Amateur Radio Equipment Development: A Historical Perspective

By Joel R. Hallas, W1ZR

The year 2014 marks the 100th year of the American Radio Relay League (ARRL, the national association for Amateur Radio). While there was Amateur Radio before there was an ARRL, it is clear that most of the advancement of radio science occurred during the period. This 100 year period saw radio move from a barely functional, unreliable set of disparate communications operations to a largely integrated system comprising major elements of the lives of everyone.

Radio amateurs have not only benefited from all the technological advances over the period, but also many contributed directly to the development and success of important advances in technology. This section highlights some of the major developmental steps of Amateur Radio that led us to the technology that we enjoy today.

First There Was Spark

Transmitters have what would appear to be a rather straightforward job — generate an ac signal and modify it in some way to carry information. In the early days, it was found that an electrical spark created in a circuit that was connected to an antenna could be received some distance away by various devices connected to another antenna. It didn't take folks very long to figure out that if the sparks were sent in patterns corresponding to Morse or other telegraphy codes, information could be sent over long distances without connecting wires. Wireless communication was born!

Figure 1 shows an advanced amateur spark station from 1907 — functional and on display at the ARRL Headquarters Museum of Amateur Radio (but not hooked to an antenna). On the lower left is the receiving side, a crystal detector with "cat's whisker," variable tuning coils and a tubular variable tuning capacitor. The top deck holds the transmitter, starting with the automotive type spark coil on the right. In the middle is multilayer flat mica capacitor with taps at various points, along with the tuning inductor. Together they defined the operating frequency. On the left of the top shelf is the actual spark gap. The lower right contains a land line sounder and switches to allow traffic to be relayed to offices connected locally.

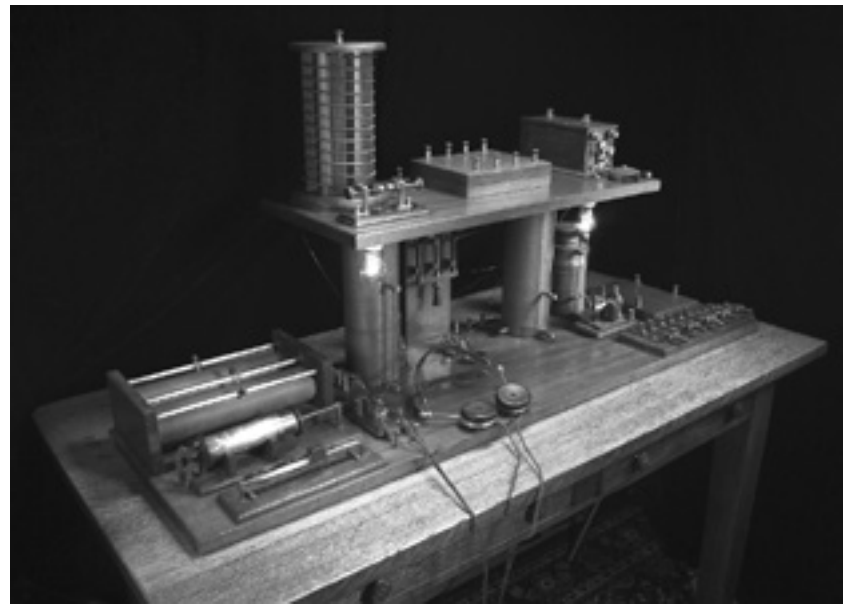

Figure 1 — An advanced amateur spark station from 1907 — functional and on display at the ARRL Headquarters Museum of Amateur Radio. [Joel Hallas, W1ZR, photo]

Technology quickly advanced through a number of variations of spark and arc transmitters and a wide variety of ingenious receiving devices. Eventually wireless telegraphy became a serious business involved with monitoring the safety of ships at sea in ways that hadn't been possible before. See **Figure 2** for a view of an early shipboard wireless station. The commercial systems of the day were more electromechanical than what we would consider electronic today.

The Receiving Detector

The primary function of a receiver takes place in a *detector*, the circuit that extracts the information from the modulated RF signal that arrives from the antenna. As shown in **Figure 3**, the simplest receiver consists of just a detector. The earliest receivers were constructed in just that way, starting with some unusual electromechanical marvels from Marconi's days. The crystal detector, in the form of the crystal set, was the first radio experienced by youngsters from the 1920s to the 1980s and was the mainstay of many commercial users until vacuum tubes became available. While even simpler configurations are possible, the circuit of **Figure 4** is typical. In an early set, a galena crystal and a cat's whisker would be employed instead of the 1N34 germanium diode.

The key performance parameters of this receiver are easy to describe. Its *sensitivity* is the signal level required at the antenna input to create an audible signal as a consequence of the diode acting as a square-law detector. The *selectivity* is determined by the loaded Q of the single-tuned resonant circuit, while the *dynamic range* extends from the sensitivity level to the point that the headphone diaphragms hit their limits or the diode opens due to high current.

While this receiver looks like it would be rather limited in performance (and it is), it does work and was even the basis for most microwave radar receivers from WWII and for some years thereafter. That was the origin of the venerable 1N34 diode. The *diode detector* is still used for full carrier amplitude modulated signals in more complex modern receivers.

Early Amateur Transmitters

Amateur transmitters evolved from those based on automotive type spark coils with

Figure 2 — Marconi shipboard radiotelegraph station from 1900. (Photo courtesy of Marconi Company, PLC)

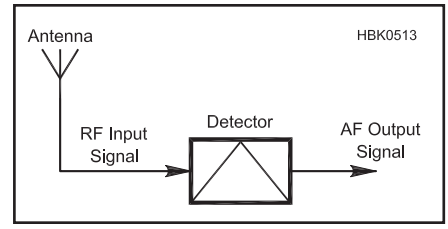

Figure 3 — The minimal receiver is just a detector.

electromagnetic interrupters (think buzzer) to high voltage transformers powered by ac mains and using motor driven *rotary spark gaps*. The rotary gaps produced sparks interrupted at a multiple of the motor speed, resulting in a form of modulation that could be heard using the popular crystal detectors of the day. The key was typically in the ac mains circuit, not a system that would pass a safety inspection today.

A high power amateur station had a transformer that could deliver 1 kW, although it's not clear how much of that power was radiated as light and noise by the gap. While the commercial — especially maritime — radio services moved to more sophisticated and complex arc transmitters, along with a number of different detector technologies, the amateur operators tended to stay with spark transmitting and crystal receiving technology until vacuum tubes became readily available.

Some commercial fixed shore stations used electromechanical *alternators* as the generators of RF for transmission. A very basic alternator might consist of a coil with a core of soft iron (pure molecular iron that will magnetize and demagnetize easily) rotating past two poles of a magnet inside a round iron mounting as in **Figure 5**. As one end of the rotating iron core coil passes the north magnetic pole, a magnetic field is built up in it. As a result of the magnetic field increasing then decreasing in the rotor coil's core, one half of an ac cycle is induced in its coil. As it continues to rotate and passes the south pole, an opposite half cycle of ac is induced in it. To produce 60 Hz ac, the coil would have to rotate 60 times per second, or 3600 revolutions per minute — a very fast rotation. The two ends of the rotating coil are connected to two slip rings on the shaft that rotates the coil. There are two brushes of fixed carbon or other material that make a constant contact with the rotating slip rings. Any ac voltage that is generated is available at these two brushes. By using multiple rings and brushes, multiple cycles of ac will be produced with each rotation.

Better systems were used to generate ac at frequencies in the lower RF range. Ernst Alexanderson, a prolific Swedish/American electrical engineer and inventor, and Canadian inventor Reginald Fessenden first produced a lower power 60 kHz RF alternator in 1906, with which they could transmit telegraphy as well as voice and music. Later the higher powered Alexanderson alternators used frequencies closer to 20 kHz. Other popular alternators of the period were the German Joly-Arco and the Goldschmidt, which was somewhat similar to the Alexanderson.[1]

The final form of the Alexanderson alternator uses a rotary toothed disc, called an *inductor* (not to be confused with a regular electronic coil). This type of inductor is a large round, flat, soft-iron disk with teeth or other regularly spaced features along the edge as shown in **Figure 6**. The many teeth of the inductor are rotated rapidly by a constant speed ac motor between the north (N) and south (S) poles of double-wound field pole electromagnets mounted one tooth-width apart all around the inside of the machine. As an iron tooth passes between the N and S ends of the dc-excited electromagnet field poles, it provides a better path for the N to S magnetic lines of force. As a result the magnetism in the field pole builds up, then collapses as the tooth moves on.

As the next inductor tooth passes the field pole, it develops another magnetic field build up and collapse. These varying magnetic fields induce ac cycles in all of the secondary RF output field-pole coils. Since there are as many field poles around the inside of the machine as teeth on the inductor, ac voltages are developed in the field pole ac pick-up coils all at the same time and are all inductively coupled to a coil/antenna/ground circuit that radiates the RF waves. The speed of the motor rotating the inductor and the number of teeth determines the output frequency.[2,3]

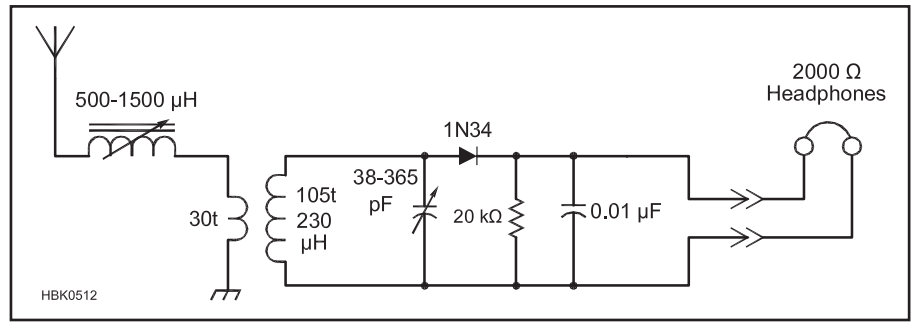

Figure 4 — Circuit of a typical "modern" crystal set with a semiconductor diode as the crystal detector.

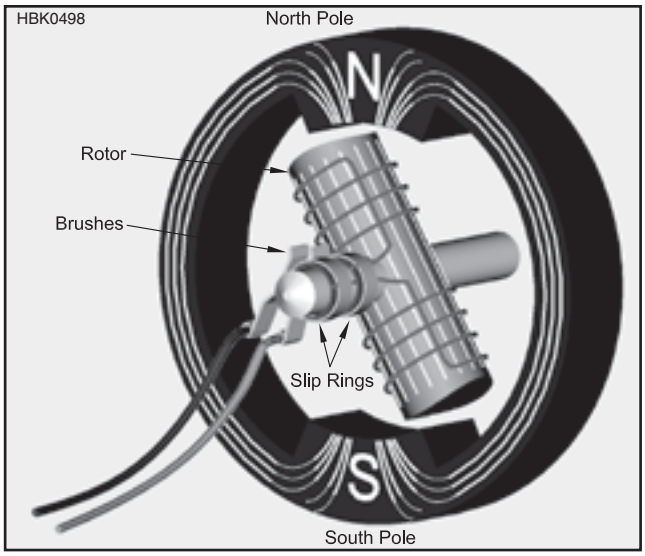

Figure 5 — Basic alternator consisting of a soft-iron cored coil rotating past two poles of a magnet.

