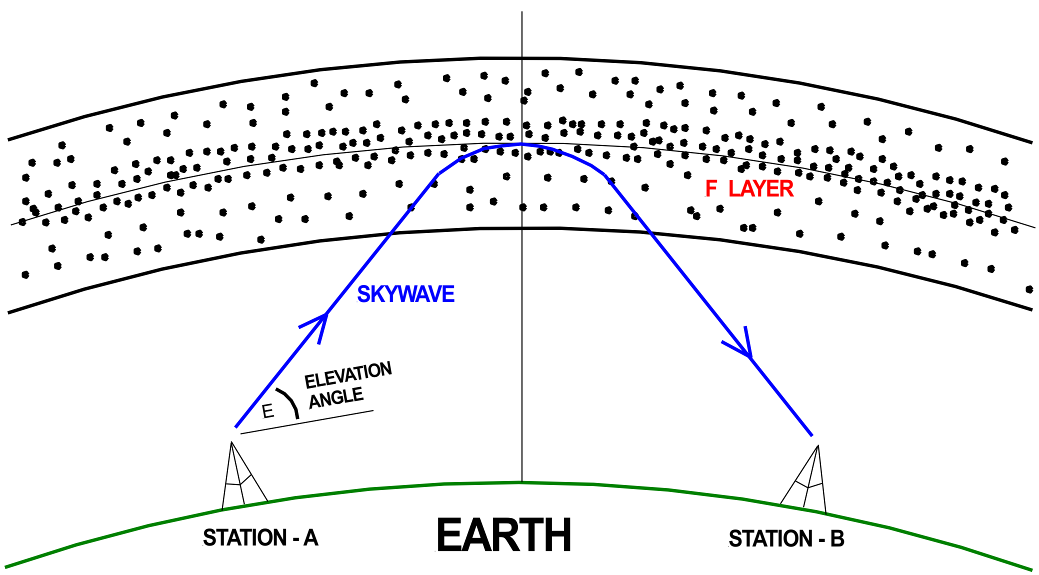
Skywave Radio Communications

Chapter 1

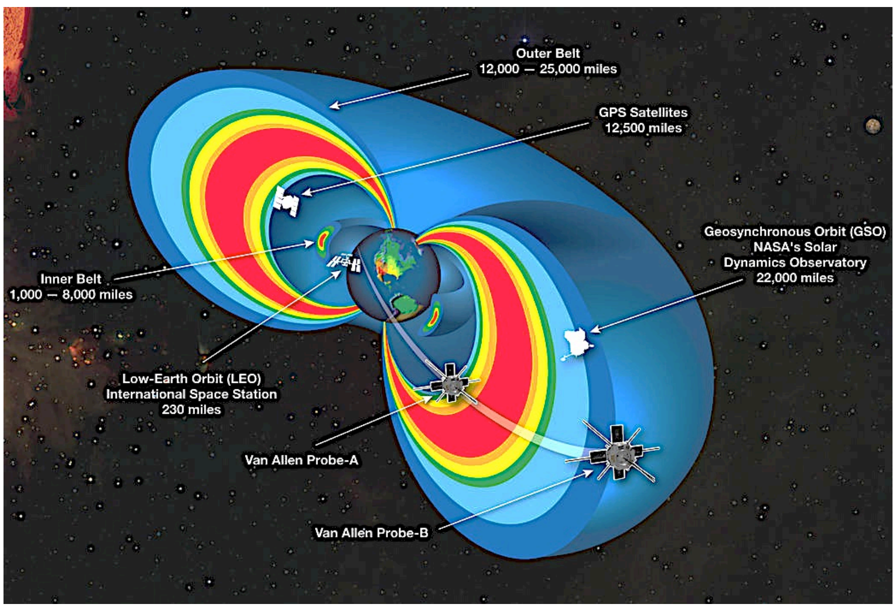
Introduction



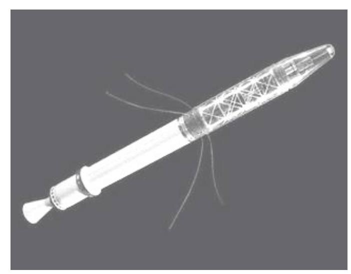
www.skywave-radio.org

1 Introduction

The discovery and ultimate understanding of radio communications via the ionosphere, better known as skywave communications, has been a long arduous process spanning over 100 years. Much of our current knowledge had to wait for the space age. Prior to that our understanding of skywave communications was based entirely on observations from the ground. Unfortunately, Earth's atmosphere, including the ionosphere itself, hide from view many of the phenomenon directly affecting ionospheric communications. Earth satellites and spacecraft probing the solar system have enabled us to go up above the atmosphere and take a look at what is going on. In fact, the very first U.S. Earth satellite, Explorer I launched in January 1958, discovered the Van Allen radiation belts (Figure 1) critical to our understanding of skywave communications.



Radiation Belts

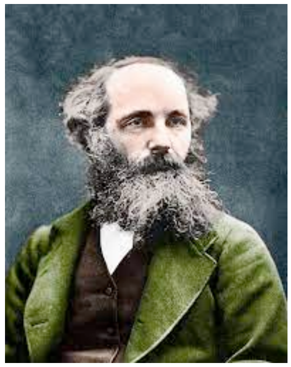


Explorer I Satellite

Figure 1 Van Allen radiation belts (source: NASA)

1.1 James Clerk Maxwell (1831 – 1879)

For centuries magnetism and electricity were believed to be completely different phenomena. In the early to mid 1800s experiments performed by Oersted, Faraday, Henry and others seemed to indicate that magnetism and electricity were somehow related. In 1865 Scottish physicist and mathematician James Maxwell, Figure 2, published a set of 20 equations that unified electricity and magnetism into a single theory of electromagnetics. Later Oliver Heaviside simplified Maxwell's equations into a set of 4 equations that we use today.



