CRYSTAL SETS TO SIDEBAND

Chapter 4B
HERTZIAN WAVES IN THE BASEMENT (Continued)

****

RECREATING HERTZ'S RADIO EQUIPMENT

Most of what I've read about the history of radio was written by non-engineers. They describe the revolutionary apparatus invented by our heroes using old-time radio terms like "earth resonances," "aether," and "coherers." They tell us how far it transmitted, but they give us only the faintest clues about how the gizmo actually worked. Was it a spark gap transmitter? A high-speed alternator? What the heck was a "Tesla oscillator" anyway?

In 1884 James Maxwell published four equations that quantified and connected magnetism with electric phenomena. These equations also predicted the existence of radio waves. The changing magnetic and electric fields related to each other with sine functions. So, once physicists had the equations to stare at, it wasn't too huge a leap to conclude that out of phase, sinewave-shaped electric and magnetic fields would generate each other in an oscillation and radio energy would propagate through space.

In 1889 Heinrich Hertz, a physics professor at the University of Bonn, Germany, was the first to demonstrate radio waves in the laboratory. Of course he might have done this in 1884 or 1887, depending on which website you visit. Ah, the glorious information age!

That's fascinating, but **HOW** did he demonstrate radio waves? Using 1880 technology, that could not have been easy. How did he know he was detecting waves and not just magnetic coupling from one coil to another? Or if his antenna was capacitive, how did he know he wasn't observing capacitive coupling? ***If I were skeptical about the existence of radio waves, but I understood the full implications of Maxwell's equations, I would be convinced if I could see communication across a distance greater than one wavelength.*** A minimum of one wavelength means that "the alleged electromagnetic wave" would change from magnetic to electric field energy then back again at least once. ***Of course, I would also want to see evidence of standing waves and a way to measure frequency.***

Demonstrating Hertzian waves

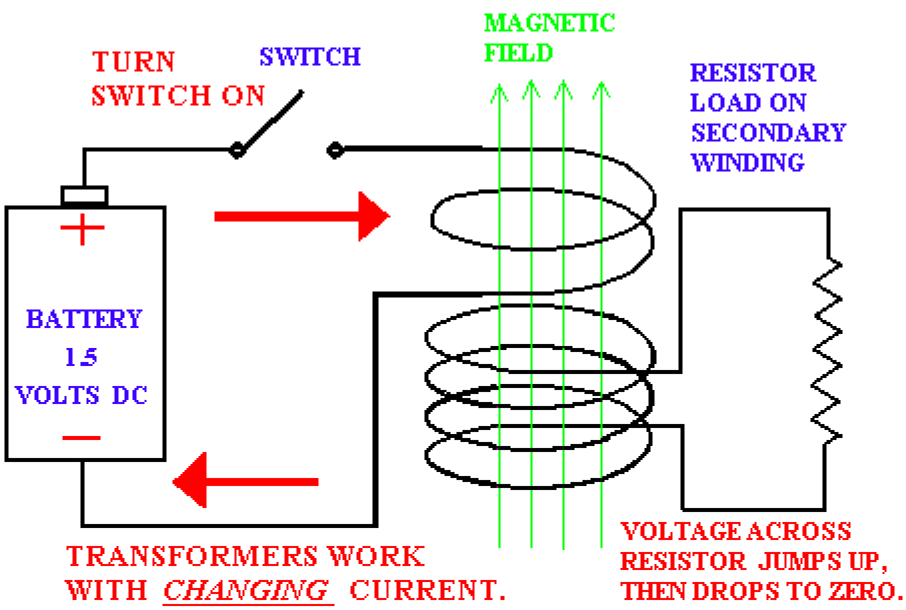
Suppose you were living in the year 1884 and Maxwell had just predicted the existence of radio waves. Using components available in that era, how would you generate Hertzian waves and get those waves named after you instead of Hertz? If you are able to generate radio waves, how could you prove to a skeptic that you had actually done it? Hertz managed this feat and apparently his demonstration was convincing. Otherwise the unit of measurement for frequency would not be the ***Hertz***. One Hertz (Hz) equals one cycle (one complete oscillation) per second.

When I first had the idea of pretending to be Dr. Hertz, I was not able to find a description of his apparatus. That was a good thing because it forced me to invent my own method to demonstrate Hertzian waves. If you already know enough about electricity to be able to handle the challenge, then get out your 1880 hardware and build a transmitter and receiver that will transmit at least one wavelength. If you don't know how to begin, keep reading.

Transmitting and receiving as simply as possible

The only detail of Hertz's apparatus I found described was that he detected his waves by means of a loop of wire. The wire had such a large current and voltage induced into it by the radio waves that a visible spark jumped across a gap in the circle of wire. Wow! It must have been a big radio signal that would induce that much energy into a loop of wire. And if the signal was that big, how far away from the transmitter had the loop been? I suspect the signal strength had to be high and the loop had to be very close, like a foot or two away.

If I were a skeptic who already knew about Faraday's transformers, this demonstration would not convince me. How would I know that radio waves had propagated across the one foot distance to the loop? Maybe all I was seeing was a big magnetic field that reached from one coil to another.



Maybe Hertz's demonstrator was just a transformer?

A transformer is a *magnetic* device that works by transmitting a changing magnetic field from one coil to another. Coils (inductors), convert the energy of an electric current moving through a wire into magnetic field energy that hovers in a cloud-like region around the coil. If a second coil is close enough to the first coil to be inside the magnetic "cloud," then when the magnetic field is changed, an electric force, a voltage, will be generated in the second coil. Transformers and the above diagram were discussed in Chapter 2.

Inductors store magnetic energy in the space around them so long as current is passing through the coil. The energy will remain in space so long as the current keeps flowing in the same direction through the coil. But when the current stops flowing, the magnetic energy becomes “stranded” in space. The magnetic energy then returns to the coil and induces a voltage in that coil in a direction that would force the current to continue flowing. That is, the induced voltage will try to maintain the earlier magnetic field in the coil.

If the current is gone permanently, then the magnetic field will completely collapse and the energy will dissipate into the coil and any circuit connected to it. But if the first coil is open circuit and even high voltages cannot restore current flow, then the field will collapse into the second coil. If the induced voltage is able, it will cause current to flow in the second coil to maintain the field. In the case of Hertz’s loop detector, the induced voltage caused a big spark to jump across a gap where the resistor is located in the illustration above.

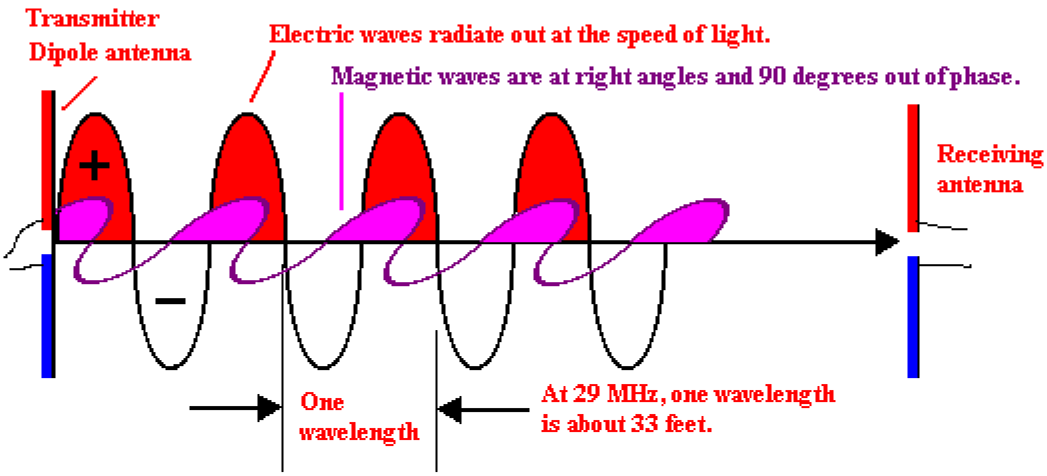
But I thought transformers were always made from iron.

