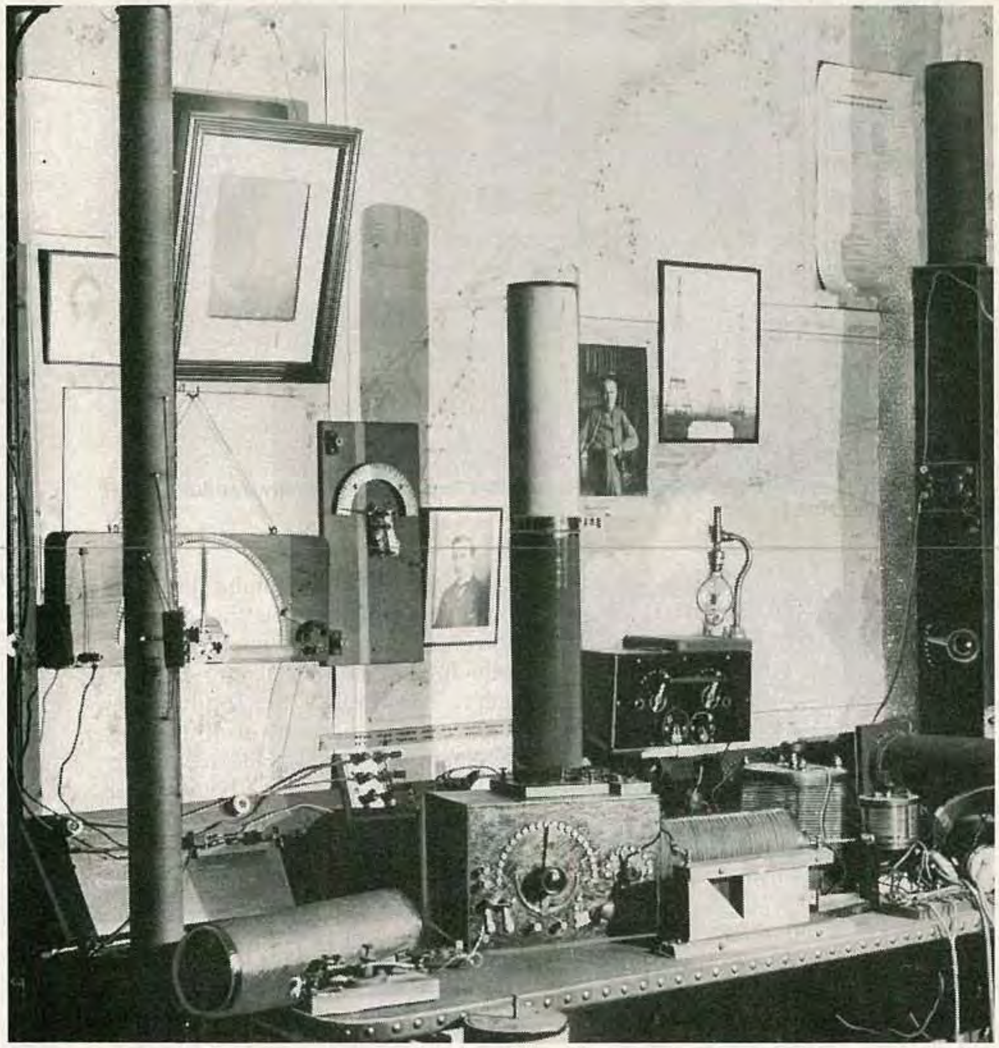
The Early Days of RADIO



MARTIN CLIFFORD

With better tubes came better radio designs.

BY THE EARLY 1920'S, RADIO DESIGN HAD progressed to the point where producing acceptable audio levels was no longer a problem and radios were being welcomed in more and more households. But with the growth of radio came the growth of radio stations. Before long there were so many radio signals, especially in urban areas, that poor selectivity no longer was merely a nuisance; it was intolerable.

At the same time, however, designers acknowledged that not everyone had the room or means to install a large, outdoor aerial system. Smaller systems, especially the indoor loops that were becoming popular in cities, could not deliver the same signal level to the receiver, so radios had to be even more sensitive to make up the difference. To solve both problems, radio designers turned to a multi-stage approach, adding sensitive and selective RF stages to their circuits. In this article we will see how their designs progressed, and eventually evolved into the circuits we use today.