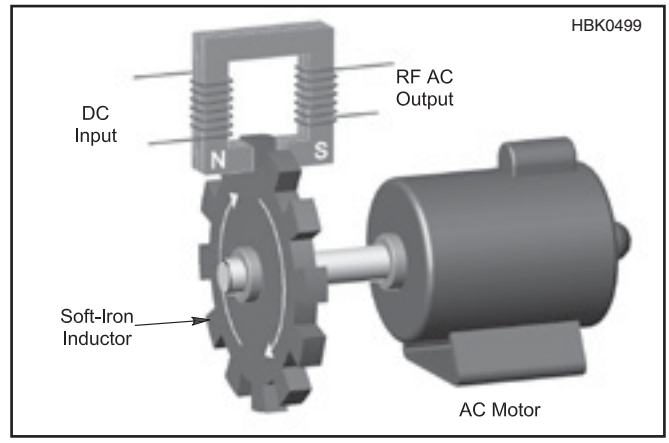

Figure 6 — High frequency alternator using an inductor disc.

Then Came Vacuum Tubes

The Diode Detector

The vacuum tube was a development based on Edison's light bulb and invented in 1904 by John Fleming, a professor of physics at the University of London. Fleming was a consultant to the Marconi Company who had worked on the development of the transmitter for their transatlantic trials. The *Fleming valve* was a diode and, as such, it could work as a receiving detector in place of the mechanically difficult galena crystal and cat's whisker arrangement. This was a great improvement, especially for shipboard stations, since vibration — not to mention recoil from cannon — would result in movement of a cat's whisker and loss of reception.

Early vacuum tubes were quite expensive and out of the reach of most amateurs. With most amateurs being able to avoid the heavy vibration and artillery issues of shipboard operation, the inexpensive but finicky crystal detector continued to serve until tubes became obtainable more reasonably.

The Triode Vacuum Tube Changes Everything

The development of the Audion by Lee DeForest in 1906 moved electronics and radio a large step forward. The Audion was the first three element (*triode*) vacuum tube. It added a grid located between the filament or cathode and the anode or plate of the earlier diode. A small change in grid voltage resulted in a larger change in plate current. This allowed the tube to amplify rather than just rectify. Weak signals applied to the grid would become stronger signals in the plate circuit.

Amplification could do a number of things that hadn't been possible before. Weak signals from a diode detector could be made strong enough to fill a room using loudspeakers. Perhaps more importantly, an amplifier with positive feedback could be made to oscillate and thus generate a sinusoidal signal. If this signal were generated at radio frequencies, it could be coupled to an antenna and keyed in Morse code, making for a very efficient transmitter.

The signal transmitted by a vacuum tube oscillator was a sustained and nearly pure sinusoidal signal referred to as continuous wave or CW. This was quite unlike the wideband chopped and damped waves of the spark transmitter. This was a mixed blessing. On the one hand, the CW signal occupied a much narrower band of frequencies, allowing other spectrum users to operate on frequencies much closer together without interference. On the other hand, the buzzing or chopped sound of the spark signal in a crystal set was replaced by a series of thumps on key-up and key-down when the set received a CW signal.

The Regenerative Receiver

Fortunately, the solution to this problem was provided by the same technology in the form of the oscillating, or *regenerative*, detector. Invented by Edwin Armstrong in 1913, the regenerative, or *autodyne*, detector became the mainstay receiver technology of amateur and commercial operators for many years.[4] The regenerative detector used an amplifying stage that was made to oscillate at a frequency slightly offset from the received CW signal. The mixing of the two signals in the tube resulted in a beat note that was clearly audible in the headphones in the plate circuit. In addition, the oscillation magnified the sharpness of the input tuned circuit, resulting in high gain and the ability to separate signals. This approach is in many ways similar to what is called a *direct-conversion* receiver by modern hams.

The regenerative receiver had a number of advantages, including simplicity and low cost. The early units used a single Audion and provided significant performance with few parts. There were some problems inherent to the design, however. A major one was that the oscillator was coupled to the antenna, making it serve as a low power transmitter while it was receiving. During World War I, allied transport ships suffered at the hands of U-boats that were able to track them by listening for their oscillating detectors. Working with limited Depression-era resources after the war, clever amateurs would actually use this "problem" as a benefit by keying the receiver to make it serve as a short range transmitter.

Amplifiers

The next element added to the detector was the *amplifier*. The first amplifiers provided bandwidth capable of amplifying only audio frequency signals and thus were inserted following the detector. In that position, they offered no improvement in sensitivity, which was still limited by the diode; however, they appeared to improve sensitivity because the amplified audio output of the receiver was now louder. It was now possible to have more than one person listen to a received signal

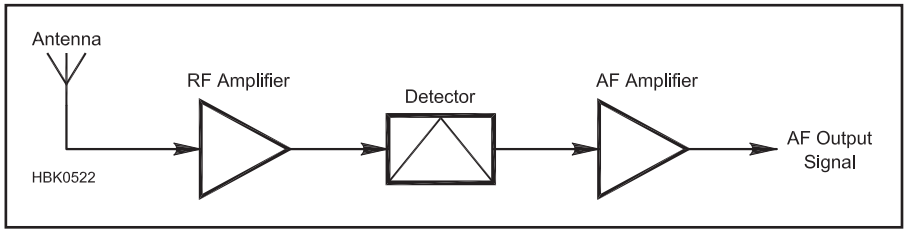

Figure 7 — Detector with RF and AF amplifiers to improve sensitivity, selectivity and power output.

through the use of a loudspeaker.

As shown in **Figure 7**, it wasn't long before vacuum tube and circuit design technology improved to the point that amplifiers could be used at RF as well as AF. The addition of RF amplifiers provided two performance improvements. First, the level of signals into the detector could be increased, providing an improvement in sensitivity. Second, if each amplifier were coupled using tuned resonant circuits, as was the usual practice, the selectivity could be improved significantly.

A receiver with one or more tuned RF amplifier stages was called a *tuned radio frequency*, or *TRF*, receiver. Some had as many as three or four RF stages, initially with separate controls, and later with up to a five-gang variable capacitor for station selection. Needless to say, a certain amount of skill was needed to adjust each stage so it would properly *track* as it was tuned over the frequency range.

There were other challenges with the TRF. With the high gain of the multiple stages all tuned to the same frequency, it was not trivial to avoid oscillation. Another concern was the selectivity provided over the tuning range. If the selectivity of the five tuned circuits could provide a bandwidth of 2% of the tuned frequency, that would result in 30 kHz at the top of the broadcast band (around 1500 kHz) but only 10 kHz at the bottom (500 kHz). Top-end receivers of the day used various methods to maintain similar bandwidth across the range automatically.

Even though the *superheterodyne* receiver became available by the 1930s, the regenerative receiver remained popular with hams through the Great Depression until World War II. The solution to the problem of transmitted signals from the receiver oscillator, described previously, was to insert an RF amplifier stage between the antenna circuit and the detector. This also improved the sensitivity somewhat, and it made the tuning more stable if the antenna moved in the wind. An audio amplifier stage was often used following the detector to permit operation with a loudspeaker. This three tube configuration was manufactured by the National Radio Company from 1931 to 1939 as the SW-3, and was one of the most popular medium and shortwave communications receivers of the era.

Early Vacuum Tube Transmitters

Vacuum tube transmitters developed in parallel with the receiving sets. Initially limited by the cost and availability of tubes that were allocated to the military, receiving type tubes began to become available after World War I. Enterprising amateurs "borrowed" tubes from the family radio and used them as single tube oscillating transmitters at first. Some receiving type audio power amplifier tubes were pressed into service as RF power amplifiers.

While transmitters are composed of many of the same named blocks as those used in receivers, it's important to keep in mind that they may not be the same size. An RF amplifier in a receiver may deal with amplifying picowatts while one in a transmitter may output up to megawatts. While the circuits may even look similar, the size of the components, especially cooling systems and power supplies, may differ significantly in scale. Still, many of the same principles apply.

Continuous Wave Telegraphy

By the early 1920s, large vacuum tubes were developed that could be used in high powered transmitters. Most of the big alternators were replaced by less massive high power tube transmitters. Amateur transmitters also quickly moved from spark to vacuum tube sets. The vacuum tube oscillators generated a single-frequency RF signal directly and continuously, creating the *continuous wave* or CW transmitter.

Although CW transmitters generally operated with less nominal power output than the spark transmitters of the day, they actually delivered more power to the antenna. Signals were of a much narrower bandwidth, allowing more usable channels once receiver capabilities caught up. The early tube transmitters were generally single-stage self-excited oscillators coupled directly to an antenna as shown in **Figures 8** and **9**.

While they were a vast improvement over early technology, such transmitters had their limitations. Note that the antenna was connected directly to the tuned circuit that controlled the frequency. This meant that every time the antenna moved in the wind, the resultant change in loading shifted the transmit frequency. If the cat walked by and was lucky enough not to be electrocuted by the exposed high voltage wiring, the frequency moved even on a calm day. The frequency would often shift with each key closure as the power supply voltage sagged with the additional current drain, resulting in signals with characteristic chirps and whoops.

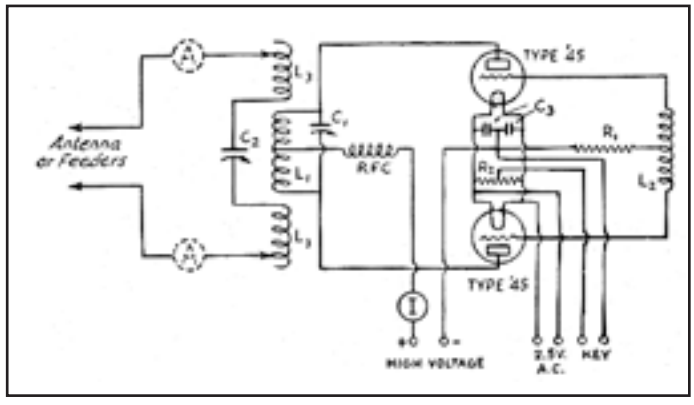

Figure 8 — Transmitter schematic as shown in an early *QST* magazine.

Figure 9 — Photo of a transmitter built from the schematic in Figure 8.

Master Oscillator-Power Amplifier Transmitters

All of these issues were at least partly addressed by adding a power amplifier between the oscillator and the antenna. Morse code transmission was accomplished by using the key to turn the power amplifier stage on and off. If properly designed and adjusted, and powered by a solid power supply, the oscillator could be left on between the transmitted dots and dashes. This stabilized the transmitted signal's frequency and the antenna was well-isolated from the frequency determining components.