You may be thinking that transformers are not air-filled flimsy coils but rather large iron things like the big steel cylinder on the power pole in the alley behind your house. Yes, those big iron things are transformers but they are designed for low frequency power lines. An AM radio frequency is on the order of a million Hz frequency, while the power company supplies current at 60 Hz. Therefore, instead of having air between the two coils, power transformers have iron. The magnetic field from the coil magnetizes the iron temporarily. Recruiting iron and turning it into a magnet increases the magnetic field a thousand times or more. With a huge magnetic field stored in the iron, the power transformer can transmit large amounts of energy with only 60 direction changes per second. A similar transformer without an iron core can transmit the same amount of energy, but it needs to repeat the magnetic field cycle perhaps a thousand times more often to transfer the same amount of total energy. In other words it must operate at a high frequency.

Why do you suppose the power company doesn’t use one million Hertz and do away with all that iron? After all, at one million Hertz the RF voltage would still be a burn hazard, but it could not electrocute anyone and would be considerably safer. Unfortunately, at one million Hz the power lines would act like antennas and radiate the energy into the sky instead of delivering it to your house. This is the same reason hams object to using the power lines to distribute high-speed Internet connections (BPL). Broadband Power Line broadcasts noise all over the shortwave bands and makes them nearly unusable.

How far should it transmit to demonstrate the existence of Hertzian waves?

To be sure that the waves are Hertzian and not just magnetic fields, I would be impressed by the demonstration if the detector (the receiver) were more than one wavelength away. A wavelength is the distance that a radio wave travels during the time it cycles from magnetic field, to electric field and back to magnetic field.



The speed of light is 186,000 miles per hour or 300,000,000 meters per second. (A meter is about 39 inches.) A wavelength is the distance a wave travels while going through one cycle of magnetic to electric energy conversion. The wavelength of the forty-meter amateur radio band (7 MHz) is obviously 40 meters. It turns out that typical, (total) antenna lengths for radio transmitters are either one half wavelength or one quarter wavelength. On 40 meters, a typical vertical pole antenna is one quarter wavelength or 10 meters (33 feet) tall. The 10 meter ham band extends from 28.0 MHz to 29.7 MHz.

What exactly is the frequency in Hertz of the 10 meter ham band? To convert wavelength to frequency, divide meters per second of light speed by the wavelength:

$$f = c / \lambda$$

Where f represents frequency, c represents the speed of light and λ represents wavelength

Speed of light/ wavelength = Frequency in Hz.

300,000,000 meters/ second / Ten meters = 30 Million Hz (30 MHz) frequency

Remember that the AM radio band extends from 550,000 Hz to 1.7 MHz. The old analog channel 2 television started at 54 MHz. So the 10 meter ham band is roughly halfway between AM Radio and TV.

Getting back to the Hertzian demonstration, if I wish to transmit one wavelength on 40 meters, my loop would have to be 132 feet away from my transmitter. Frankly, I don't think Hertz's loop detector will work at that range. And if it did, I would be arrested for using a transmitter that powerful. As a rule of thumb, the American FCC will not object to experiments like this if the radio waves don't go past 50 feet at easily detected signal strengths. On the other hand, maybe I could use a higher ham band like 10 meters. Now I only have to go 33 feet. If I go up to UHF frequencies, the wavelength could be a foot or less, but those frequencies would be hard to generate and harder to measure with 1884 technology. I have since been told that Hertz actually used 4 meters wavelength for his demonstrations.

Designing the 10 meter transmitter

For the sake of symmetry I used two identical LC circuits for my transmitter and receiver. To imitate what little I knew about Hertz's apparatus, I used one loop of wire about a foot in diameter. I knew from experience that seeing sparks on the receiver side was hopeless, so I also made the receiver an LC circuit that I knew would trap an oscillation from the transmitter.

At this point, you may want to depart from the year 1880. It all depends on how pure you wish to be playing the historical game. For a capacitor I used a modern 140 pF variable capacitor. That way, I could adjust the capacitance and tune the oscillation to a particular frequency. If I wanted to be a purist, it would not be hard to make a homemade capacitor made from sheets of metal with paper for insulation between the plates. Personally, I was confident that such a capacitor would work. I just didn't want to spend hours to make one.

My first problem was how to start the oscillation in the transmitter LC. In theory, by shorting a battery across the loop, it will charge the loop with a big current limited only by the internal resistance of the battery. Then when the battery is removed, the coil's magnetic field will discharge, forcing a voltage to appear across the capacitor. The loop will then be shorting out the capacitor and the oscillation will begin.

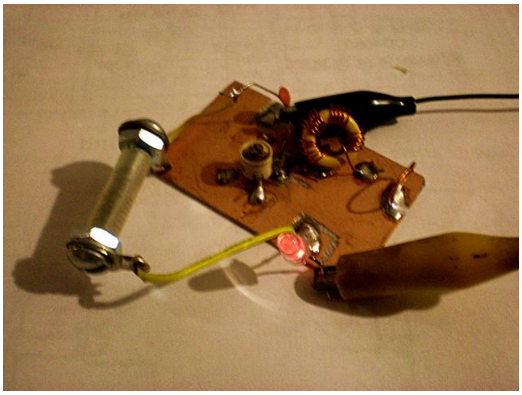
Like the variable capacitor, this project will go faster if you use modern tools to make sure your components are working. For example, to see if my transmitter was really transmitting, I used a ham band receiver tuned to ten meters. Sure enough, when I clicked the battery on the capacitor terminals, I could hear a click in the receiver loudspeaker. And when I tuned the capacitor, I could get the sound to reach a sharp maximum volume at a specific setting of the capacitor. Of course, if this were 1880, I would have to do everything by guess, trial and error. Those old guys were smart and persistent.

Designing the receiver

The next problem was how to detect when the receiver resonant loop was oscillating due to waves from the transmitter. What to use for a detector? The 1880 solution would be to use a detector called a ***coherer.*** Coherer detectors were developed for the wire telegraph. By the time a signal had propagated miles down a telegraph wire, the signal was often too weak to close a mechanical relay. Coherers were used to "amplify" a weak Morse code signal. They were described in Chapter 1, page 12. Unfortunately, coherers are a low frequency device, under 20 Hz. They are suitable for detecting a weak DC Morse code signal on a cross-country telegraph wire, but will not respond to modulation in the audible range. I doubted that they would be useful for tiny radio frequency signals. Being lazy, I didn't build one to find out. Besides, my crystal detector made out of local rocks certainly fit the 1880 criterion. Rocks had been discovered by 1880. I decided to build a 10 meter crystal set.

Are coherers adequate?

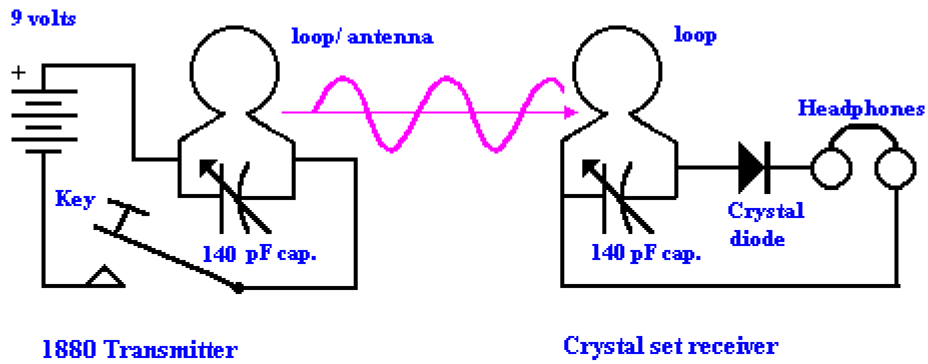
After I originally wrote this I met Ashish Derhgawen, N6ASD. Ashish built his own coherers out of a pinch of steel filings in a small plastic tube. Long machine screws were threaded into the ends of the tube and gently touch the powdered steel. He used this detector to turn on an LED from across the room. Of course the LED stayed lit until he bumped the coherer and shook the powder. Next he added an old fashioned electromagnetic door bell mechanism to shake the coherer and reset it automatically. The buzzing door bell assembly provided the "BFO function" so that the Morse code would be audible. Ashish posted a YouTube video of his spark gap transmitter sending Morse to his coherer receiver. It is receiving Morse code from 3 meters away.