Resistance-coupled RF amplifier

Early RF-stage designs were little more than barely modified audio-amplifier circuits. One early multi-stage receiver design is shown in Fig. 1. That radio used a pair of resistance-coupled RF amplifiers. Those did nothing for selectivity, but they did make receivers more sensitive. Unfor-

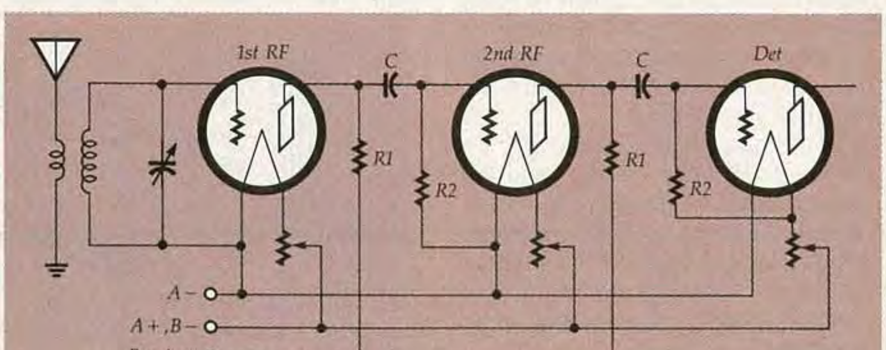


FIG. 1—RESISTANCE-COUPLED RF amplifiers made receivers more sensitive, but did nothing for selectivity.

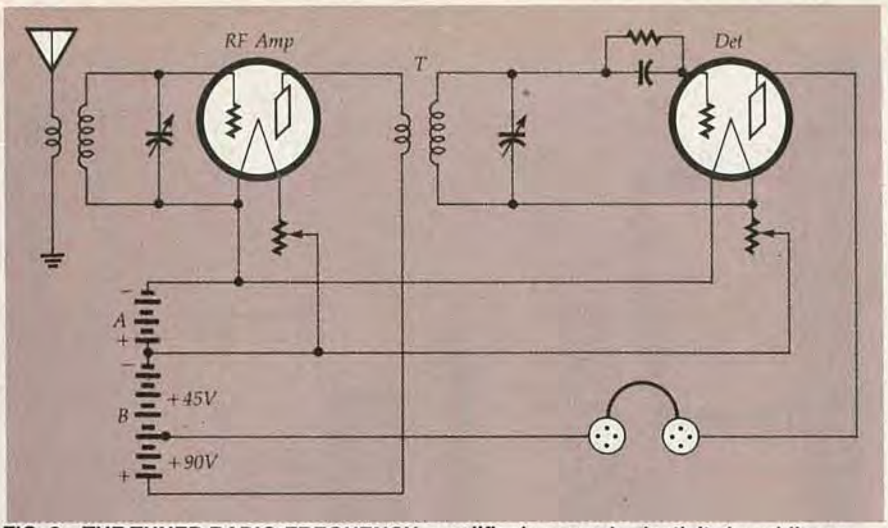


FIG. 2—THE TUNED RADIO-FREQUENCY amplifier improved selectivity by adding a tuning capacitor.

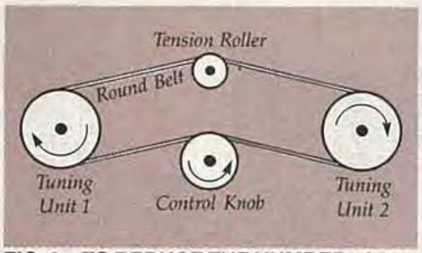


FIG. 3—TO REDUCE THE NUMBER of controls, some receivers linked the tuning capacitors via a belt drive.

tunately, the added sensitivity compounded the problem of selectivity; more signal reached the detector, but without being sorted. The resistance-coupled RF amplifier did not become popular and soon went on its way.

TRF amplifier

To improve selectivity, a second tuned circuit was added, yielding a circuit similar to the one shown in Fig. 2. The set was called a *T*uned *RF* (TRF) receiver. Selectivity was slightly improved with that design; before long, receivers with a second tuned-RF stage (three tuned circuits) began to appear. With two stages, it was possible to tune the set so that the background signals from competing stations were subdued.

Both single- and two-stage TRF receivers tuned quite broadly, and the tuning capacitor settings were not too critical. However, with the two-stage design, the user had three settings to fiddle with and every time you changed stations, all three had to be readjusted.

One-dial tuning

Even though tuning wasn't critical in the TRF receiver, tuning three capacitors every time you changed stations was a tedious affair, and one that even intimidated some prospective users. To simplify the tuning procedure, various one-dial tuning schemes were tried. One technique was the pulley and drive-belt system shown in Fig. 3. Another was a rack and pinion drive system; it was not as simple as the drive-belt method and was much more expensive to implement, so it wasn't used to any extent. The problem was solved to everyone's satisfaction by mounting the tuning capacitors on a common shaft, a tuning technique that is still in use today.