This was known as the *master oscillator-power amplifier* (MOPA) configuration and remained popular in WWII military radios and amateur equipment into the 1950s. The popular (for amateurs as war surplus) ARC-5 series military aircraft single band HF transmitters were of MOPA design using a single-triode variable-tuned oscillator and a pair of tetrode amplifier tubes in parallel. While the military ran these at more conservative ratings, amateurs used them to provide up to 100 W output, all from a box the size of a small loaf of bread. Two examples are shown in **Figure 10**.

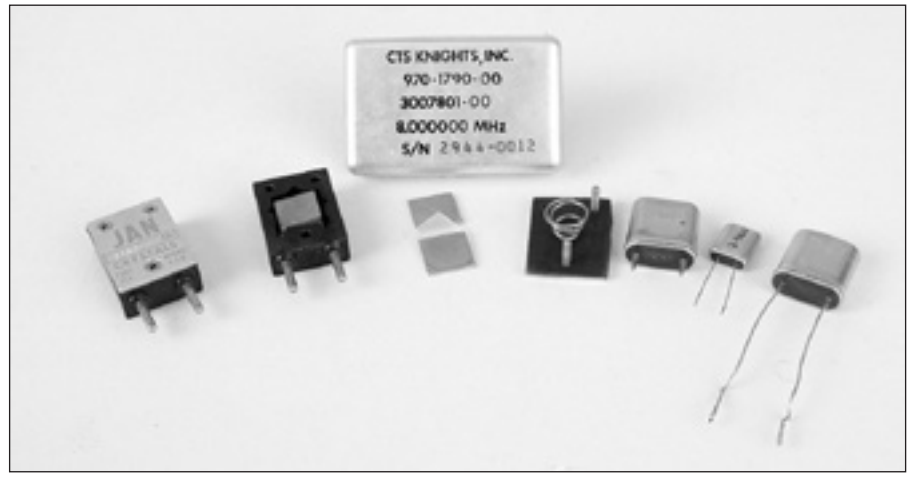

Figure 11 — Sample of typical frequency control crystals for amateur use. The units on the right are modern sealed types. On the far left is a WWII era crystal with a disassembled type FT-243 holder in the center.

Crystal-Controlled Transmitters

The transmitters discussed so far operated on frequencies under the control of variable-tuned LC circuits. This provided a certain amount of flexibility, but had the drawback of making the operating frequency uncertain. Stability over time, and with variations in supply voltage and temperature, was not as good as might be desired.

Properties of crystal structures were studied at least as far back as the 18th century, but the first major application of the piezoelectric effect that converts crystal motion to voltage and back occurred during WWI. The French developed a piezoelectric ultrasonic transducer that was used as the transmitter in an acoustical submarine detection system. Between WWI and WWII, the use of resonant wafers of quartz crystal as frequency determining elements in oscillators became feasible and packaged crystals became readily available.

Crystal frequency control was particularly suitable for radios that operated on assigned fixed channels, common to most services other than Amateur Radio. Still, crystal-controlled transmitters were popular in amateur service, with most stations having a large collection of crystals that could be selected to change frequency. For many years, the old entry-level Novice class amateur required crystal control of the transmitter. **Figure 11** shows a selection of popular crystal types.

Figure 10 — A pair of 7 to 9.1 MHz ARC-5 transmitters from WWII. On the left is the US Army Air Forces version (BC-459), and a Navy equivalent (T-22, ARC-5) is shown on the right. These were available for about $5 in "new in box" condition after WWII.

Multiband Transmitters

The transmitters discussed so far generally operated over a single band or frequency range. In order to operate on multiple bands, a number of approaches were used. The early amateur bands were selected to be harmonically related in order to avoid interference to other services from spurious harmonic signals. The amateur bands of the early years were 160, 80, 40, 20, 10, 5 and 2½ meters.

A popular approach to early vacuum tube transmitters designed for multiple bands was to use frequency multiplier stages to raise the frequency to a desired harmonic. A frequency multiplier was just an amplifier operated in class C with the output circuit tuned to a multiple of the input frequency. The plate current pulses would cause the output tank circuit to ring at the selected harmonic, filling in the cycles between the pulses. If the Q of the tank circuit were high enough, harmonics through the third or fourth could be generated with acceptable distortion. Additional tuned stages would reinforce the desired harmonic and reject the other signals.

A block diagram of a typical period transmitter for 80, 40 and 20 meters is shown in **Figure 12**. Note that a particular advantage of this configuration is that a single crystal at say 3.505 MHz can also be used at 7.010 and 14.020 MHz. This arrangement was quite common in transmitters designed for frequency modulation in which not only the

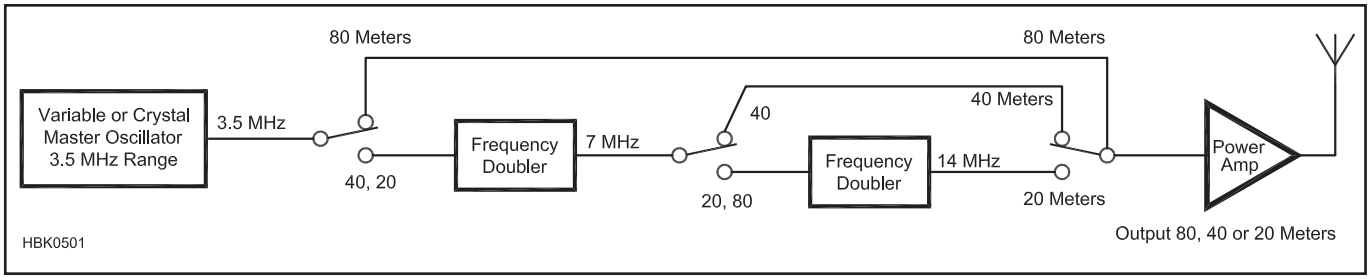

Figure 12 — Block diagram of typical frequency multiplier transmitter.

frequency, but also the deviation, is multiplied at each step.

While the figure shows cascaded doublers, some used circuits switchable between doublers, triplers and quadruplers to avoid the need for additional stages. Note that a limitation of this arrangement is that any drift or instability is also multiplied, making it necessary that particular attention is paid to oscillator stability. The tuning rate of variable oscillators, if used, is also different on each band.

Heterodyne Receivers Solve Many Problems

A receiver design that avoids many of the issues inherent in a TRF receiver is called a *heterodyne*, or often *superheterodyne* (*superhet* for short), receiver. This design was proposed by then Major Edwin Armstrong near the close of World War I for use in direction finding equipment. The superhet combines the input signal with a locally generated signal in a nonlinear device called a *mixer* to result in the sum and difference frequencies as shown in **Figure 13**. The combination of the mixer and oscillator is often called a *frequency converter*. The receiver may be designed so the output signal is anything from dc (a so-called *direct-conversion* receiver) to any frequency above or below either of the two frequencies. The major benefit is that most of the gain, bandwidth setting and processing can then be performed at a single frequency.

Changing the frequency of the local oscillator (ie, turning the VFO control in a modern radio) makes a corresponding change in the input frequency that is translated to the output, along with all its modulated information. In most receivers the mixer output frequency is designed to be an RF signal, either the sum or difference — the other being filtered out at this point. This output frequency is called an *intermediate frequency* or *IF*. The IF amplifier system can be designed to provide the selectivity and other desired characteristics centered at a single fixed frequency — much easier to process

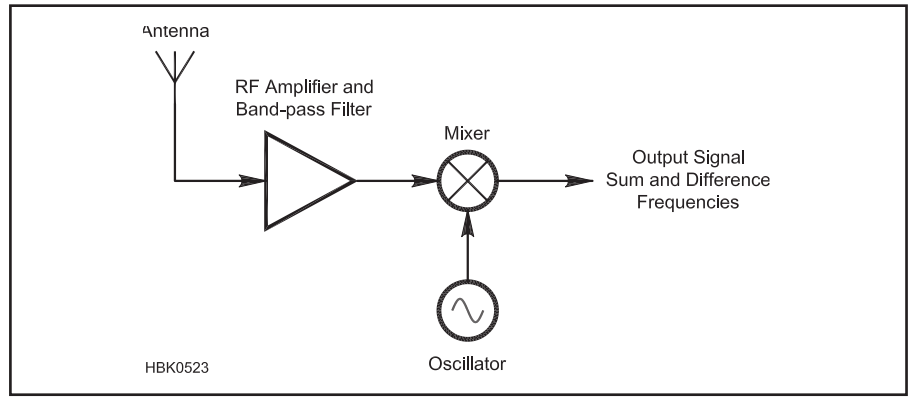

Figure 13 — Basic architecture of a heterodyne conversion stage. The output signal is at a fixed intermediate frequency.

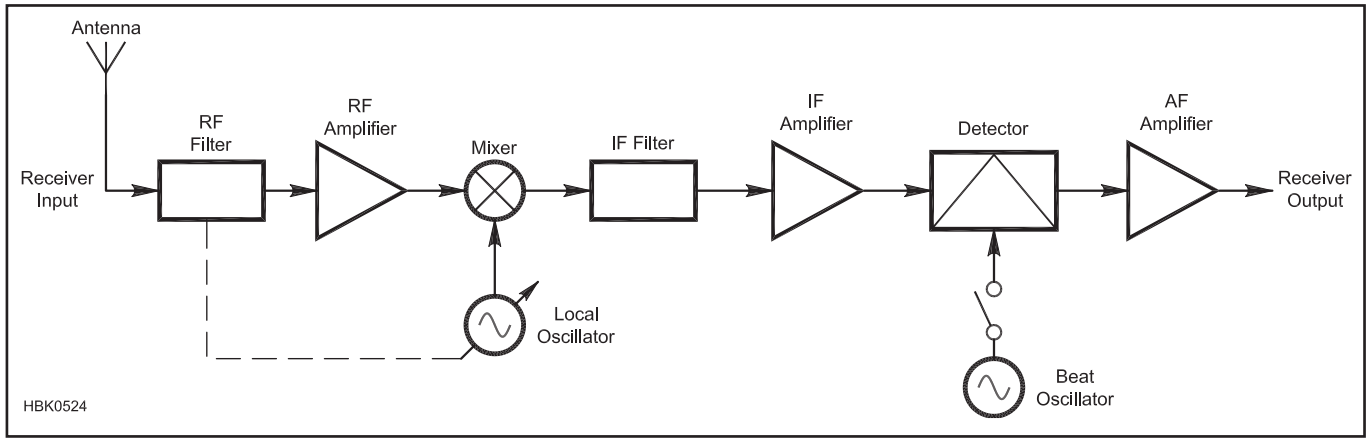

Figure 14 — Elements of a traditional superheterodyne radio receiver.

than the variable arrangement of a TRF set.

A block diagram of a typical superhet is shown in **Figure 14**. In traditional form, the RF filter is used to limit the input frequency range to those frequencies that include only the desired sum or difference but not the other — the so-called *image* frequency. The dotted line represents the fact that in receivers with a wide tuning range, the input filter often tracks along with the local oscillator. While this is similar to one of the issues with TRF receivers, note that generally there are fewer tuned circuits involved in a superhet, making the tracking easier to accomplish. The IF filter is traditionally used to establish operating selectivity — that required by the information bandwidth.