The photo above shows the coherer I made following Ashish's method. It worked, but adjusting the screw pressure was extremely critical and it wasn't reliable. Here it is responding to an RF signal. The L-C resonant circuit is tuned to the RF frequency. I quickly learned that an LED also works quite well as an HF detector. Here, the coherer latched ON, stayed ON, and after that the LED simply rectified the RF. I'm pretty sure LEDs weren't available in 1880. My simple coherer experiment didn't begin to work nearly as well as Ashish's sophisticated communicator. But of course I only invested a fraction of the effort that he did. Patience and persistence are the key to success!

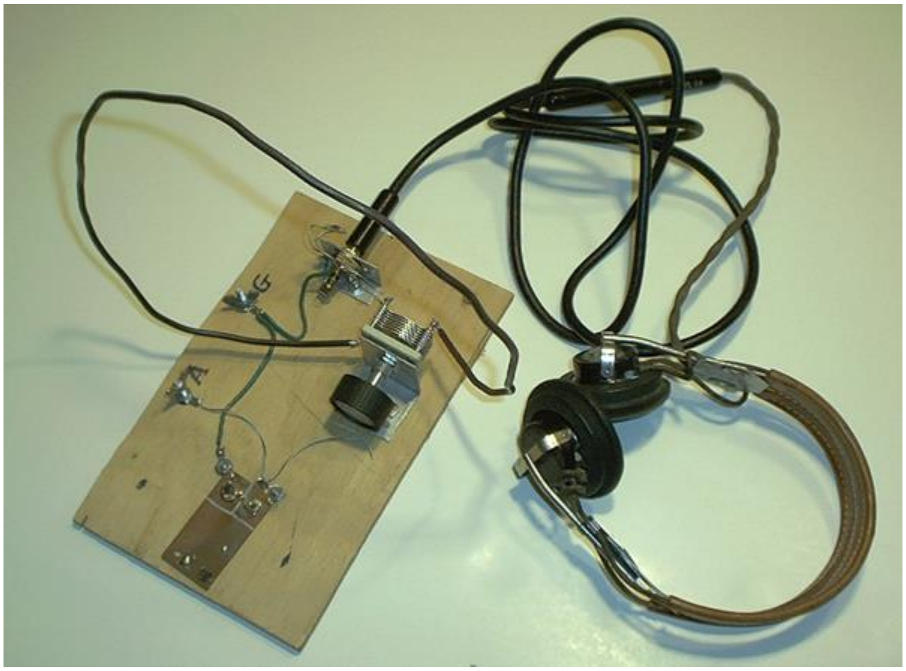
The headphones

Are headphones 1880 technology? Yes, just barely. Alexander Bell built his first telephone in 1879. He used a headphone designed like the homemade device described earlier. Actually, for my 10-meter receiver I used old commercial high impedance headphones instead of the homemade earphone. I couldn't afford to waste any sensitivity.



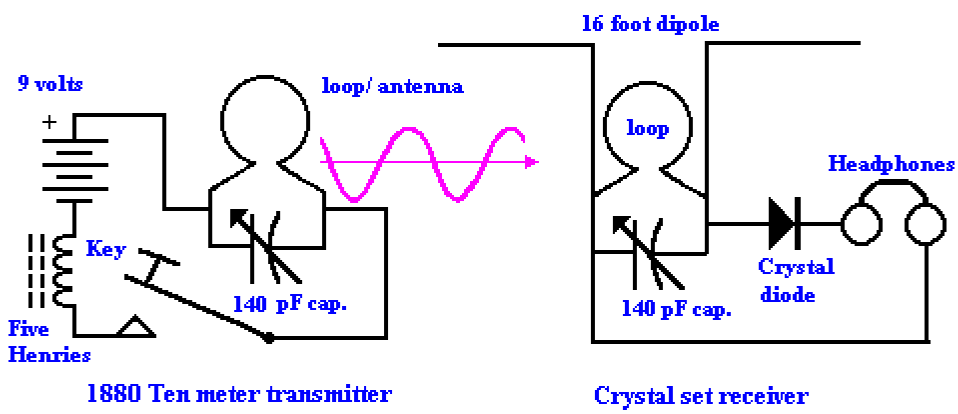
The simplest 10-meter communicator

I began with identical loop antenna/ inductors about 1 foot diameter for both receiver and transmitter. Instead of Hertz's spark gap, I put 140 pF variable capacitors across both loops to give me tunable LC circuits.



10-meter crystal set receiver

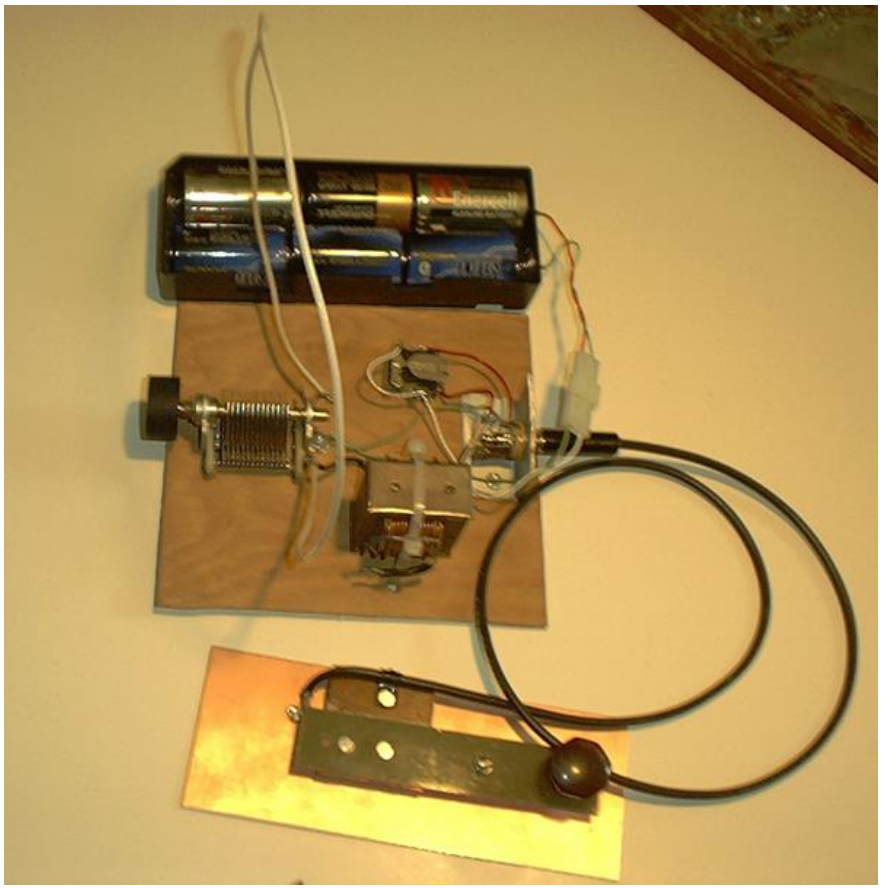
To charge the LC circuit, I used a crude ***spark gap***. I just touched battery terminals across the LC circuit while listening to the crystal set. I positioned the crystal set a foot away from the transmitter loop. It wouldn't be "radio" communication, but it would at least tell me if I was on the right track. I made sparks on the transmitter loop while tuning the capacitor. When tuned to just the right spot, I could suddenly hear obvious clicking in the headphones. I was surprised how sharp the tuning had to be. The big loop had relatively low inductance, so the capacitor had a tuning range of over 30 MHz. Tuning was probably not sharp by modern standards, but the adjustment was critical. In any case, I achieved a range of 12 inches from the transmitting loop. Progress! Well, it's much farther than the apparent range of a refrigerator magnet.