The loop antenna

As was alluded to earlier, a key to the popularity and practicality of radio was the elimination of large aerial systems. In fact, the crystal set's absolute need for such an aerial was as big a factor as any other in its eventual fall from popularity.

Owners of early multi-stage vacuum-tube receivers soon discovered that it was possible to forsake the large outdoor designs in favor of compact indoor antennas. Some experimenters used bedsprings or window screens as the antenna; others ran large loops of wire around their baseboards or in their attics. However, a more elegant answer soon became popular—the loop antenna. Those were generally positioned on top of the receiver or nearby and were of two types: the flat or pancake loop shown in Fig. 4-*a* or the box loop shown in Fig. 4-*b*-. One advantage of the loop antenna was that it helped minimize competing signals. The loop was part of a tuned circuit (see Fig. 5) that could be tuned for resonance at a particular frequency, thereby improving selectivity. Further, the loop could be physically rotated or positioned to minimize interference. As shown in Fig. 5, a "C" battery often was used for bias, but grid leaks and grid capacitors were also used for that purpose.

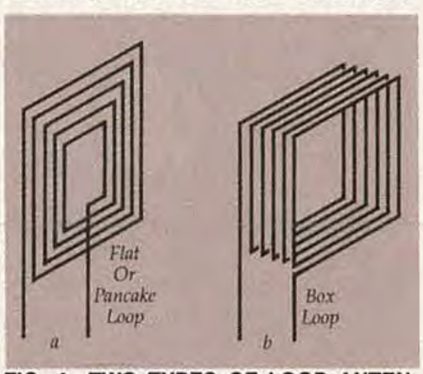


FIG. 4—TWO TYPES OF LOOP ANTENNAS. A flat or pancake type is shown in *a*, a box type is shown in *b*.

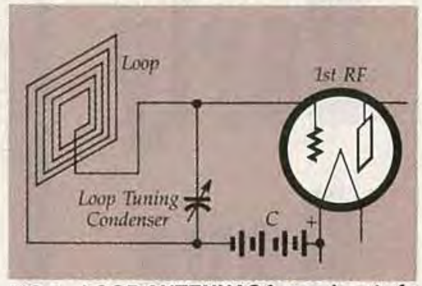


FIG 5—LOOP ANTENNAS formed part of a tuned circuit and could be tuned for a specific frequency, thereby increasing selectivity.

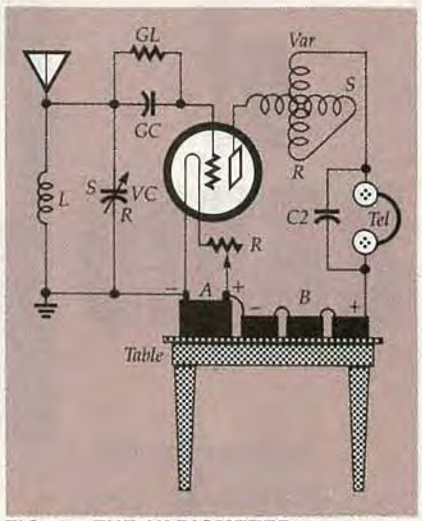


FIG. 6—THE VARIOMETER was a carryover from crystal sets.

Variometer tuning

While the appearance of vacuum-tube triodes spelled the end of crystal-set popularity, that end did not come overnight.

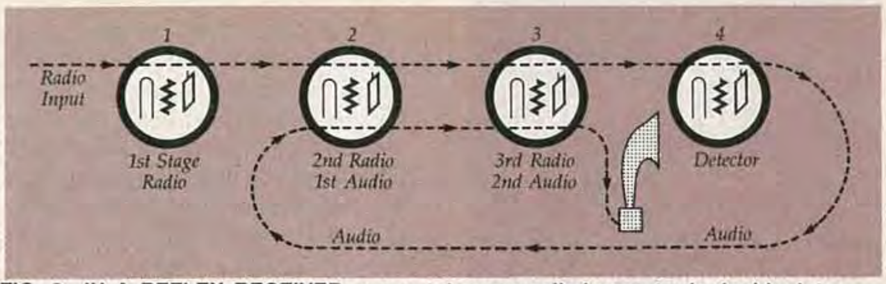


FIG. 8—IN A REFLEX RECEIVER, some tubes are called upon to do double duty.

Both types of sets co-existed for many years. And many sets incorporated circuits and features from both types of receivers. For instance, some sets combined a crystal detector with a vacuum tube audio amplifier.