In many instances, it is not possible to achieve all the receiver design goals with a single-conversion receiver, so multiple conversion steps are taken. Traditionally, the first conversion is tasked with removing the RF image signals, while the second allows processing of the IF signal to provide the information-based IF processing. A second mixer is used to detect the IF signal, translating it to audio. The configuration is shown in Figure 14.

In a typical configuration, the local oscillator (LO) and RF amplifier stages are adjusted so that as the LO is changed in frequency, the RF amplifier is also tuned to the appropriate frequency to receive the desired station. An example may help. Let's pick a common IF frequency used in an AM broadcast radio, 455 kHz. Now if we want to listen to a 600 kHz broadcast station, the RF stage should be set to amplify the 600 kHz signal and the LO should be set to 600 + 455 kHz or 1055 kHz.[6] The 600 kHz signal, along with any audio information it contains, is translated to the IF frequency and is amplified. It is then detected, just as if it were in a TRF receiver at 455 kHz.

Note that to detect standard AM signals, the second oscillator, usually called a *beat frequency oscillator* or BFO, is turned off since the AM station provides its own carrier signal over the air. Receivers designed only for standard AM reception, the typical "table or kitchen radio," generally don't have a BFO at all.

It's not clear yet that we've gained anything by doing this; so let's look at another example. If we decide to change from listening to the station at 600 kHz and want to listen to another station at, say, 1560 kHz, we can tune the single dial of our superhet to 1560 kHz. With the appropriate ganged and tracked tuning capacitors, the RF stage is tuned to 1560 kHz, and the LO is set to 1560 + 455 or 2010 kHz and now that station is translated to our 455 kHz IF. Note that the bulk of amplification can take place at the 455 kHz IF frequency, so not as many stages must be tuned each time we change to a new frequency. Note also that with the superheterodyne configuration the selectivity (the ability to separate stations) occurs primarily in the intermediate-frequency (IF) stages, and is thus the same no matter what frequencies we choose to listen to. The superhet design has thus eliminated the major limitations of the TRF at the cost of two additional building blocks.

It should be no surprise, then, that the superhet, in various flavors has become the primary receiver architecture in use today. Just as WWI was coming to a close, the concept was introduced by Major Edwin Armstrong, a US Army artillery officer. The superhet quickly gained popularity and, following the typical patent battles of the times (and hardly unknown today) became the standard of a generation of vacuum-tube broadcast receivers. These were found in virtually every US home from the late 1920s through the 1960s, when they were slowly replaced by transistor sets, but still of superhet design.

Significant development of the superhet to optimize its potential for use by amateurs occurred at the ARRL in the early 1930s. James Lamb, W1CEI, (later W1AL, now deceased) was *QST* Technical Editor from 1929 to 1939, a period of significant activity. While he had a number of inventions, leading to nine patents, perhaps his most significant invention was the single-signal receiver, described in a series of *QST* articles of the period (and available online from the *QST* archive).[5-8]

The single-signal concept made use of a single piezoelectric crystal resonant at the IF frequency. This formed a filter that was part of the IF amplifier. By carefully adjusting the frequency of the BFO and the crystal phasing control, settings could be found that would eliminate the response on the other side of the zero beat, thus reducing by half the apparent number of signals that the receiver could respond to. Lamb's prototype single-signal receiver is on display at the ARRL Headquarters Museum of Amateur Radio.

This technology was a significant part of communications receivers for at least 50 years, although the use of band-pass filters, rather than the single crystal type of Lamb's design, started to take over with the shift of voice operation to single sideband starting in the 1950s. While Lamb's approach was focused on CW operation with its narrow bandwidth, his filter was also useful for AM voice. By use of the phasing control, a sharp notch in the passband could be used to null out an interfering carrier, such as the heterodyne that resulted from the carriers of two AM transmitters on adjacent frequencies.

Voice Enters the Picture

AM Voice — Like the "Big Boys"

In 1920, KDKA in Pittsburgh, Pennsylvania, became the first federally licensed commercial AM broadcast station in the country. KDKA operated on 1020 kHz, initially using a 100 W vacuum tube transmitter. Amateur Radio operators were using voice communication in the same era with similar technology.

The typical AM transmitter used the same RF equipment that was used for CW. To apply the voice modulation, an audio amplifier stage, typically with a transformer coupled output, was inserted in series with the high voltage supply, providing power to the plate of the final stage of the RF deck. The audio amplifier was adjusted so that the plate voltage would swing from zero to twice the normal anode voltage of the RF stage with the audio voltage. This required an amplifier that could deliver 50% of the dc input power of the final RF stage. A block diagram of the setup is shown in **Figure 15**.

In this scheme, the nonlinear (class C) RF amplifier acted like a high level mixer, providing *high-level modulation* by multiplying the RF signal times the audio signal in its plate circuit. Multiplying (in other words, modulating) a carrier with a single tone results in the tone being translated to frequencies at the sum and difference of the two. An analog view of the signals is shown in **Figure 16**. A representative design of an early modulator was described by James Lamb in a 1931 *QST* article.[9]

Thus, if a transmitter were to multiply a 600 Hz tone by a 600 kHz carrier signal, we would generate additional new signals at 599.4 and 600.6 kHz. If instead we were to modulate the 600 kHz carrier signal with a band of frequencies corresponding to human speech of 300 to 3300 Hz (also called *toll quality* — a term carried over from long-distance telephone systems), we would have additional bands of signals carrying the information and extending from 596.7 to 603.3 kHz, as shown in **Figure 17**. These bands are called *sidebands,* and some form of sidebands is present

in any AM signal that is carrying information.

Note that the total bandwidth of this AM voice signal is twice the highest frequency transmitted, or 6600 Hz. If we choose to transmit speech and limited music, we might allow modulating frequencies up to 5000 Hz, resulting in a bandwidth of 10,000 Hz or 10 kHz. This is the standard channel spacing that commercial AM broadcasters use in the US. In actual use, stations on adjacent channels are generally separated geographically, so broadcasters can extend some energy into the adjacent channels for improved fidelity. We refer to this as a *narrow-bandwidth* mode.

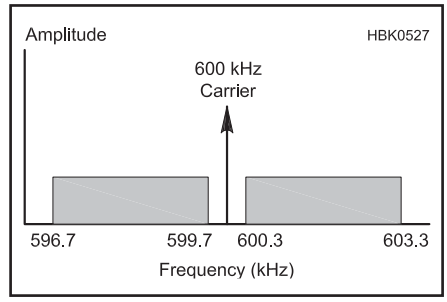

Figure 17 — Graphical representation of spectrum of the amplitude modulation used to send a voice signal on a 600 kHz carrier.

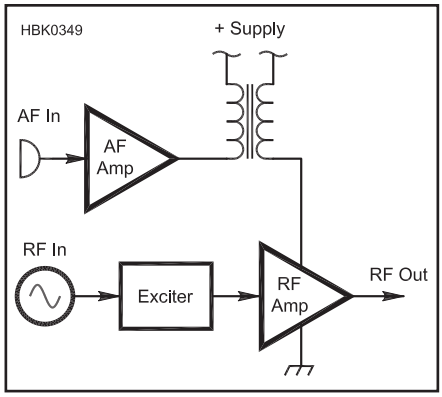

Figure 15 — Elements of an amplitude modulated transmitter using high-level modulation.

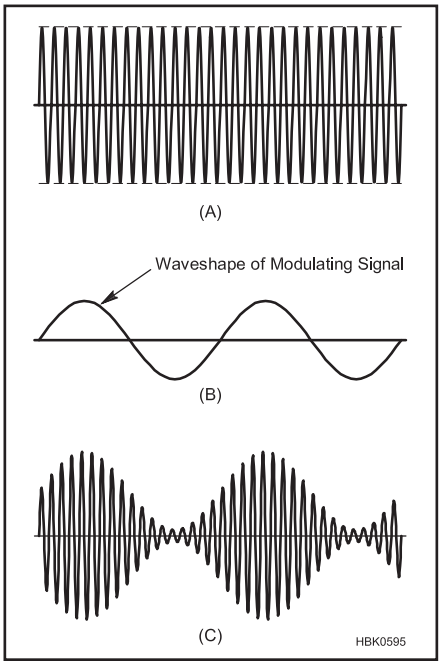

Figure 16 — Graphical representation of the amplitude modulation process that occurs in the system of Figure 15. At (A) the RF signal, at (B) the audio (modulating) waveform and at (C) the RF signal modulated by the audio signal using high-level

What does this say about the bandwidth needed for a receiver? If we want to receive the full information content transmitted by a US AM broadcast station, then we need to set the bandwidth to at least 10 kHz. What if our receiver has a narrower bandwidth? Well, we will lose the higher frequency components of the transmitted signal — perhaps ending up with a radio suitable for voice but not very good at reproducing music.

On the other hand, what is the impact of having too wide a bandwidth in our receiver? In that case, we will be able to receive the full transmitted spectrum, but we will also receive some of the adjacent channel information. This will sound like interference and reduce the quality of the desired signal we are receiving. If there are no adjacent channel stations, we will receive any additional noise from the additional bandwidth and minimal additional information. The general rule is that the received bandwidth should be matched to the bandwidth of the signal we are trying to receive to maximize *signal-to-noise ratio* (*SNR*) and to minimize interference.

As the receiver bandwidth is reduced, intelligibility suffers, although the SNR is improved. With the carrier centered in the receiver bandwidth, most voices are difficult to understand at bandwidths less than around 4 kHz. In cases of heavy interference, full carrier AM can be received as if it were SSB, as described later, with the carrier inserted at the receiver and the receiver tuned to whichever sideband has the least interference. A view of a typical amateur AM and CW station from the 1950s is shown in **Figure 18**.

Single Sideband Suppressed Carrier

In looking at Figure 17, you might have noticed that both sidebands carry the same information, and are thus redundant. In addition, the carrier itself conveys no information. It is thus possible to transmit a *single sideband* and *no carrier*, as shown in **Figure 19**, relying on the BFO (beat frequency oscillator) in the receiver to provide a signal with which to multiply the sideband in order to provide demodulated audio output.

The implications in the receiver are that the bandwidth can be slightly less than half that required for double sideband AM (DSB). There must be an additional mechanism to carefully replace the missing carrier within the receiver. This is the function of the BFO, which must be at just exactly the right frequency. If the frequency is set improperly, even by a few Hz, a baritone can come out sounding like a soprano and vice-versa!

The standard commercial AM format is very convenient for receiver design, since the carrier needed to demodulate the received sidebands is sent along with the sidebands. Applying the total received signal to a detector or multiplier allows the audio to be recovered without having to worry about any of the finer points we will discuss later. This is very cost-effective in a broadcast environment in which there are many inexpensive receivers and only a relatively few expensive transmitters. For amateur or commercial point-to-point use, there is usually a single receiver associated with the transmitter on any particular link, and it makes sense to divide the complexity more evenly to reduce the total system cost.

