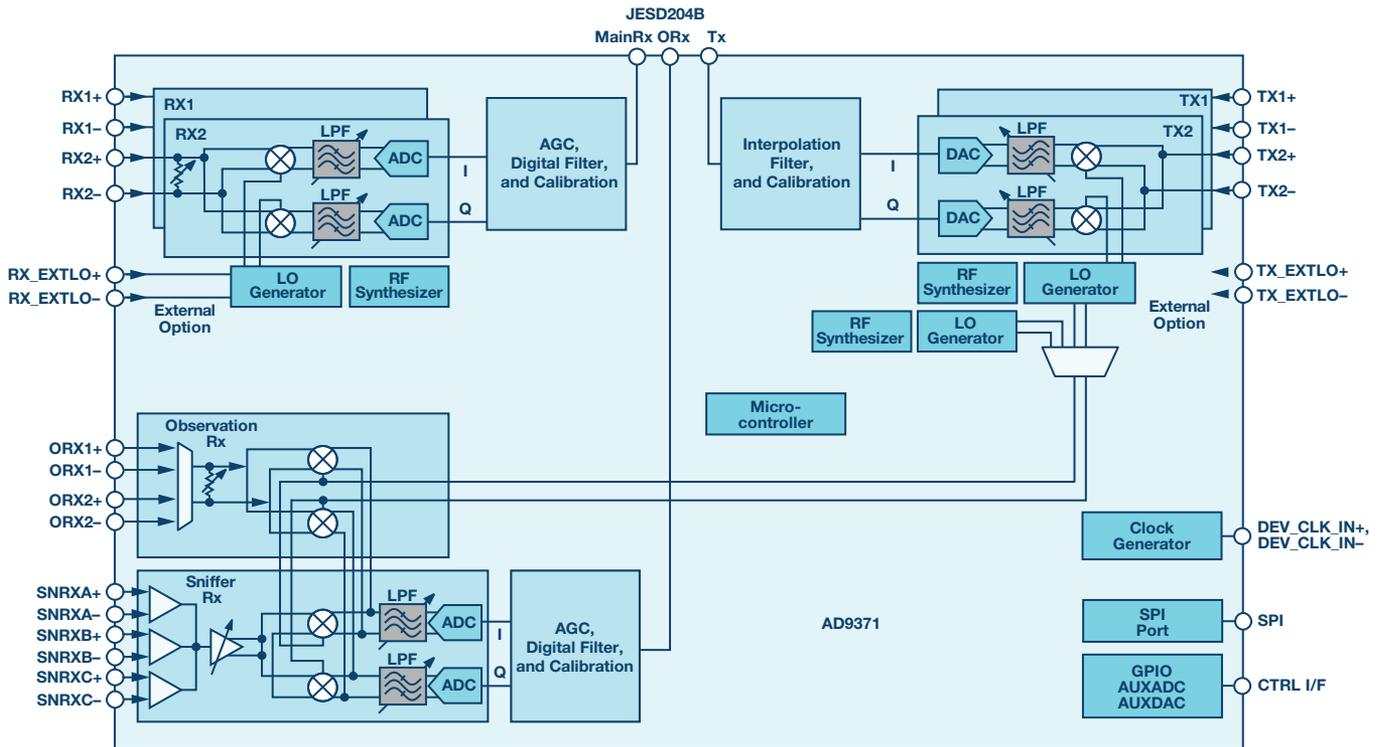
Figure 22. AD9371 ZIF transceiver.

One of the key elements that has made this possible is the detector fulfilled by the high speed ADC in today's radio technology. Improvements over the last few years in data converter and other technology has ushered in our connected world, which is changing our daily lives and the fabric of modern society. The exciting part is that this core technology is continuing to evolve, which will continue to enable new wireless solutions that may not be known today. The next 100 years carry as much potential to the next generation of wireless as Armstrong and Levy's inventions provided to the last 100 years.

References

[1] Tom Lewis. *Empire of the Air: The Men Who Made Radio*. Harper Collins, 1991.

[2] F1jmm. Recepteur tube limaille. March 1902.

[3] By Alessandro Nassiri. Detector magnetico Marconi 1902—Museo scienza e tecnologia Milano. Museo nazionale scienza e tecnologia Leonardo da, December 2012.

[4] JA.Davidson. CrystalRadio. September 2007.

[5] Fred Harris. "Exact FM Detection of Complex Time Series." Electrical and Computer Engineering Department, San Diego State University.

[6] Walter Tuttlebee. *Software-Defined Radio: Enabling Technologies*, Chapter 4: Data Conversion in Software Defined Radios. Wiley, 2002.

[7] Edward H. Armstrong. U.S. patent 1342885, "Method of Receiving High Frequency Oscillations." Application filed February 8, 1919, issued June 8, 1920

Brad Brannon

Brad Brannon [brad.brannon@analog.com] has worked at Analog Devices for 32 years following his graduation from North Carolina State University. At ADI he has held positions in design, test, applications, and in system engineering. Brad has authored a number of articles and application notes on topics that span clocking data converters, designing radios, and testing ADCs. Currently Brad is responsible for system engineering for 4G and 5G receive architectures.

Also by this Author:

Where Zero-IF Wins: 50% Smaller PCB Footprint at ⅓ the Cost

Volume 50, Number 3