Improved spark gap transmitter & receiver

By adding a big iron core inductor in series with the battery, I got a much bigger, more sustained spark and a much louder signal in the crystal set. The inductor was the primary of an iron core filament transformer that I had in my junk box. The secondary of the transformer was left open circuit. Actually, I tried several transformer windings and inductors from my junk box until I found the one that gave me the biggest visible spark. The battery was six D-cell alkaline batteries in a plastic battery holder from Radio Shack. My telegraph key was made out of two pieces of printed circuit board separated from each other by a piece of wood. (See Chapter 9.)

Using an extra inductor in this way is analogous to the automobile ignition system described in Chapter 2. Without the big inductor, the only inductance to store energy was the single loop of wire. So when I tapped the battery wire on the LC circuit, very little energy was stored in the capacitor and inductance. The spark in a spark gap transmitter happens when the wire charging the inductor is broken, just like opening the breaker points in a car ignition. The more energy is stored in the system, the bigger the spark when the connection is opened.



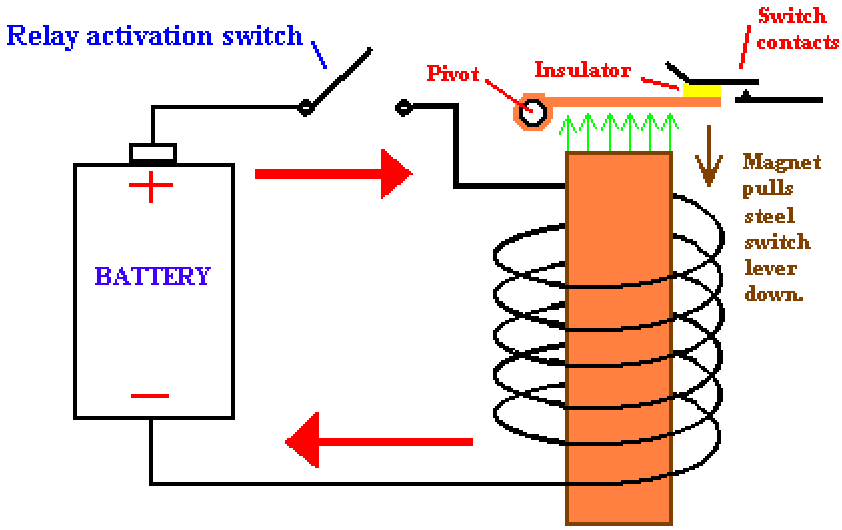
10 meter sparkgap transmitter

I enlisted my XYL to listen to the headphones while I moved the transmitter across the room. (Wives are known as “XYL” in Morse code. XYL stands for “former young lady.") Now that I had the inductor and a larger spark, I produced a range of 10 feet! I explained to Katie that she was doing the same job as Marconi’s assistant, Mignani. When Mignani heard the repeated “S” in Morse, he fired a rifle into the air. “Don't I get a rifle?” Katie asked.

A relay for automatic keying

Unfortunately, a wavelength at 29 MHz is 33 feet and I was still short. Hmmmm ... How to get the last 23 feet of range? First I added a relay to key the inductor. Yes, they had relays in

1880. The relay had nothing to do with extending the range, but it did enable me to continue doing experiments without having Mignani do the listening for me. You know, fire the rifle and all that.



RELAY CONSTRUCTION

A relay is an inductor/ electromagnet wound around an iron core. A hinged piece of steel is suspended by a spring near the iron core. When current passes through the electromagnet, the steel hinge is attracted to the iron with an audible "click." The hinge, in turn, mechanically closes a switch that can be entirely independent of the electromagnet circuit. In this way, one current can control an independent circuit. In my transmitter, the switch contacts on the relay became my "spark gap." Every time the switch opened, a big spark jumped across the switch contacts, not unlike the spark in an automobile ignition spark plug.

Using a relay, I could use my electronic telegraph key, a "bug" set on "dots," to key the transmitter automatically. This homemade bug is described in Chapter 9. If you don't have one of those, you'll have to have your "Mignani" key the transmitter for you. With the transmitter making a continuous buzzing signal, I could move the receiver around the house. The signal sounded just like the automobile ignition noise that you sometimes hear in your AM radio.

More range = bigger, higher antennas plus bigger batteries

I could easily increase the range of the transmitter by using more and bigger batteries and a larger series inductor. If I really wanted to extend the range, I could add a dipole designed for 10 meters and put it up in the air about 50 feet. In fact, this is exactly what the early guys did – they made bigger and bigger transmitters and antennas. However this was 2002, not 1880. The trouble with using a wavelength of 10 meters is that, if I were to increase the effectiveness of my transmitter, I might easily hear it with my crystal set 33 feet away. Unfortunately, someone else might also hear it in Australia. That would be bad since sparkgap transmitters have been banned since 1927.

The simplest improvement I could make to the receiver was to add a 16-foot dipole antenna. The dipole consisted simply of two eight-foot wires soldered onto the sides of the receiver tuning capacitor. The dipole was oriented at right angles to the direct path to the transmitter. That did it. Now I could plainly hear the signal from the basement to the other end of the 2nd floor of my house, over 50 feet. That was well beyond one wavelength range. Success!

Looking for standing waves

To measure wavelength, I set up a long wire transmitting "antenna" about 50 feet long across the floor and upstairs. I reduced the transmitter batteries from 9 volts down to 3 volts. Then I turned on the transmitter. I took the dipole off the receiver and then used the receiver loop as a "probe." Walking along the wire, I was able to hear peaks and dips in reception every 6 feet or so along the wire. What I was hearing was "standing waves." When the RF current reaches the end of an open wire, it bounces back along the wire. The returning waves cancel and reinforce the outgoing waves, which explained the peaks and nulls that I was hearing. A large number of peaks means the wire length is different than one wavelength and the standing waves are complicated. If the wire were exactly one wavelength, I would hear just two peaks – they would be the two positive and negative humps of a single sinewave cycle.

Next I cheated. Since I already knew the frequency was 29 MHz, I calculated what the wire length should be for one wavelength. I trimmed the wire to exactly that distance and tried again. As expected, there was a single pronounced dip in the center of the wire. The sinewave signal was reflecting back and forth from one end of the wire to the other, with a dip, the zero crossing, in the middle. When the reflections don't come out even, you get multiple dips and peaks.

Of course, knowing the answer before he started was not what Hertz experienced. He had to figure out all the details the hard way. Also, knowing the answer ahead of time biases the result. The exact alignment and distance of the receiver loop with respect to the wire were critical, so it can be argued that I was just hearing what I wanted to hear. Craftsmanship and scrupulous honesty are essential when doing science. My frequency measurement obviously needs more work.

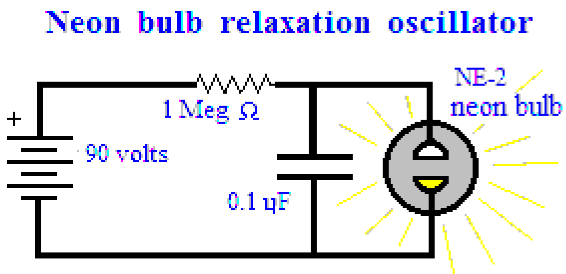
****

RESONANT SPARK GAP TRANSMITTERS

Ashish, N6ASD, sent me an article about "negative resistance" spark gap oscillators. Rather than generate a low frequency spark remote from the L-C frequency filter, the spark itself is used as a switching element. The sparks occur at the actual operating frequency with the L-C circuit controlling that frequency. It seemed to me that this design had the potential to be more efficient and narrow band when compared with generating low frequency sparks and extracting the desired high RF components.

Analogy with neon light relaxation oscillators

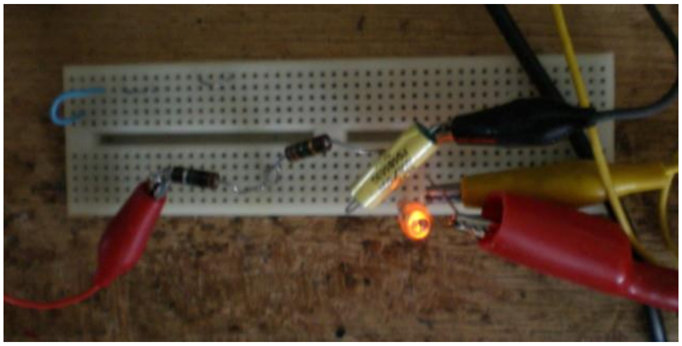
When I was a kid, one of my first electronic projects that I actually understood was a neon light relaxation oscillator. These consist of a DC voltage power supply and a series resistor-capacitor circuit with a neon gas light bulb wired across the capacitor. Back in ancient times appliance stores sold 90 volt batteries for use in portable vacuum tube radios.