For reception of CW, suppressed-carrier single-sideband voice (SSB) or on-off or

Figure 18 — A typical 1950s amateur AM and CW station used a separate transmitter and receiver. Shown here is W1ZR's 1950s-era station with an E. F. Johnson Viking II transmitter (100 W) with separate VFO on the left and a National HRO-60 receiver on the right.

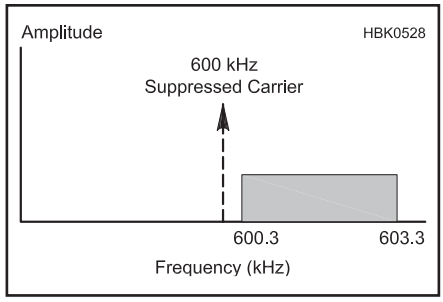

Figure 19 — Graphical representation of the spectrum of a single sideband AM voice signal adjacent to its suppressed 600 kHz carrier.

frequency-shift keyed (FSK) signals, a second *beat frequency oscillator* or BFO is employed to provide an audible voice or an audio tone (or tones) at the output for the operator to receive by ear (or for further processing of digital modes such as FSK). This is the same as a heterodyne mixer with an output centered at dc, although the IF filter is usually designed to remove one of the output products.

This creates a requirement for a much more stable receiver design with a much finer tuning system — a more expensive proposition. An alternate is to transmit a reduced level carrier and have the receiver lock on to the weak carrier, usually called a *pilot* carrier. Note that the pilot carrier need not be of sufficient amplitude to demodulate the signal, just enough to allow a BFO to lock to it. These alternatives are effective, but they tend to make SSB receivers expensive, complex and most appropriate for the case in which a small number of receivers are listening to a single transmitter, as is the case of two-way communication.

Note that the bandwidth required to demodulate an SSB signal effectively is actually less than half that required for an AM signal because frequencies immediately adjacent to the AM carrier need not be received. Thus the toll quality spectrum of 300 to 3300 Hz can be received in a bandwidth of 3000 not 3300 Hz. Early SSB receivers typically used a bandwidth of around 3 kHz, but with the heavy interference frequently found in the amateur bands, it is more common for amateurs to use bandwidths of 1.8 to 2.4 kHz with a corresponding loss of some high and low frequency speech components.

Single-sideband makes it possible to transmit the information with just one of the sidebands and no carrier. In so doing, we use somewhat less than half the bandwidth, a scarce resource, and also consume much less transmitter power by not transmitting the carrier and the second sideband.

Managing Selectivity and Images

After WWII, activity increased dramatically on the amateur bands and across the HF spectrum generally. This generated higher performance requirements for receiver selectivity and for the rejection of *images* — undesired reception of signals at alternate frequencies that also generate mixing products at the receiver's IF. Advances in receiver design during the 1940s were applied and several improvements of the superheterodyne architecture appeared.

Double and Triple Conversion

In the 1950s, a new technique called *double-conversion* was defined to solve the image problem. Rather than having to decide between a high IF frequency for good image rejection, or a low IF frequency for narrow channel selectivity, some bright soul decided to do both! As shown in **Figure 20**, a conversion of the desired signal to a relatively high IF is followed by a second conversion to a lower IF to set the selectivity. This arrangement solved the image problem nicely, while the rest of the receiver could be pretty much kept the same as before. A number of manufacturers offered revised receivers in the fifties that added an additional conversion stage, often just above 10 or 15 MHz, regions that had poor image rejection in the original design. Note that major changes were not required in the receiver, and some looked just like their single-conversion predecessors from the outside. The extra stage was shoehorned in, with just a retuning of the first local oscillator required.

Some manufacturers decided that if two conversions were good, three could be even better. These designs often began with the same basic architecture, but converted the 455 kHz second IF down to a third IF, often in the 50 to 100 kHz range, for even sharper selectivity. This approach lasted until crystal lattice filters that outperformed the low frequency LC circuits became available, eliminating the need for the additional conversion stage.

The Collins System

In the 1940s, a visionary radio pioneer, Arthur Collins, WØCXX, founder of Collins Radio in the 1930s, came up with another approach to double-conversion. One problem with earlier MF and HF receivers was that they used a standard LC (first) oscillator using a variable capacitor of the type used in

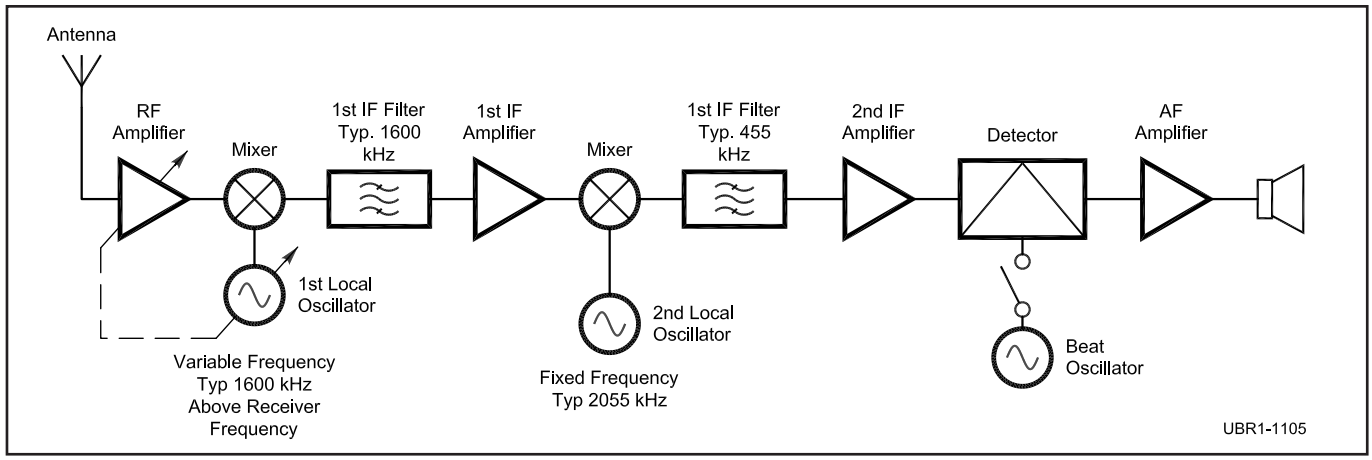

Figure 20 — An early type double-conversion superhet receiver.

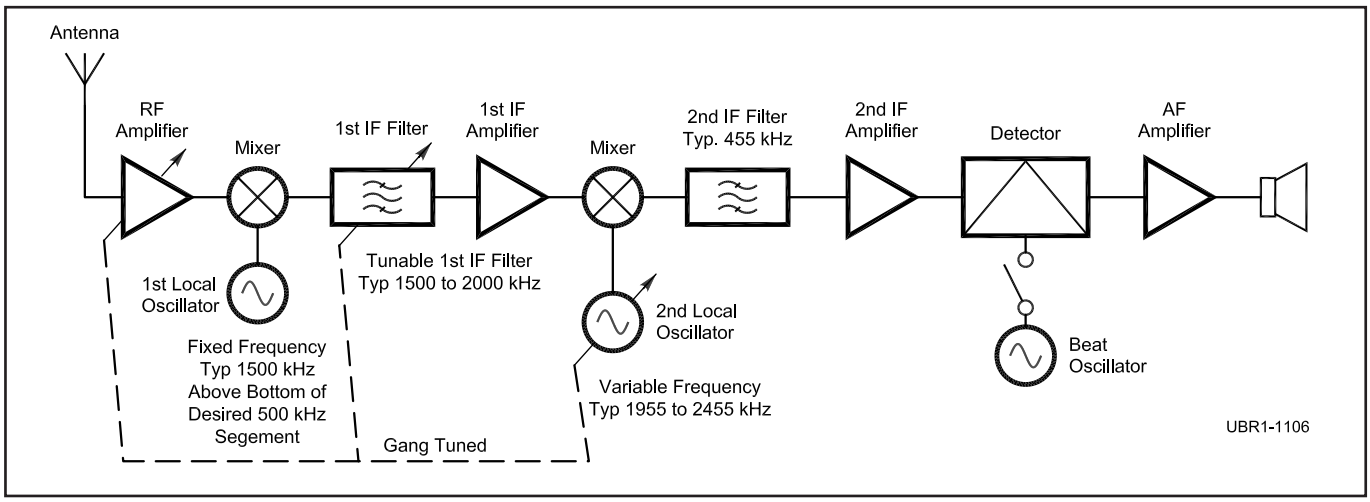

Figure 21 — A double-conversion superhet receiver using the Collins system desribed in the text.

broadcast receivers. These capacitors typically had a 9:1 capacitance range, resulting in a 3:1 frequency range. The typical tuning arrangement for multiband receivers was 0.5 to 1.5, 1.5 to 4.5, 4.5 to 13.5, and 13.5 to 30 MHz. This covered frequencies from the AM broadcast band to the top of the HF range in four bands.[7] With this arrangement, the tuning rate and dial calibration marks became less and less precise as higher bands were selected, making tuning difficult.

The Collins system (see **Figure 21**) switched the variable oscillator to the second mixer and used crystal-controlled oscillators rather than variable oscillators, in the first position. Although many more bands were required (30 for the famous 51J series that covered 0.5 to 30.5 MHz), each tuned with exactly the same tuning rate. Collins went a step further and designed an inductance-tuned oscillator (usually called a *permeability-tuned oscillator* or *PTO*) that was linear throughout its range. These oscillators could be tuned to the nearest 1 kHz from beginning to end, avoiding the tuning uncertainty of other receiver types. For this to work properly, synchronized or gang tuning is required between the oscillator, first IF variable filter and RF stages. This was no problem for the engineers at Collins Radio, but their equipment brought a premium price.

The Pre-Mixed Arrangement

A third approach to double-conversion was a mix of the first two (no pun intended). The pre-mixed arrangement (**Figure 22**) uses a single variable oscillator range, as with the Collins design, but does the mixing outside of the signal path. This avoids the need for variable tuning at the first IF, allowing a tighter roofing filter (first IF filter), but the trade-off is the need for filtering at the output of the pre-mixer (not shown).

High Frequency Crystal Lattice Filters

Just as double-conversion receivers were settling into becoming the "way things were done," the use of piezoelectric crystals as elements of band-pass filters became feasible. While crystal lattice filters at 455 kHz were popular following WWII because of availability of large stocks of surplus crystals at that range, they didn't address the image response issue. They just provided more selective filter responses at the same IF frequency. Starting in the 1950s, crystal lattice filter technology moved higher in the MF region and into HF, improving receiver selectivity dramatically.