One crystal set component that was commonly found in tube sets was the *variometer*, which consisted of a continuously variable inductance formed by connecting two coils in series and mounting them so that one rotated within the other. The overall inductance could be varied over a wide range, with a very low minimum value. As shown in Fig. 6, the variometer was inserted in the plate circuit of the tube. Both the variometer and the variable capacitor in the grid circuit were used to tune a signal. In time, with the introduction of receivers with two and three tuned-RF stages, the variometer joined the growing heap of discarded electronics components.

The neutrodyne

It seems that every advance in receiver circuitry was accompanied by some unexpected problems; the triode

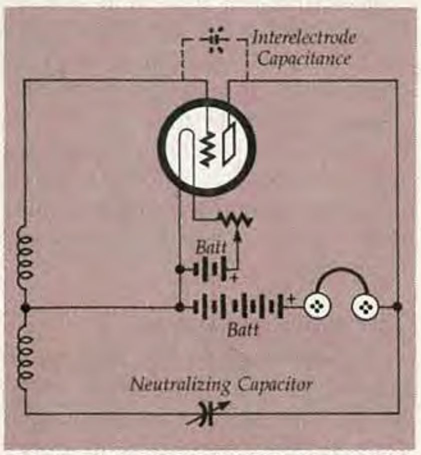


FIG. 7—A NEUTRALIZING CAPACITOR was needed to compensate for the triode's interelectrode capacitance.

and the TRF were no exception. Since the elements of the triode were made of metal, they formed a capacitor. And like all other capacitors, the *interelectode capacitance* had a reactance that varied inversely with frequency. When used as an audio amplifier, the interelectrode capacitance was not a great concern. At radio frequencies, however, it was quite definitely a concern. At those frequencies, the interelectode capacitance provided a feedback path between the receiver's input and output circuits. Since the feedback was in phase with the input, the result was oscillation, and a howling in the headphones or the speaker.

Fortunately, the problem was solved by Professor Hazeltine of the Stevens Institute in Hoboken, NJ when he developed what he called the neutrodyne principle. The capacitor shown in dashed lines in Fig. 7 indicates the feedback path provided by the tube's interelectrode capacitance. Hazeltine's technique was to use an adjustable capacitor to intentionally feed back part of the signal, but out of phase with the input. The out-of-phase feedback signal was equal to or greater than the in-phase feedback, and the two canceled out.

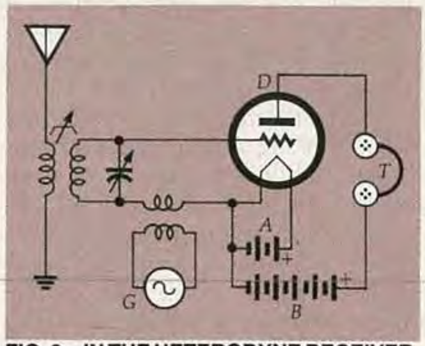


FIG. 9—IN THE HETERODYNE RECEIVER, a locally generated signal is mixed with the received signal.

The introduction of the multigrid tube relegated the neutrodyne to receiver history. However, neutralization is still used today in transmitters, including those with multi-grid tubes.

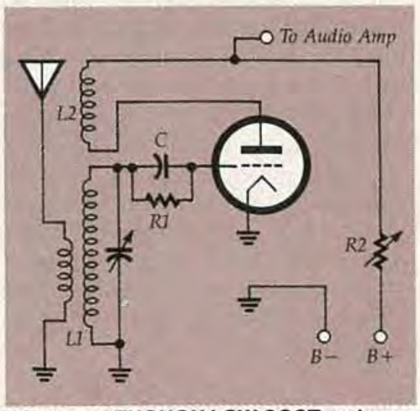


FIG. 10—ALTHOUGH LOW COST and easy to build, the regenerative receiver suffered from many drawbacks.

Enter the tetrode

The need for receiver tube neutralization was eliminated by the introduction of the tetrode. That four-electrode tube was invented by A.W. Hull. His name for the tube was the *Pliodynatron*, which never did grip the public's imagination. But, although the name was soon forgotten, the tube remained.