Neon bulbs give off light when a certain ignition voltage is reached and the neon gas begins to ionize. In the neon oscillator, once the capacitor voltage charges up to the ignition threshold, current begins to flow and the bulb lights. The current flows through the bulb and discharges the capacitor. After losing about 30% of the voltage, the ionization ends. The current flow stops, the light goes out and the capacitor voltage is free to rise again.

What does "negative resistance" really mean?

For decades I have been hearing about "negative resistance" and how it can be used to make oscillators. There are many examples of so-called "negative resistors." These include neon light bulbs, magical quantum tunnel diodes, the low plate voltage characteristic of old-time tetrode vacuum tubes and spark gaps. "Negative resistance" is a terrible name for these devices. Most of them are voltage sensitive switches that are OFF until the voltage across them rises to a certain level. Then they turn ON and short out the voltage that is applied to them. With appropriately sized resistors and capacitors, a neon bulb can flash on and off at any slow rate you choose, such as once per second.



Because I was interested in making an RF generator, I wondered how high a frequency the neon bulb could handle. Others may have better luck, but the highest I could manage was 18 KHz. Above 18 KHz, the oscillation disappeared and the neon remained continuously lit.

Can resistance really be negative?

To me, negative resistance implies a "resistance with a minus sign in front." Ohm's law says Voltage = I (current) x Resistance. Therefore negative resistance would be:

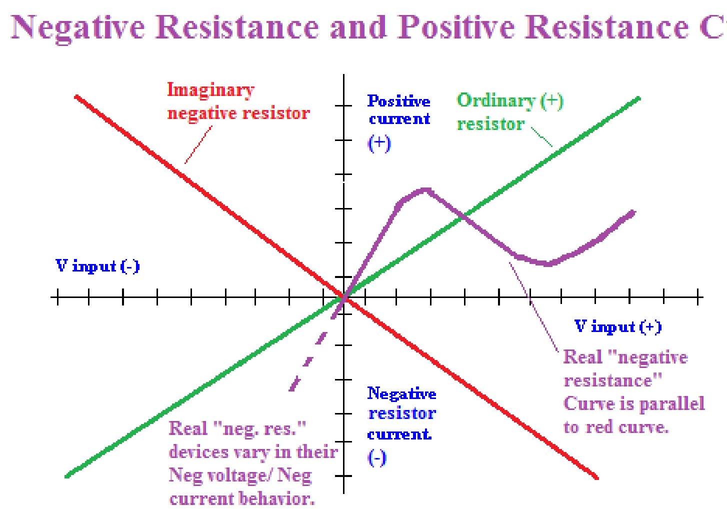
Voltage = I (current) x (**minus** Resistance)

For example, let V = **plus** 10 volts, R = **minus** 100 ohms

Therefore, I = **plus** 10 volts ÷ **minus** 100 ohms

I (current) = **minus** 100 milliamperes

In other words, ***negative resistance makes current flow backwards?*** It pushes current in a direction opposite to the driving voltage. Real negative resistors would resemble battery chargers ... which is why the term is misleading.



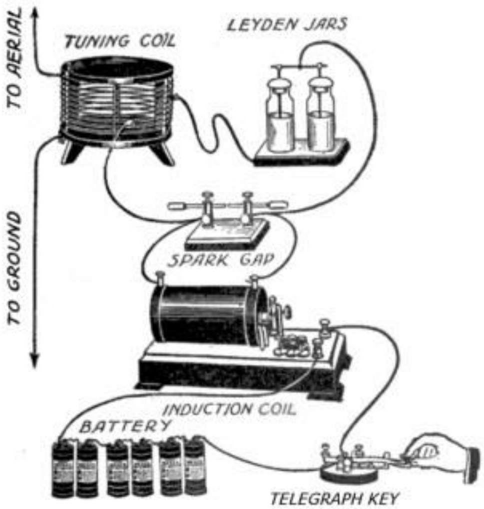
"Negative resistance" implies a magical passive electronic component that makes current flow backward. Real "Negative resistors" have a small percentage of their V/I curves in which rising voltage causes current to fall instead of increase.

The volt ampere characteristic of a simple resistor is shown above in green. When voltage goes up, the current through an ordinary resistor increases. The volt-ampere characteristic of an imaginary negative resistor is shown in red. The volt-ampere characteristic of a real "negative resistor" like a tunnel diode or tetrode tube is shown in purple.

The only characteristic that real negative resistors share with "minus resistance" is the ***SLOPE*** of the graph over some interval of voltage. At any given fixed voltage, a real negative resistor has a fixed current flow in the normal, positive direction. The negative feature of the negative resistor is that value of the resistance changes when the voltage changes. The negative resistance has a downward, negative slope instead of positive slope.

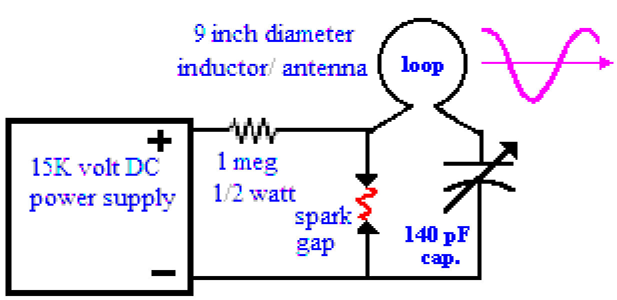
A resonant spark oscillator

Supposedly in the 1910-era, resonant spark-gap oscillators were more like tube or transistor oscillators. Because there was a separate spark for each sinewave cycle, presumably the frequency could be sharply tuned, like a continuous sinewave oscillator. Having experimented with simple spark gap transmitters, I wondered if I could make this concept work:



Ashish sent me this drawing of an old time spark gap transmitter. An induction coil, similar to an automotive spark coil, generates pulses of very high voltage, 8 or 10 kilovolts. The secondary of this transformer has a ***series*** L-C tuned circuit across it. The drawing is a bit confusing because the upper coil is also a transformer, probably an autotransformer. The spark is generated by an oscillating series L-C circuit and not the operator's key.

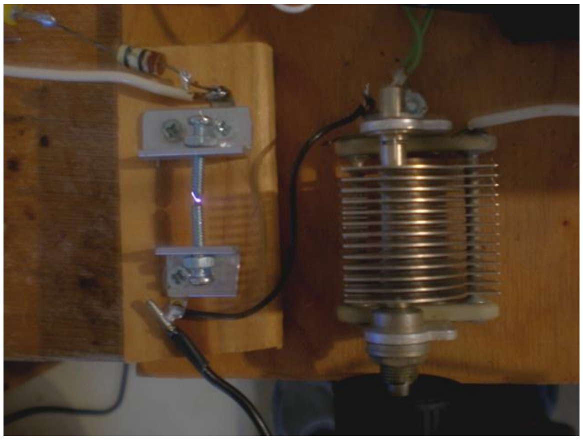
In my original spark gap transmitter the current pulses were produced manually. I have read that practical spark gap transmitters, like the one on the Titanic, used motorized commutator-like spark gaps to generate a continuous buzz. In this way dots and dashes were continuous like the CW we hear today. Without the motor, an AM crystal set receiver only hears clicking noises. A BFO doesn't help, because there is no continuous carrier. Notice that the Titanic did *NOT* use a resonant spark system. To make my first spark gap continuous, I put a weight on the paddle of an electronic keyer. That way I didn't need someone to press the key constantly.