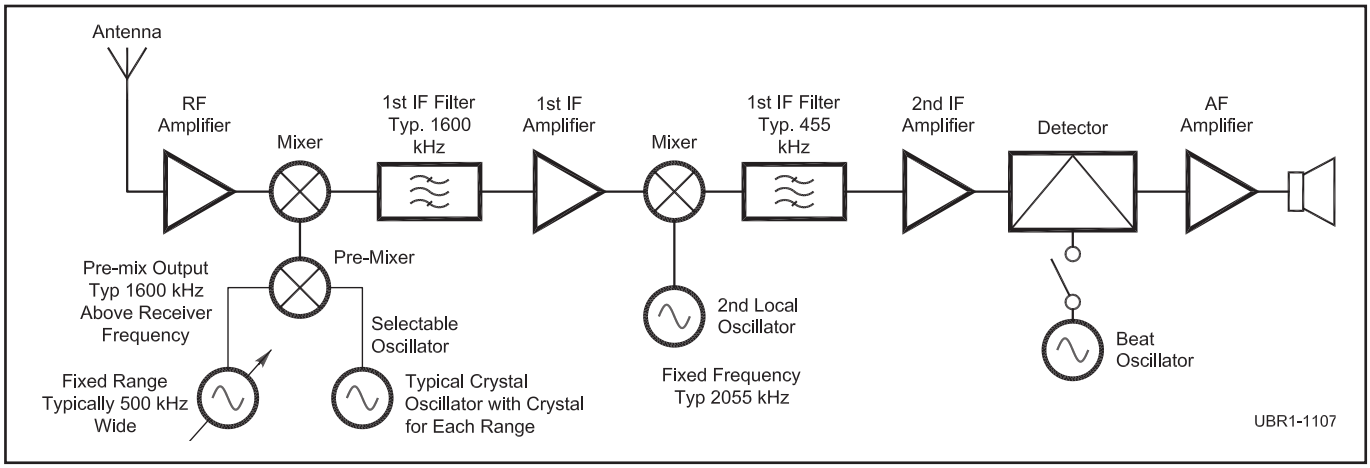

Figure 22 — A double-conversion superhet receiver with the pre-mixed local oscillator configuration.

Transmitting Single Sideband

There are a number of ways of generating a single sideband signal for transmission. While other approaches are possible, the most popular technique used in Amateur Radio, and most other services, is called the *filter method*, in which a selective filter is used to eliminate the undesired sideband from a DSB suppressed carrier signal. The next most frequently encountered technique is called the *phasing method*, which takes advantage of the trigonometric properties of the sinusoidal waves that we're dealing with. We will describe both here. (Both methods are described in more detail in the **Modulation** chapter of *The ARRL Handbook*.)

The Filter Method of SSB Generation

The block diagram of a simple single-sideband suppressed-carrier (SSB) transmitter is shown in **Figure 23**. This transmitter uses a balanced mixer as a *balanced modulator* to generate a double-sideband suppressed-carrier signal without a carrier. (See the **Mixers, Modulators, and Demodulators** chapter of *The ARRL Handbook* for more information on these circuits.) That signal is then sent through a filter designed to pass just one (either one, by agreement with the receiving station) of the sidebands.[10] Depending on whether the selected sideband is above or below the carrier frequency, the signal is called *upper sideband* (*USB*) or *lower sideband* (*LSB*), respectively. The resulting SSB signal is amplified to the desired power level and we have an SSB transmitter.

While a transmitter of the type in Figure 23 with all processing at the desired transmit frequency will work, the configuration is not commonly used. Instead, the carrier oscillator and sideband filter are often placed at an intermediate frequency that is heterodyned to the operating frequency as shown in **Figure 24**. The reason is that the sideband filter is a complex narrow-band filter and most manufacturers would rather not have to supply a new filter design every time a transmitter is ordered for a new frequency. Many SSB transmitters can operate on different bands as well, so this avoids the cost of additional mixers, oscillators and expensive filters.

Note that the block diagram of our SSB transmitter bears a striking resemblance to the diagram of a superheterodyne receiver as shown in Figure 14, except that the signal path is reversed to begin with information and produce an RF signal. The same image rejection requirements for intermediate frequency selection that were design constraints for the superhet receiver apply here as well.

The Phasing Method

Most current transmitters use the method of SSB generation shown in Figure 23 and discussed in the previous section to generate the SSB signal. This is done in two steps — first a balanced modulator is used to generate sidebands and eliminate the carrier, then a filter is used to eliminate the undesired sideband, and often to improve carrier suppression as well.

The *phasing method* of SSB generation uses two balanced modulators and a phase-shift network for both the audio and RF carrier signals to produce the upper sideband signal as shown in **Figure 25**. By a shift in the sign of either of the phase-shift networks, the opposite sideband can be generated. This method trades a few phase-shift networks and an extra balanced modulator for the sharp sideband filter of the filter method. While it looks deceptively simple, a limitation is in the construction of a phase-shift network that will have a constant 90° phase shift over the whole audio range. Errors in phase shift result in less than full carrier and sideband suppression. Nonetheless, there have been some successful examples offered over the years.

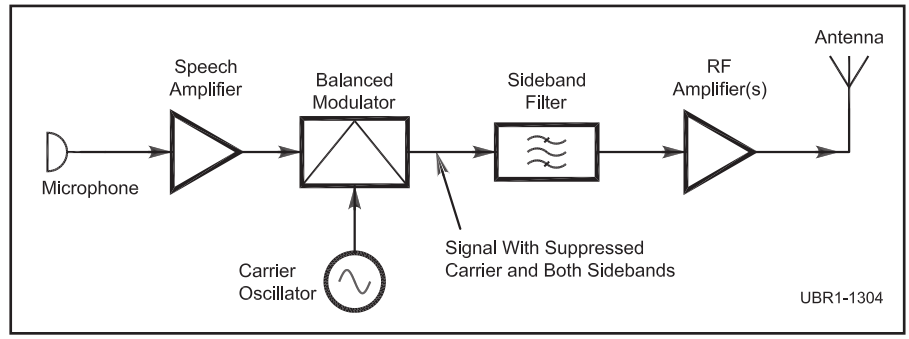

Figure 23 — Block diagram of a single-frequency filter-type single sideband suppressed carrier (SSB) transmitter.

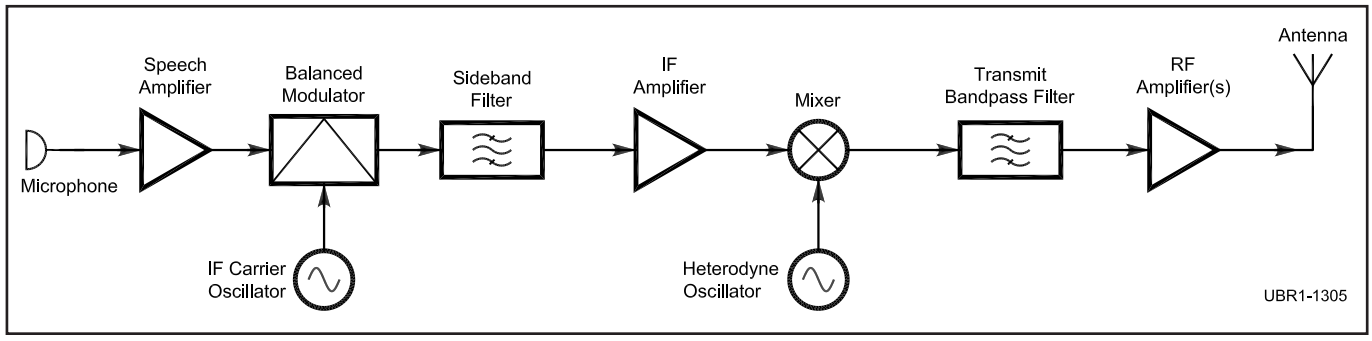

Figure 24 — Block diagram of a heterodyne filter-type SSB transmitter for multiple frequency operation.

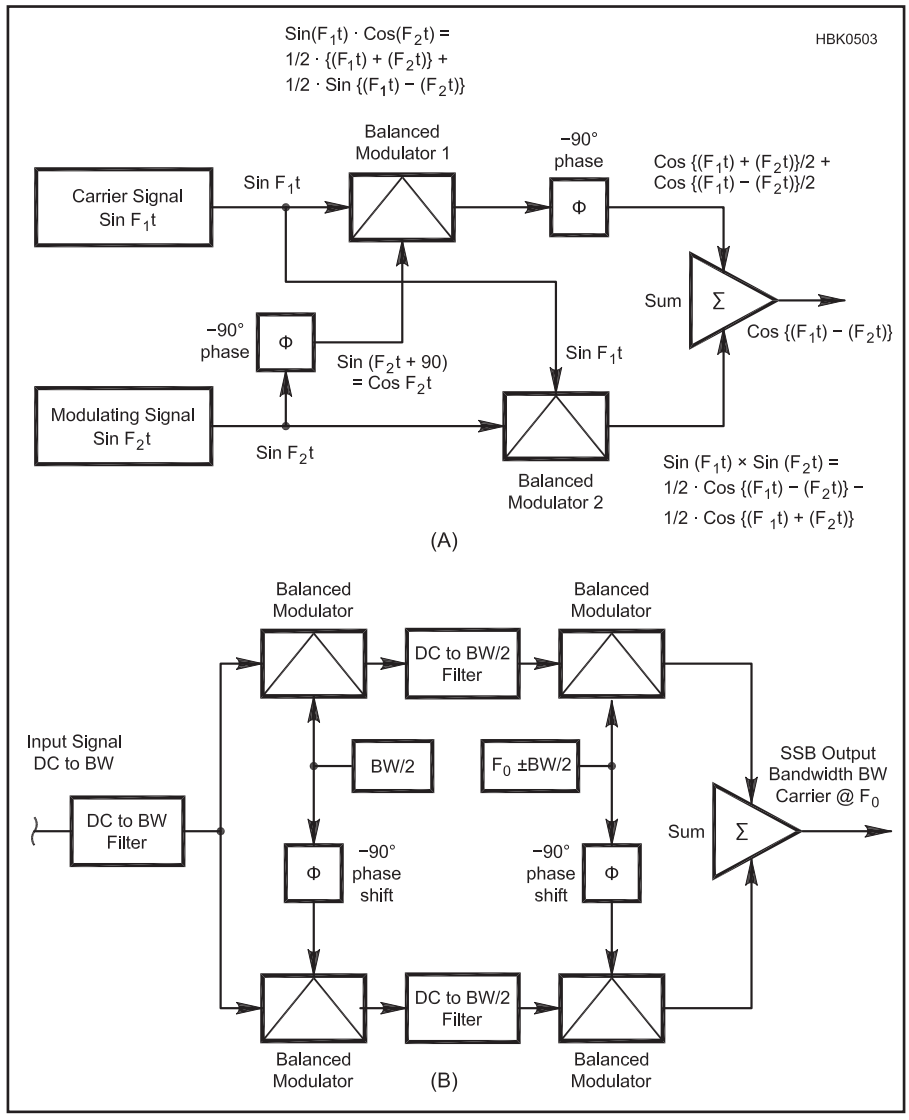

Figure 25 — Block diagrams of phasing type SSB transmitters for single frequency operation.