James Clerk Maxwell Oliver Heaviside

Figure 2 Maxwell and Heaviside (source: Wikipedia)

Maxwell is regarded as one of the greatest theoretical physicists of the 19^{th} century. Maxwell died young at the age of 48. However, in this short life, he formulated the unifying theory of electromagnetics, developed the kinetic theory of gases, made significant contributions to our understanding of color vision and. studied the nature of Saturn's rings.

In developing his set of equations, Maxwell integrated together laws of electricity, magnetism, and induction originally developed by Gauss, Faraday, and Ampere.

Gauss's Law:

$$\oint \vec{E} \cdot \vec{dA} = \frac{Q}{\varepsilon}$$

Gauss's Law of Magnetism:

$$\oint \vec{B} \cdot \vec{dA} = 0$$

Faraday's Law of Induction:

$$\oint \vec{E} \cdot \vec{dl} = -\frac{d\phi_B}{dt}$$

Ampere and Maxwell's Law

$$\oint \vec{B} \cdot \vec{dl} = \mu_0 I + \varepsilon_0 \mu_0 \frac{\partial \phi_E}{dt}$$

Maxwell's equations showed that it was possible for time varying electric and magnetic fields to be self-sustaining, each inducing the other. That is, a time varying magnetic field could induce a time varying electric field which in turn induced the magnetic field. Based on this result, Maxwell predicted the existence of electromagnetic waves, each wave consisting of an alternating electric wave perpendicular to an inseparable magnetic wave, and both orthogonal to the direction of wave travel as illustrated in Figure 3. Since each wave induced the other, the electromagnetic wave itself could theoretically propagate forever in its direction of travel through empty space.

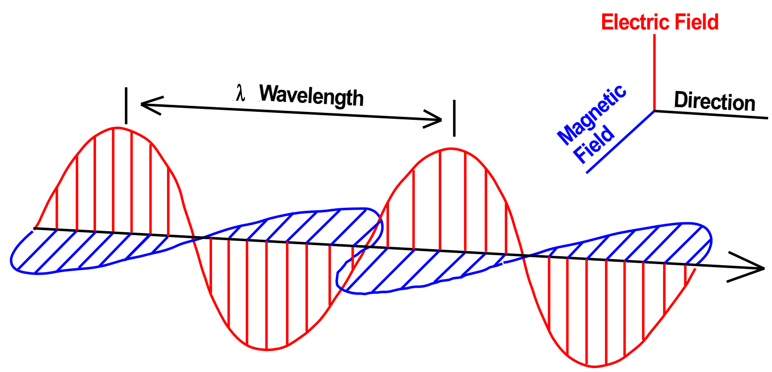


Figure 3 Electromagnetic Wave (source: author)