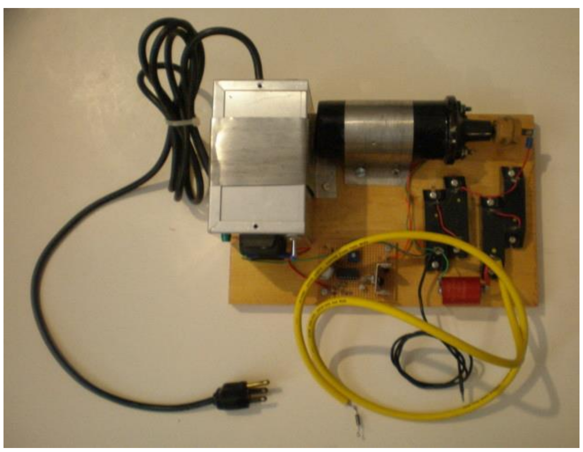


10 meter resonant spark oscillator

Spark frequency = resonant frequency

The above circuit is my version of the drawing Ashish sent me. It is powered by a continuous, very high DC voltage, 15,000 volts. It works! There is a continuous blueish-pink spark in the gap. Listening to my old Collins HF radio, I can hear a strong buzzing signal wherever I tune the L-C circuit. The signal is extremely strong at the exact resonant frequency, but the buzz can be heard over the entire 10 meter band. My receiver has no S-meter, but the interference on 12 and 15 meters is many decibels down. I turned it on next to my VHF receiver while I was listening to the local 2 meter ham club roundtable and there was no discernable interference.



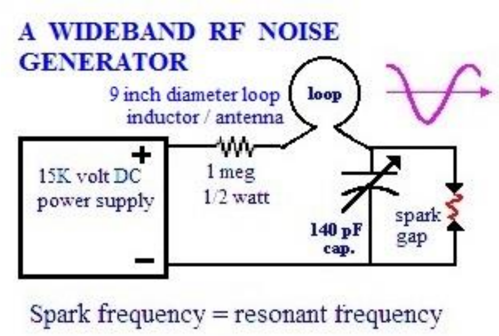


The 15,000 volt DC supply is shown above. It consists of a 12 volt, 10 watt power supply powering a MOSFET switcher. The MOSFET transistor pulses a Buick automobile ignition spark coil at about 11 KHz. The pulses from the spark coil are rectified and doubled by the 4 large, black, high-voltage diodes. High voltage, 15 KV or more, is stored in a very high voltage mica capacitor (the red cylinder). I salvaged the capacitor from an old cathode ray tube TV set. I originally tried to use a 350 volt DC power supply for this circuit. Unfortunately, the spark gap width had to be more precise than I could build and may not even be practical. I don't know any legitimate ham radio applications for a 15,000 volt DC supply. However, the circuit illustrates some useful power supply design concepts so I added it to the end of the power supply chapter, Chapter 8.

An RF noise generator

In my first attempt at a resonant spark gap I put the spark gap across a ***parallel*** L-C circuit, much like the L-C used in my original spark gap. This produced few sparks and little signal. I wasn't surprised, considering the low resistance and impedance of a simple wire loop. Maybe I should have tried putting a relatively large high voltage capacitor, like 500 pF, in series with the parallel L-C. That would prevent the DC voltage from being directly across the simple wire loop.

Next I tried to use a circuit analogous to an R-C-neon light bulb relaxation oscillator. In the neon flasher the resistor is in series with the capacitor being charged. When the capacitor reaches the neon ignition voltage, it discharges into the neon light. Putting the spark gap across the capacitor produces strong buzzing across the entire upper end of the HF spectrum.



Listening on my old Collins radio, this noise generator pretty much wipes out the 15, 12 and 10 meter bands. By tuning the variable capacitor I can concentrate the cacophony on "just one" of those bands. This is definitely an invention we don't need! In summary, my primitive experiments suggest that resonant spark generators don't narrow the signal bandwidth, no matter how they are wired. I prefer transistor oscillators, thank you.

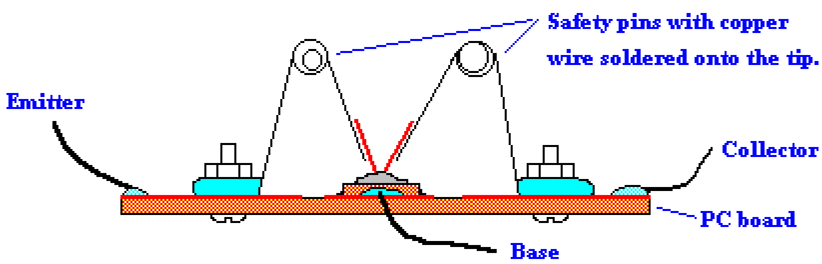
****

HOMEBREW TRANSISTORS

Here is another project you may enjoy. It will also introduce you to the basic principles of bipolar transistors. Schockley and Bardeen first invented bipolar transistors while working for the Bell Laboratories in 1947. Actually, I've read that the basic principles for field effect transistors were described in German patents in the 1930s. However, field effect transistors (FETs) were not developed into useful components until the late 1960s. We shall first use an FET for a ham project in Chapter 6, so FETs are discussed there.

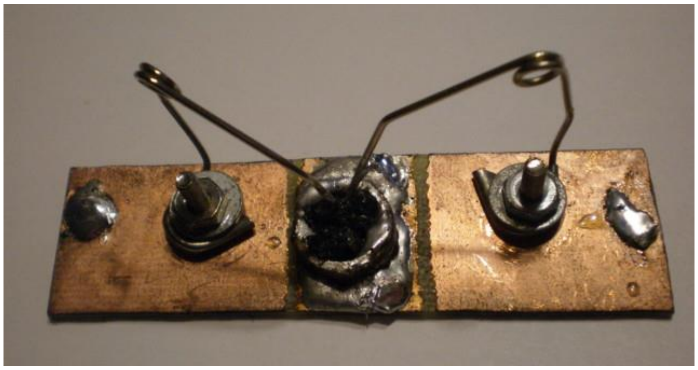
The purpose of transistors is to "amplify" small signals. Saying it another way, transistors control big currents or big voltages using tiny control signals. Transistors can amplify a tiny signal that is audible only with sensitive headphones. They can make it strong enough to run a loudspeaker or even deafen a crowd in a stadium during a rock concert. Alternatively, a transistor can be used to allow a little control signal to turn on a huge current and voltage. For example, an engineer at a power plant might push a keypad on a computer with a fraction of a milliampere of current flowing through the switch. This action is amplified and results in megawatts of power at hundreds of thousands of volts flowing toward a city.

The homebrew transistor



A point contact transistor

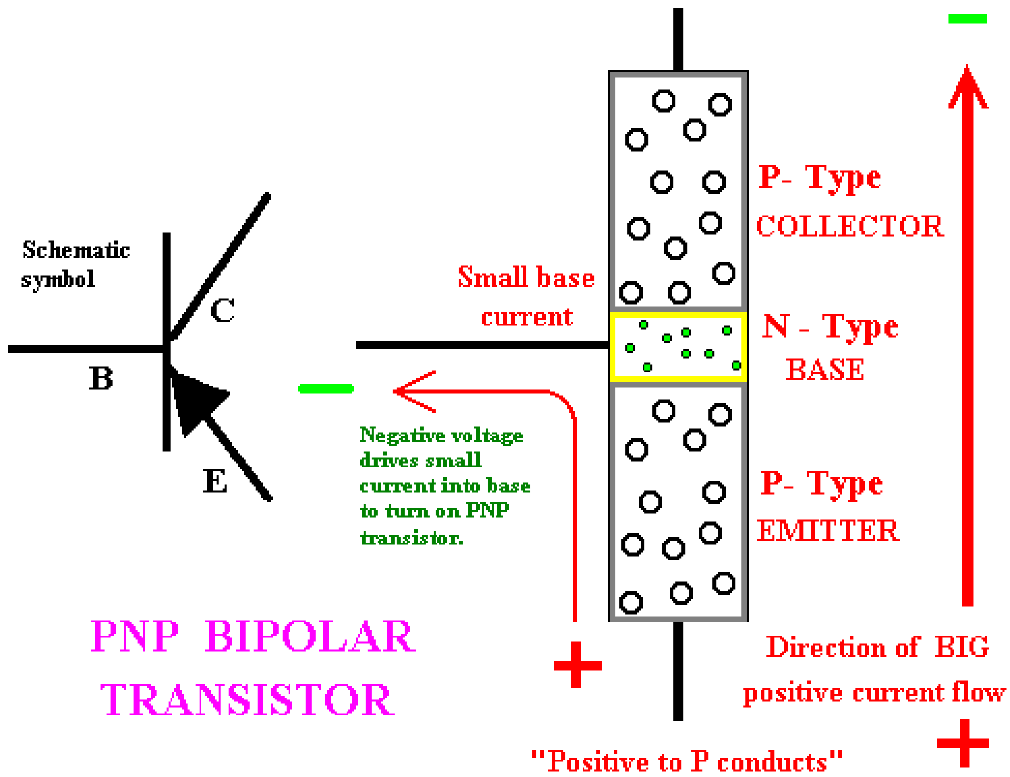
The first bipolar transistors were the "point contact" type. They were much like the galena diode described earlier. After my diodes worked so well, I wondered if I could make a transistor. I also recall that the early point contact transistors were quite fragile - not surprising!