The SSB Transceiver is Born

A major evolution step of Amateur Radio technology was the migration from separate transmitters and receivers for SSB to a single package that contained both. This was called a *transceiver*, a combination of the words *trans*mitter and re*ceiver*. This was based on the observation that, unlike AM transmitters of the period, most SSB transmitter designs looked a lot like the architecture of a modern receiver, except configured in reverse, as in the form of a heterodyne transmitter as shown in Figure 24.

Some early heterodyne SSB transmitters and transceivers took advantage of the fact that the 9 MHz heterodyne oscillator frequency could actually be used on both 80/75 and 20 meters with the same 5.0 to 5.5 MHz

Figure 26 — An early SSB transceiver, the 1957 vintage Collins KWM-2(A). [Photo courtesy of the Collins Collectors Association]

frequency range in the VFO. The difference product at 9 MHz minus the VFO frequency (4 to 3.5 MHz) was used for the lower band, while the sum product at 9 MHz plus the VFO frequency (14 to 14.5 MHz) was used for 20 meters. This was a popular dual-band scheme in the early days of SSB voice. One drawback was that the VFO tuned across the band from low to high frequencies in the opposite direction on each band. For many hams of the day, it was a small price for the simplicity.

The transmit-receive (TR) switching arrangements in transceivers allowed use of the same oscillators and filters for both receiving and transmitting, resulting in the transmitter automatically being tuned to the receive frequency. This was a significant advantage for SSB operation because the two frequencies needed to be close together to provide intelligible speech. One of the earliest successful HF SSB transceivers was the KWM-2 series by Collins Radio, shown in **Figure 26**.

Most SSB transceivers did not include a built in power supply, and with the elimination of the need for the heavy modulation transformers of AM transmitters, the resulting transceiver could be light and compact. Different power supplies could be selected for operation from ac mains for home station operation or from dc systems for mobile operation in vehicles. To say that the HF SSB transceiver revolutionized Amateur Radio would not be an overstatement. Fortunately, all the benefits of the transceiver for SSB operation were also beneficial for CW and data modes.

Narrowband Frequency Modulation

Another popular voice mode is *frequency modulation*, or FM. FM can be found in a number of variations depending on purpose. In Amateur Radio and commercial mobile communication use on the shortwave bands, it is universally *narrowband FM* or *NBFM*. In NBFM, the frequency deviation is limited to around the maximum modulating frequency, typically 3 kHz. The bandwidth requirements at the receiver can be approximated by 2 × (D + M), where D is the deviation and M is the maximum modulating frequency. Thus 3 kHz deviation and a maximum voice frequency of 3 kHz results in a bandwidth of 12 kHz, not far beyond the requirements for broadcast AM. (Additional signal components extend beyond this bandwidth, but are not required for voice communications.)

In contrast, broadcast or *wideband FM* or *WBFM* occupies a channel width of 150 kHz. Originally, transmitted signal bandwidth was extended to provide high-fidelity reproduction of up to 15 kHz audio with an improved signal-to-noise ratio (SNR). However, with multiple channel stereo and sub-channels all in the same allocated bandwidth, the primary channel is today limited to about the maximum audio signal bandwidth.

In the US, FCC amateur rules limit wideband FM use to frequencies above 29 MHz. Some, but not all, HF communication receivers provide for FM reception. For proper FM reception, two changes are required in the receiver architecture as shown within the dashed line in **Figure 27**. The fundamental change is that the detector must recover information from the frequency variations of the input signal. The most common type is called a *discriminator*. The discriminator does not require the BFO, so the BFO is turned off or eliminated in a dedicated FM receiver. Amplitude variations convey no information in FM, so they are generally eliminated by a *limiter*. The limiter is a high-gain IF amplifier stage that clips the positive and negative peaks of signals above a certain threshold. Most noise of natural origins is amplitude modulated, so the limiting process also strips away noise from the signal.

Transmitters using frequency modulation (FM) or phase modulation (PM) are generally grouped into the category of *angle modulation* since the resulting signals are often indistinguishable. An instantaneous change in either frequency or phase can create identical signals, even though the method of modulating the signal is somewhat different. To generate an FM signal, we need an oscillator whose frequency can be changed by the modulating signal.

We can make use of an oscillator whose frequency can be changed by a "tuning voltage." If we apply a voice signal to the tuning voltage connection point, we will change the frequency with the amplitude and frequency

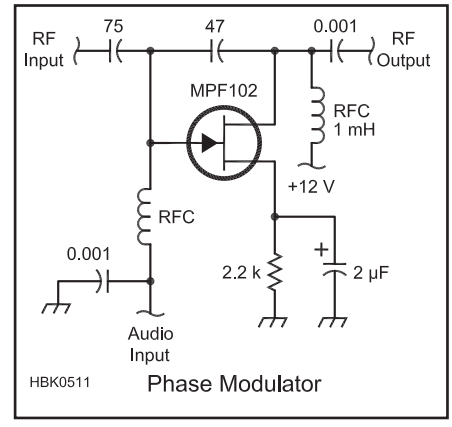

Figure 28 — Simple FET phase modulator circuit.

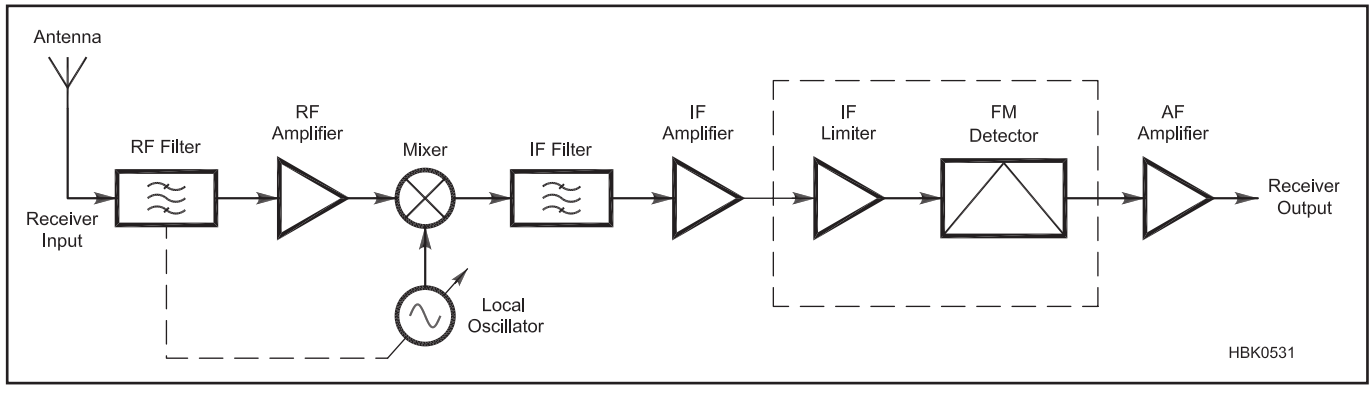

Figure 27 — Block diagram of an FM superhet receiver. Changes from an AM receiver are shown in the dashed box.

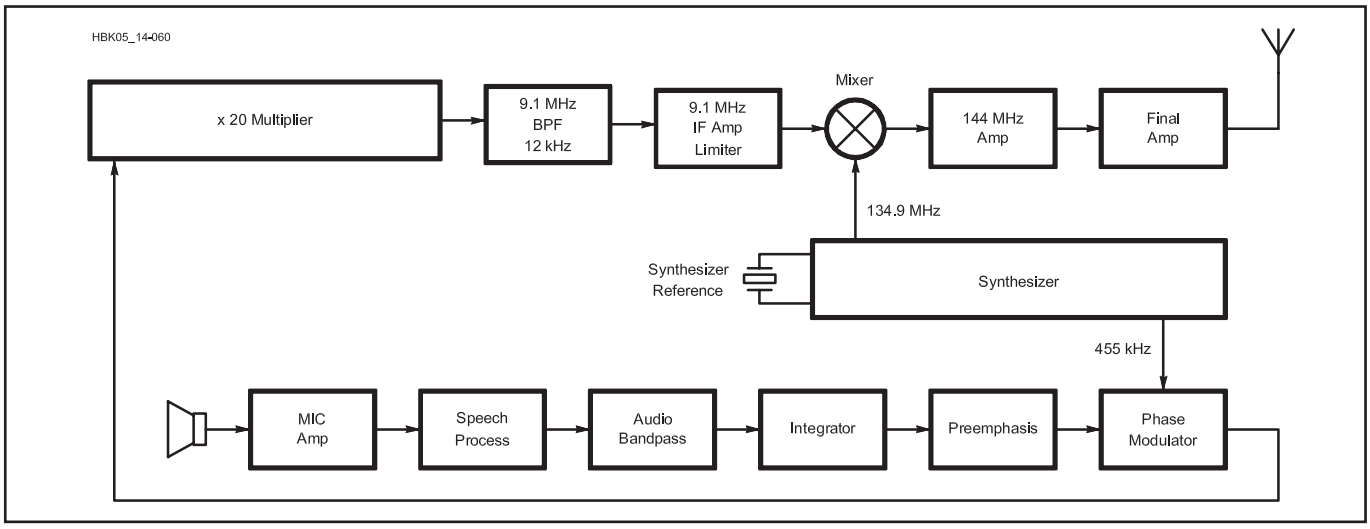

Figure 29 — Block diagram of a VHF/UHF NBFM transmitter using the indirect FM (phase modulation) method.

of the applied modulating signal, resulting in an FM signal.

The phase of a signal can be varied by changing the values of an R-C phase-shift network. One way to accomplish phase modulation is to have an active element shift the phase and generate a PM signal. In **Figure 28**, drain current through the MPF102 field-effect transistor is varied with the applied modulating signal, varying the phase shift at the stage's output. Because the effective load on the stage is changed, the carrier is also amplitude-modulated and must be run through an FM receiver-type limiter in order to remove the amplitude variations. A block diagram of a complete NBFM transmitter using the phasing method is shown in **Figure 29**.

Computers Enter the Scene

Computers have had a strong impact on all facets of life, and Amateur Radio has not escaped the influence of computer technology in a number of areas.

Operation Using Digital Modes

Communication via digital signaling has been popular with radio amateurs since teletype equipment became readily available on the military surplus market following World War II. The availability of personal computers (PCs) with sound cards and specialized software has made many forms of digital transmission easy for amateurs to accomplish without the need for special and often complex hardware. While most new digital modes are more robust and capable than traditional *radioteletype* or RTTY (pronounced "ritty"), it still remains one of the most popular modes bridging the divide between old and new technology.

Radioteletype transmission makes use of the Baudot code constructed with sequences of ON-OFF elements or bits as shown in **Figure 30.**[11] The state of each bit — ON or OFF — is represented by a signal at one of two distinct frequencies: one designated *mark* and one designated *space*. (These states are named after the early Morse tape readers that placed a pen *mark* on a paper tape when the key was down and made a *space* when the key was up.) This is referred to as *frequency shift keying* (*FSK*). The transmitter frequency shifts back and forth with each character's individual elements.