In the tetrode, another grid (the "screen grid") was positioned between control grid and plate. That reduced interelectrode capacitance by putting the tube's capacitances between grid and plate in series. Instead of being from grid to plate it was from grid to screen grid and from screen grid to plate. Some early attempts

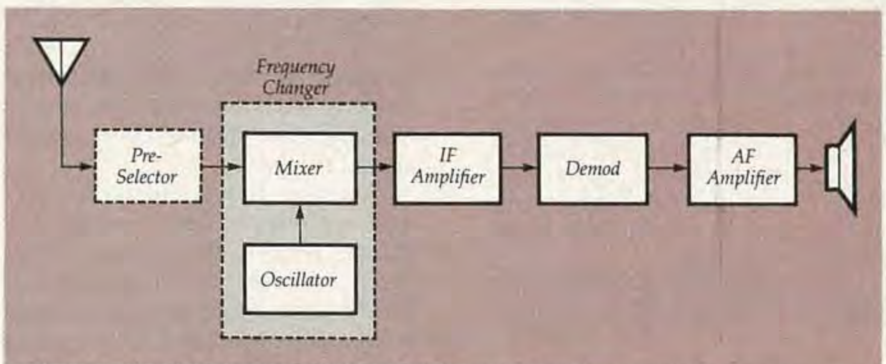


FIG. 11—BLOCK DIAGRAM of a super heterodyne receiver.

connected the screen grid to ground which of course made it an effective shield. But it was discovered that if a positive potential was put on that new element, and was held constant, feedback would be practically zero. A capacitor was connected between screen grid and ground effectively putting the grid at signal ground.

The addition of the screen grid tube also had a second, pleasant effect. It increased the amplification factor to as much as 400, as compared to only about 10 for a triode.

During the mid 1920's, volume was still one of the most important criteria in the public's evaluation of a receiver's quality. One approach to increasing the volume without increasing the component count, and hence the cost, was to use a *reflex* circuit. There were many variations of that idea, and one popular one is shown in Fig. 8. In that circuit, the incoming signal was brought into an RF amplifier, proceeded through two more RF stages, and was then fed into a detector. From the detector the signal was brought back into the second RF amplifier, which was then also used as the first audio amplifier. The signal was then injected into the third RF amplifier, which now also functioned as the second audio stage, and whose output was delivered to a speaker. Thus, two of the four tubes did double duty, functioning as both RF and AF amplifiers.

The heterodyne receiver

Spark transmitters generated large numbers of damped waves having all kinds of harmonics. Those waves heterodyned or beat together in the receiver's detector, and therefore could be referred to as self-modulating. That property was what made it possible to hear dits and dashes using simple crystal receivers.

However, those same receivers were not suitable for receiving *C*ontinuous *W*ave (CW) signals. A CW signal is simply an RF wave that is broken up into long and short pieces using a telegraph key and is not self modulating. That problem was overcome by the invention of the heterodyne receiver by Reginald A. Fessenden in 1905. The hetrodyne circuit was the basis on which Major Edwin H. Armstrong devised the superheterodyne circuit in 1918.

As shown in Fig. 9, the principle behind the heterodyne receiver is quite simple. In it, a local signal is generated (G) and injected into the detector. There, it beats or heterodynes with the received signal, producing a number of sum and difference signals. One of those, a difference signal in the audio range, is heard in the headphones.

The regenerative receiver

The triode vacuum tube was full of surprises. It was initially regarded solely as an audio amplifier. Subsequently, it was learned that the same tube could be used, simultaneously, as a detector. And then it was discovered that the triode also could simultaneously function as an oscillator as well. A circuit that used one tube for all three functions was the regenerative receiver shown in Fig. 10.

In that circuit, some of the amplified signal output is fed back to the input in-phase with the incoming signal. The grid leak, R1, its shunting capacitor C1, the tuned circuit, and the filament/cathode form a diode detector, with R1 functioning as the diode's load. The amplified signal is inductively coupled from L2 into L1, a process that is allowed to continue by adjusting of R2 until the circuit is at the point of oscillation; that's the point where its sensitivity (gain) is greatest. Since the tube can be made to oscillate, that self-oscillation produces a frequency that beats with the incoming signal, resulting in an audio tone that can be heard in headphones or speaker.

The superheterodyne

A block diagram of a basic superheterodyne receiver is shown in Fig. 11. The preselector consists of one or more tuned RF stages, but in less expensive receivers it is often omitted. As in the heterodyne receiver, the incoming signal is fed to a mixer, where it beats with a signal supplied by a local oscillator. However, unlike the heterodyne receiver, the local oscillator is designed so that the difference frequency is on the order of several hundred (or more) kHz. Known as the *I*ntermediate *F*requency (IF), the difference signal is then amplified, demodulated to recover the audio signal, and then the audio is amplified again. The great advantage of the superheterodyne design was, and still is, the large amount of selectivity that can be supplied by the IF stages, which are fixed-tuned to a selected frequency. **R-E**