Construction of a PNP bipolar transistor

A ***bipolar*** junction transistor consists of two diodes made on the same semiconductor crystal. In the drawing shown above, I am attempting to use galena as the N-type semiconductor. The diodes are connected back-to-back so that it will look like an open circuit from the terminals called "emitter" and "collector." This is supposed to be a P-N-P transistor. The two copper points are supposed to touch the galena so close together that the tiny N-type semiconductor region between the two points can be biased by the base current. The bias current is supposed to electrically convert the semiconductor region into a "conductor" and thus turn the two back biased diodes "ON." Don't forget to cut notches in the copper sheet to separate the three terminals.

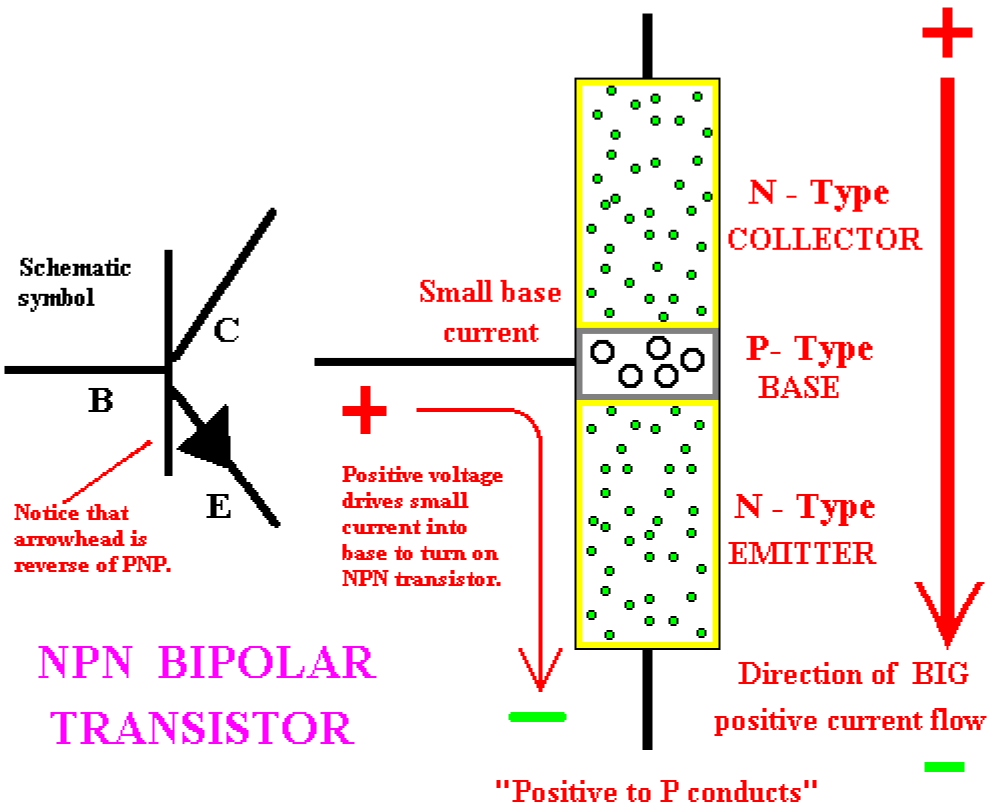
Below is a diagram of a modern commercial PNP transistor. Manufactured transistors have all three components made from semiconductors - no metal points delicately probing the N-type semiconductor. Since the P-type semiconductor components can be quite large, they can tolerate high currents and heating.



In a point contact transistor the metal from the copper points diffuses into the surface of the crystal and makes a tiny region of P-type semiconductor where the copper touches the semiconductor. Unfortunately the "emitter" and "collector" are identical and aren't optimized for their different roles as they are in commercial bipolar transistors. I also tried making a copper/steel/copper transistor, but without any measurable breakdown voltage, it just acted like a short circuit.

NPN transistors

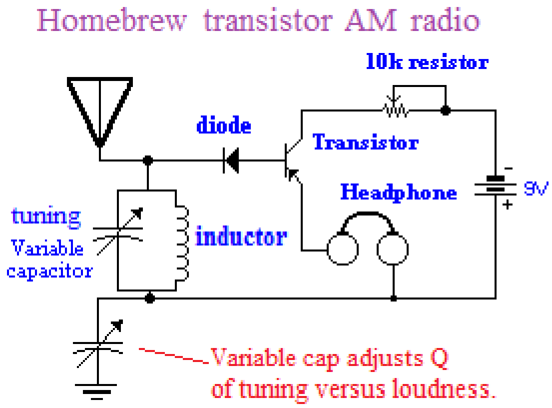
A nifty advantage of bipolar transistors is that they can be built two ways. By reversing the P-type and N-type semiconductors, an NPN transistor can be built that operates exactly like a PNP transistor, except all the polarities and current directions are reversed. The advantage of having two polarities is that circuits can often be simplified by using both kinds in the same circuit. In practice, NPN transistors are usually slightly more robust and less likely to fail at high power loads. For this reason, the power amplifier stages in modern transmitters are almost always NPN and N-channel devices. On the other hand making an NPN transistor out of crude crystals and safety pins is inherently difficult!



Proving the homebuilt PNP transistor has gain

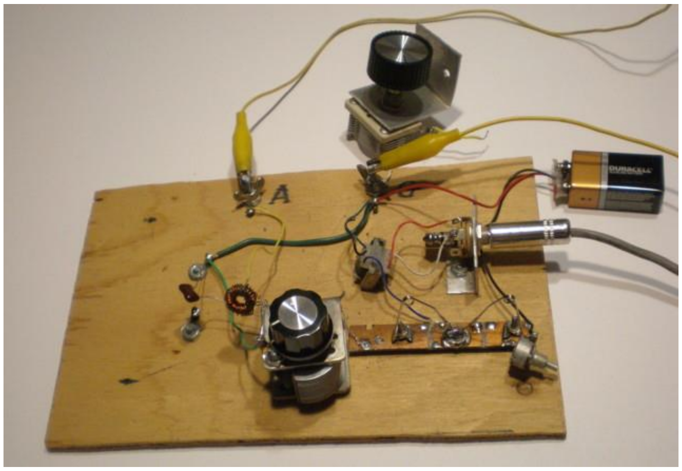
I measured the static volt/ ampere characteristics of my galena, point-contact transistor but couldn't show any gain with static DC currents even in the microampere range. Before I gave up, I thought I would try it as an amplifier in the crystal set. Maybe I could demonstrate gain in the subtle world of RF detection.

As we now know, these crude PN junctions usually fail at just one or two volts, so attempting to build an amplifier with voltage gain is not promising. I decided that my "transistor" was most likely to work as an ***emitter follower.*** In an emitter follower there is no voltage gain, only current amplification. The load, the headphones, would be located between the emitter and the positive side of the battery. That would match the impedance between the high impedance detector and my low impedance (8 ohm) commercial headphones. Because galena diodes break down with typically 1 volt of reverse bias, I used a 9 volt battery with a 10K pot in series so I could limit the voltage on the collector to less than 1 volt.