Amateur Radio operators typically use a 170 Hz separation between the mark and space frequencies, depending on the data rate and local convention, although 850 Hz is sometimes used. The minimum bandwidth required to recover the data is somewhat greater than twice the spacing between the tones. Note that the tones can be generated by directly shifting the carrier frequency (*direct FSK*), or by using a pair of 170 Hz spaced audio tones applied to the audio input of an SSB transmitter (*audio FSK* or AFSK), often supplied by a computer sound card. Direct FSK and AFSK sound the same to a receiver.

Note that if the standard audio tones of 2125 Hz (mark) and 2295 Hz (space) are used, they fit within the bandwidth of a voice channel and thus a voice transmitter and receiver can be employed without any additional processing needed outside the radio equipment. Alternately, the receiver can employ detectors for each frequency and provide an output directly to a computer.

Some receivers provide a narrow CW bandwidth filter with the center frequency shifted to midway between the tones (2210 Hz). The

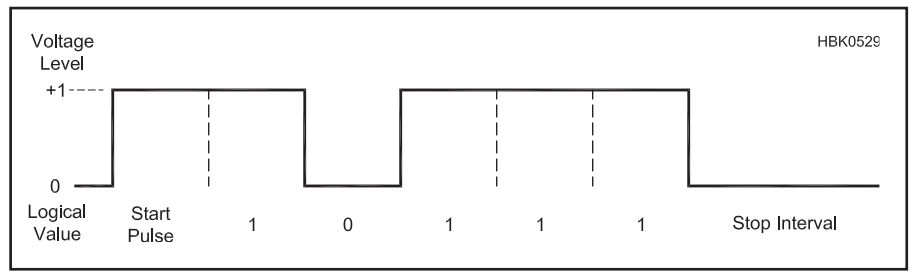

Figure 30 — Voltage pulses corresponding to the letter "A" in Baudot code, including start and stop pulses.

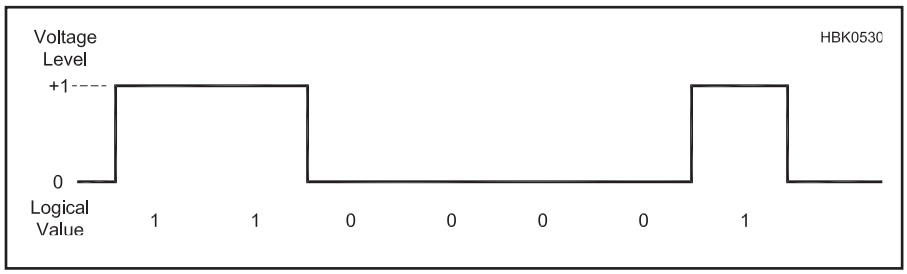

Figure 31 — Voltage pulses corresponding to the lower case letter "a" in ASCII code.

most advanced receivers provide a separate filter for mark and space frequencies, thus maximizing interference rejection and signal-to-noise ratio (SNR). Using a pair of tones for FSK or AFSK results in a maximum data rate of about 1200 bits per second (bps) over a high-quality voice channel.

Phase shift keying (*PSK*) can also be used to transmit bit sequences, requiring good frequency stability to maintain the required time synchronization to detect shifts in phase. If the channel has a high SNR, as is often the case at VHF and higher, telephone network data-modem techniques can be used.

At HF, the signal is subjected to phase and amplitude distortion as it travels. Noise is also substantially higher on the HF bands. Under these conditions, modulation and demodulation techniques designed for "wireline" connections become unusable at bit rates of more than a few hundred bps. As a result, amateurs have begun adopting and developing state-of-the-art digital modulation techniques. These include the use of multiple carriers (MFSK, Clover, PACTOR III, and others), multiple amplitudes and phase shifts (QAM and QPSK techniques), and advanced error detection and correction methods to achieve a data throughput as high as 3600 bps over a voice-bandwidth channel. (Spread-spectrum techniques are also being adopted on the UHF bands, but are beyond the scope of this discussion.)

Another code that is in common use is the ASCII (American Standard Code for Information Interchange) code that was developed for computer-to-computer communication. This is easy to generate on a PC, since it is the code that PCs use to communicate with each other. Unlike the five-unit Baudot code that can send 2^5, or 32, distinct characters and thus uses a special character to toggle between *letters* and *figures*, the seven unit ASCII code (see **Figure 31**) can send 2^7 (128) characters, enough for most printing characters including upper and lower case letters of most languages, punctuation, numerals, special characters and control signals.

The bandwidth required for data communications can be as low as 100 Hz for PSK31 to 1 kHz or more for the faster speeds of PACTOR III and Clover. Beyond having sufficient bandwidth for the data signal, the primary requirements for receivers used for data communications are linear amplitude and phase response over the bandwidth of the data signal. The receiver must also have excellent frequency stability to avoid drift, and excellent frequency resolution to enable the receiver filters to be set on frequency.

Computer Control of Radios

Most modern amateur HF transceivers provide a connection to allow information exchange with PCs. Depending on the internal architecture of the transceiver, it may be possible to completely control the transceiver through the use of the PC. This can permit remote operation using Internet communications between local and distant PCs, allowing radio operation far from the home station. In addition, logging software can make use of the link to receive frequency and mode information from the radio to automatically fill in the log, or to change frequency to find a particular station that has been reported to the PC from a remote network.

Software Defined Radio

Moving beyond processing of digital communications signals and the control of transceivers, computers have taken over many functions within radio transceivers through the use of *digital signal processing* (*DSP*). Early adaptations used PCs to perform digital signal processing on audio signals from traditional receivers, while newer technology involves building the entire radio around specialized high performance DSP integrated circuits.

Modern receivers using DSP for operating bandwidth and information detection often have an additional conversion step to a final IF at a frequency in the tens of kHz. This is established by the maximum sampling rate of a finite frequency response analog-to-digital converter (ADC). Advances in the art have resulted in ADC and processor speeds fast enough that the ADC has been moving closer and closer to the RF frequency — resulting in fewer required conversions, and in some cases, converting the RF signal directly to digital data.

The design choices described above are still found in current receiver architectures, but some recent advances have added a few new twists. While advances in microminiaturization of all circuit elements have made a radical change in the dimensions of communication equipment, perhaps most significant technology impacts on architecture come in two areas — the application of digital signal processing and direct digital synthesis (DDS) frequency generation. Both are discussed in detail in the chapter on **DSP and Software Radio Design** in *The ARRL Handbook*. An overview is presented here.

Digital signal processing provides a level of bandwidth setting filter performance not practical with other technologies. While much better than most low frequency IF LC bandwidth filters, the very good crystal or mechanical bandwidth filters in amateur gear are not very close to the rectangular shaped frequency response of an ideal filter, but rather have skirts with a 6 to 60 dB response of perhaps 1.4 to 1. That means if we select an SSB filter with a nominal (6 dB) bandwidth of 2400 Hz, the width at 60 dB down will typically be 2400 × 1.4, or 3360 Hz. Thus a signal in the next channel that is 60 dB stronger than the signal we are trying to copy (as often happens) will have energy just as strong as our desired signal.

DSP filtering approaches the ideal response. **Figure 32** shows the response of a DSP bandwidth filter with a 6 dB bandwidth of 2400 Hz as measured by the ARRL Lab. Note how rapidly the skirts drop to the noise level. In addition, while analog filtering generally requires a separate filter assembly for each desired bandwidth, DSP filtering is adjustable — often in steps as narrow as 50 Hz — in both bandwidth and center frequency. In addition to bandwidth filtering, the same DSP can often provide digital noise reduction and digital notch filtering to remove interference from fixed-frequency carriers.

Early DSP filters operated in the audio frequency range, based on the frequency

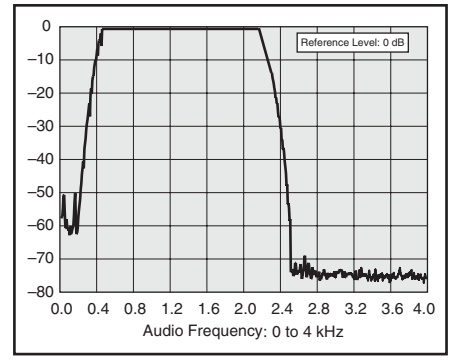

Figure 32 — Frequency response of an aftermarket 2400 Hz bandwidth DSP audio filter as measured in the ARRL Lab.

response limitations of the analog-to-digital converter (ADC) that provides the digitization for further processing. The limitation was based on the Nyquist sampling theorem (AT&T, 1924) that established that a periodic signal could be sampled and then reconstructed from its samples if sampled at least twice during each cycle of its highest frequency component. Early ADCs were limited to sampling rates less than about 50 kHz. Such a DSP filter could be added to the audio signal at the output of any receiver and work its magic on the signals in the audio stream. Unfortunately, at that time ADC and DSP that worked at typical receiver IF frequencies were either unavailable, or prohibitively expensive for amateur use.

Receiver designers quickly resurrected the earlier triple conversion receiver architecture —now with IF frequencies even lower, in the 12-20 kHz range — so that DSP could be employed in the IF rather than audio section of receivers. This allowed the receiver to eliminate interfering signals before the application of automatic gain control, avoiding gain reduction in the presence of off-channel signals. As time and technology have advanced, the clocking speed and input bandwidth of ADCs continues to increase while prices continue to fall, allowing the third IF frequency to move toward the second. Soon the additional conversion step will no longer be required, and we will see high-performance HF receivers that apply signals at the input frequency directly to the ADC.

NOTES

[1]R. Shrader, W6BNB (SK), "When Radio Transmitters Were Machines," *QST*, Jan 2009, pp 36-38.

[2]The only remaining Alexanderson alternator transmitter is the Grimeton VLF transmitter (**http://en.wikipedia.org/wiki/Grimeton_VLF_transmitter**) that operates on 17.2 kHz.

[3]R. Shrader, W6BNB (SK), "Radio Gear of Yesteryear," *QST*, Mar 1994, pp 41-43, 57.

[4]ARRL Staff, "A Short Wave Regenerative Receiver," *QST*, Dec 1916, pp 3-5.

[5]**www.arrl.org/arrl-periodicals-archive-search**

[6]J. Lamb, W1CEI, "What's Wrong With Our C.W. Receivers?," *QST*, Jun 1932, pp 9-17.

[7]J. Lamb, W1CEI, "Short-Wave Receiver Selectivity to Match Present Conditions," *QST*, Aug 1932, pp 9-16.

[8]J. Lamb, W1CEI, "An Intermediate-Frequency and Audio Unit for the Single-Signal Superhet," *QST*, Sep 1932, pp 9-15.

[9]J. Lamb, W1CEI, "High-Power Performance from the Small 'Phone Transmitter," *QST*, Dec 1931, pp 10-23.

[10]Amateur practice is to use USB above 10 MHz and LSB on lower frequencies. The exception is 60 meter channels, on which amateurs are required to use USB.

[11]The Baudot code is still the standard employed in the wire-line teletype (TTY) keyboard terminals used by hearing impaired people.