Maxwell noticed that the inverse square root of electric permittivity in free space, ε_0, and free space magnetic permeability, μ_0, had the units of velocity. Maxwell concluded that the velocity, v, of his proposed electromagnetic wave must be

$$v = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$$

Coulomb and Ampere had previously determined the values of ε_0 and μ_0 through careful static electric and magnetic measurements. Using their measurements Maxwell calculated the speed of an electromagnetic wave to be 284,000,000 meters per second, a value very close to the speed of light known at the time. Maxwell concluded that light itself must be an electromagnetic wave.

1.2 Heinrich Hertz

20 years later in 1888 physicist Heinrich Hertz tested Maxwell's equations to see if electromagnetic waves really existed. To verify Maxwell's equations Hertz had to demonstrate three things:

1. A rapidly changing electrical current, such as that produced by an electric spark, would generate an electromagnetic wave of the type predicted by Maxwell,

2. The electromagnetic wave would travel through space from one place to another, and

3. That it would do so at a finite speed near the speed of light.

Hertz used the apparatus shown in Figure 4 to perform these experiments. The apparatus worked as follows. A relay device, called an interrupter, placed in series with the battery "chopped" the electrical current flowing from the battery through the primary winding of the induction coil into electrical pulses. The pulses produced voltage spikes on the order of several thousand volts in the induction coil secondary winding. A capacitor consisting of two metal sheet was connected across the induction coil secondary as shown in Figure 4. The capacitor was charged to a very high voltage by the secondary winding voltage spikes. When charged to a high enough voltage, air molecules in the gap between two closely spaced spherical balls were ionized. The ionization allowed an electrical current to flow across the gap from one ball to the other producing a high voltage spark. The voltage spark discharged the capacitor. The process of repeatedly charging and discharging the capacitor continued as long as the switch in the battery circuit was closed. The capacitance of the two metal sheets plus the inductance inherent in the secondary circuit wiring formed a series resonant circuit which in the experiments conducted by Hertz was around 150 to 200 MHz. This circuit satisfied the first of the three conditions needed to verify Maxwell's equations.

The second condition required verification that the electromagnetic wave presumably produced by the spark gap circuit did indeed travel across Hertz's laboratory. His laboratory was a large lecture hall. To verify this condition Hertz constructed a wire loop with a very small air gap as shown in Figure 4. Hertz held the loop at the opposite end of the lecture hall from his radiating spark gap transmitter. When he did so, small sparks developed across the air gap in the receiving loop verifying that electromagnetic waves were in fact traveling across the laboratory.

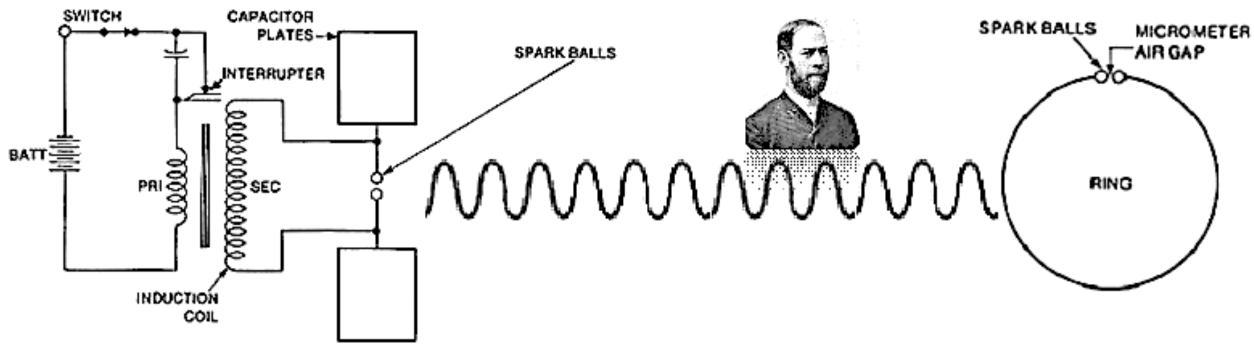


Figure 4 Experimental apparatus used by Hertz. (source: Wikipedia)