Rock radio schematic

As I moved the emitter pin around on the crystal, a loud radio station abruptly appeared in the headphones. I disconnected the battery. Sure enough, the music was <u>much louder</u> with the battery connected.



An accidental crystal microphone

With the battery in place, but the signal diode disconnected, I heard a faint "sea shell" roaring sound - you know, like a live microphone. Tapping on the transistor assembly, I heard the scratching sound greatly amplified in the headphones. I seemed to have built a "crystal microphone!" I replaced the homebrew transistor with a real PNP transistor, a 2N3906. It amplified just as well but had no microphone-like characteristic.

Commercial crystal microphones consist of a lump of Rochelle salt held between two electrodes. When exposed to sound, mechanical vibrations cause the salt to generate a tiny audio frequency electric voltage. Crystals like this are called ***piezoelectric crystals*** and perhaps I had just built one. Piezoelectric quartz crystals are used to regulate frequency and are discussed in Chapter 6.

Repairable transistors

While listening to a radio station, I slowly increased the battery voltage by lowering the resistance of the 10K pot. As the collector-to-emitter DC voltage rose higher, the volume increased higher and higher. I monitored the average collector to emitter DC voltage with a high impedance voltmeter. Then suddenly the voltage and sound crashed. I lowered the voltage again, but the sound didn't return. Good grief! I blew my transistor! No sweat. I just scratched the collector pin around on the crystal until I found a new "sweet spot" and I was back in business. Repairable transistors! Now there's a concept. After several trials I found sweet spots as high as 5 volts before the transistor died.

Now that I had an amplifier on my crystal set, I replaced the commercial headphone with the homebuilt "Caribou headphone" which I described earlier. You may remember that this headphone was made from a piece of magnetite ore and had a tin can lid for a diaphragm. Success! The sound was loud enough to understand actual words, rather than just distant music. More progress! I tried replacing the sulfide ore crystal with a carborundum crystal. It worked as well as the ore, but I had more difficulty finding active contact points for the little copper probes.

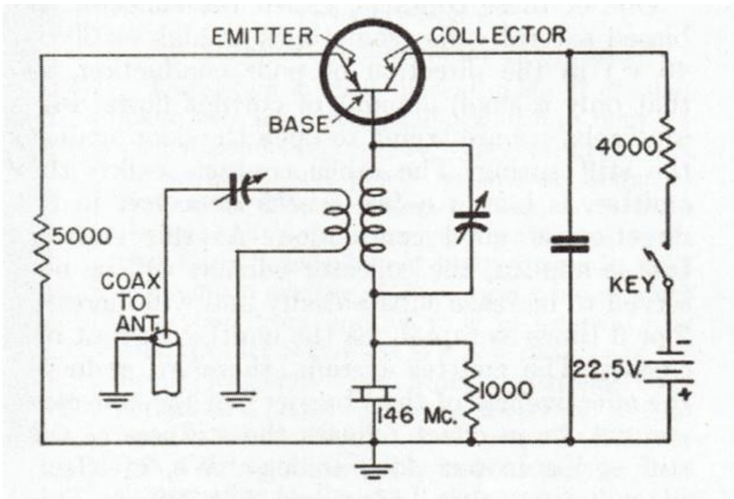
Is this gain or what?

In fairness, most of the "gain" or amplification I observed was just DC bias to the headphone which helps overcome the hysteresis of the steel components in the headphone. I replaced the transistor with a variable resistor so that the battery's sole function was to bias the headphone. That produced extra sound, especially for the homemade headphone that has such a weak magnet. After switching back and forth between resistor and transistor, the transistor was clearly louder, about 5 dB (decibels) louder measured with my multi-meter.

As I fiddled with the contact points, I eventually ruined the crystal. Under the microscope I could see copper smeared on the surface of the galena - all the sweet spots were spoiled. When I built replacements, they never worked as well. I spent another morning trying to conclusively demonstrate gain by building an RF oscillator. I never got so much as a peep out of it, although as you would expect, a real 2N3906 PNP transistor worked just fine.

The first transistorized ham transmitter

A 1953 QST article described the first transistorized CW ham transmitter. It operated on 2 Meters! *(see QST March 2015, p.50)* It used a prototype point contact transistor built in the RCA laboratory by George Rose, K2AH. The semiconductor crystal was germanium. The two contact points were phosphor-bronze needles positioned less than a millimeter apart. Apparently the germanium was assumed to be N-type semiconductor. It was packaged in a tiny 3 lead package, not much different from modern leaded transistors. The resistance in series with the 22.5 volt battery totaled 9K ohm, not including the resistance of the transistor. Average transistor voltage during operation was 10 volts. It was wired as an emitter follower circuit and has an eerie similarity to my crude experiments. It oscillated at 146 MHz with a 30 milliWatt output. The output was sampled from the base LC circuit. The crystal operated on the 9th overtone of a 16 MHz primary frequency crystal that had been ground to be a 5th overtone crystal. George fed it into a 12 element Yagi beam antenna and sent a message 25 miles, RST 559. Wow! I'm impressed. Couldn't they have at least started with 40 meters?



I reproduced this circuit with a homemade transistor, a 40 meter quartz crystal and the appropriate L-C tuning. So far, I have not been able to make it oscillate. Darn!

Commercial bipolar transistors

Above is an assortment of bipolar transistors. The little black ones on the lower left include 2N3906 and 2N3904 devices used for low power oscillators and amplifiers. The medium sized transistors are used for amplifiers in the range of ¼ to 5 watts. Notice that two of them are wearing black aluminum heat sinks. The power handling capacity of a transistor or other component can be increased by getting rid of the waste heat more efficiently. The large power transistors in the back row are designed for power supplies and high power amplifiers. A mounting kit for case style TO-204 (formerly TO-3) transistors at the upper right consists of a mica insulator, silicon grease and a socket with screws. The insulator and screws mount the transistor on large heat sinks. Heat sinks for high power transistors usually have flat, thick aluminum plates to receive the transistors, while arrays of molded cooling fins on the opposite side dissipate the excess heat into the air.

The beautiful gold device is an obsolete high power germanium transistor. These transistors use germanium instead of silicon for the basic semiconductor crystal. They have the advantage that when turned on (saturated) there is less voltage across them than comparable silicon bipolar transistors. So back in the 1960s and 1970s, they were preferred for high current, low frequency switching circuits. Unfortunately, they are much more easily damaged by heat. Today, modern MOSFET silicon transistors achieve better performance without this disadvantage, so germanium has gone the way of the crystal set. In theory, it would also be possible to build transistors made from graphite, another semiconductor crystal. However, these devices would be way too heat sensitive to be useful.

An audio amplified vacuum tube toy receiver

I started Chapter 4A by describing a toy crystal set from about 1940. The toy receiver kit pictured below also belonged to my brother-in-law, Peter Hill. It appears to be from the late 1940s. It uses low voltage, portable radio, vacuum tubes to detect and amplify the AM signals. This receiver is a comparable in principle to my crude transistorized crystal set. I'll bet it worked much better than mine and it couldn't help be more reliable.

It is powered by carbon-zinc batteries. Of course, the large "Eveready" 45 volt portable radio battery is stone dead. Batteries like this haven't been manufactured for many decades. However, the 1.5 volt filament battery could be replaced with a modern 1.5 volt alkaline D-cell. When we want to build something, there is usually a way. Perhaps someday I'll retrieve Peter's old toy from the attic and get it running again. For now, Chapters 7 and 14 describe homebrew ham radio receivers that attempt to achieve reliability and (relative) simplicity.

In conclusion

Yes, Virginia, there are homemade transistors. But science that isn't reproducible isn't science. Without better basement technology, my homemade transistors have no future except perhaps as microphones. Oh, well. The reward is the journey, rather than the destination. Keep reading, thinking and dreaming!

****