The third step, verifying that the electromagnetic waves were traveling at the speed of light was more difficult to prove. To do this, Hertz placed a large metal sheet on the wall at the far end of the lecture hall from his transmitter. The hope was that electromagnetic waves reflected from the metal sheet would interfere with the waves radiating from the transmitter forming standing wave patterns exactly the same as those obtained in reflecting light. To detect the standing waves, Hertz walked toward his transmitter holding his receiving loop and observing the strength of the sparks occurring across the receiving loop's air gap. As he did so he observed strong sparks across the air gap at certain distances from his transmitter, indicating standing wave voltage maximums. At other locations there were hardly any sparks at all signifying standing wave minimums. The distance from one peak to the next represented the wavelength λ of the standing wave. Using this result, and estimating (guess at) his transmitting frequency f, Hertz was able to calculate the velocity v of his electromagnetic wave according to

$$v = \lambda \cdot f$$

which he found to be nearly equal to the speed of light c. While this experiment was admittedly crude, it verified that electromagnetic waves did travel through space at a finite speed which was on the order of the speed of light.

Other scientists at the time were also experimenting with Maxwell's equations at lower frequencies without much luck. What caused Hertz to succeed was that he was operating, intentionally or otherwise, at relatively high frequencies permitting standing waves with two or more maximums to develop within the confines of his laboratory.

Hertz performed his experiments, including development of his test apparatus, over a considerable period of time. It was not a quick couple day project! In the end Hertz verified that Maxwell's equations predicting the existence of propagating electromagnetic waves was indeed correct.

Hertz was a physicist focused on the theoretical aspects of Maxwell's equations. It is unlikely that the commercial ramifications of his electromagnetic experiments ever occurred to him. With his experiments complete, he went on to other things. Hertz died in 1894 of blood poisoning at the early age of 36.

1.3 Guglielmo Marconi

In 1894, when he was 20 years old, Marconi read an obituary on Heinrich Hertz describing the work that Hertz had done. Marconi realized the commercial ramifications of the electromagnetic experiments performed by Hertz and began repeating the experiments at his father's estate (Villa Griffone) near Bologna, Italy. Like Hertz, Marconi initially experimented at approximately 2 meters. At these high frequencies he was unknowingly limiting himself to line of sight distances.

Marconi's spark gap transmitter (Figure 5) was pretty much the same as that used by Hertz.

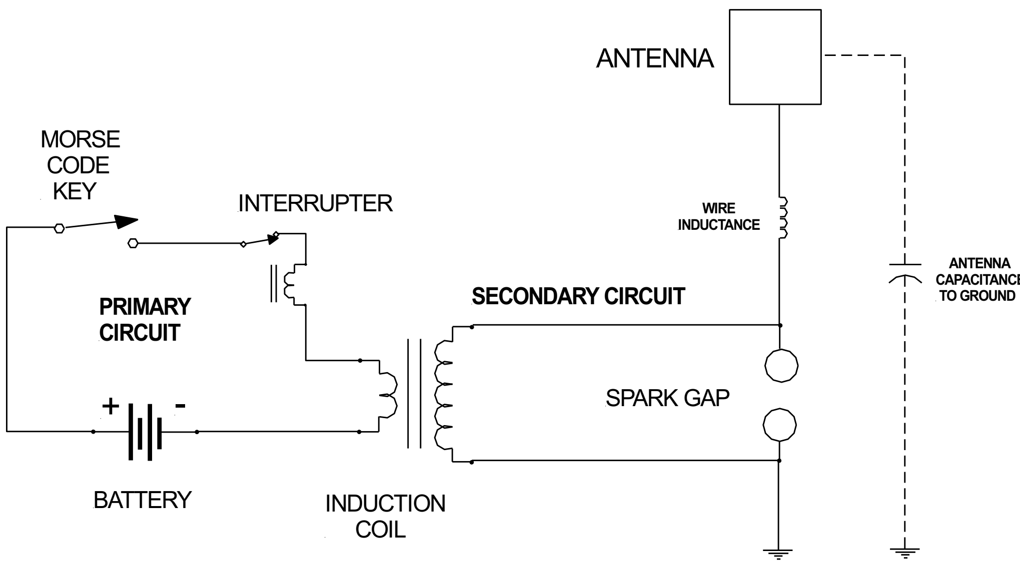


Figure 5 Marconi's first spark gap transmitter (source: author & Wikipedia)

While his spark gap transmitter was similar to that used by Hertz, Marconi's spark gap receiver, shown in Figure 6, was vastly different than the simple receiving loop used by Hertz.

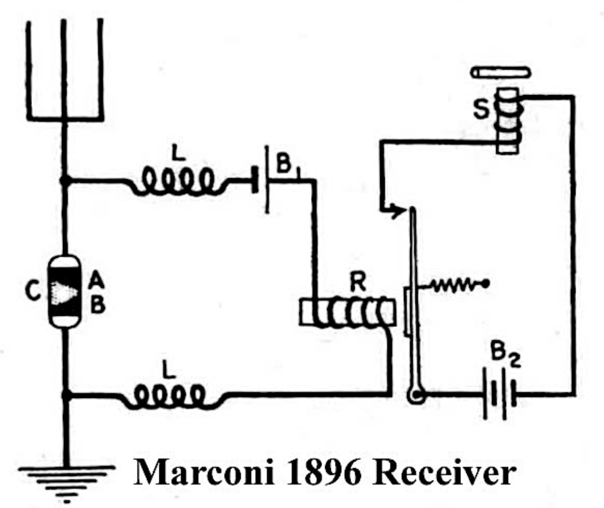


Figure 6 Marconi's 1896 spark gap receiver (source: Wikipedia)

A practical receiver required a detector that could activate a relay when an electromagnetic wave was detected. Such a device, called a coherer, consisted of a tube containing metal filing filling the gap between two electrodes as illustrated in Figure 7. In the presence of an electromagnetic wave the metal filings became fused together allowing an electrical current to flow through the coherer activating a relay as shown in Figure 6. That relay in turn activated a telegraph sounder relay causing the receiving circuit to sound exactly like the clickity clack of a telegraph system when receiving morse code. In Figure 6 the device marked A B C is the coherer, R is the relay, S the sounder, B_1 and B_2 batteries, and L being circuit inductance. The glass tube in the picture on the right in Figure 6 is the coherer while the two tan colored cylinders are the relay R.

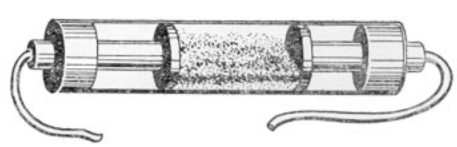


Figure 7 A coherer (source: electronics.stackexchange.com)

In 1890 Edouard Branly had discovered that the resistance measured across the ends of a tube filled with metal filings was normally very high. But, in the presence of an electromagnetic wave the metal filings became fused together resulting in a very low resistance across the tube. The resistance remained low after the electromagnetic wave disappeared until the tube was tapped returning the particles to their original state. The coherer was in affect a switch that was turned on by the presence of an electromagnetic wave and turned off by some sort of mechanical tapping arrangement. Branly was interested in why this occurred. He had no interest in wireless communications. In 1894 Oliver Lodge expanded on Branly's work developing a superior cohere which Lodge used in experimenting with the similarities between light and electromagnetic waves. It was Lodge that gave the device its name.

The coherer that Marconi used in his receiver was an improved version of Lodge's cohere. Marconi's improvement consisted of enclosing the metal filings in a tube partially evacuated of air. For a number of years, Marconi's cohere was the best detector available.

What drove Marconi personally, and in obtaining money from his father and others for his experiments, was transmitting long distances. At first a few feet, then progressively further. In 1894 Marconi transmitted a distance of 2 miles near his home in Bologna, Italy. In March 1899 he transmitted a signal across the English Channel from Dover, England to Boulogne, France. His greatest achievement came on December 12, 1901 when he and his associate transmitted a signal more than 2,000 miles across the Atlantic from Poldhu in Cornwall, England, to St. John's, Newfoundland, Canada.

To Marconi, transmitting longer distances meant using larger antennas. In 1896, as the result of "cut and try" experimentation, Marconi built a vertical antenna by connecting the upper part of the spark gap to a long vertical wire, and the other part to ground, as illustrated in Figure 8. In this figure the transmitting station is shown on the left and the receiving station on the right. Using a vertical antenna considerably increased transmission distance.

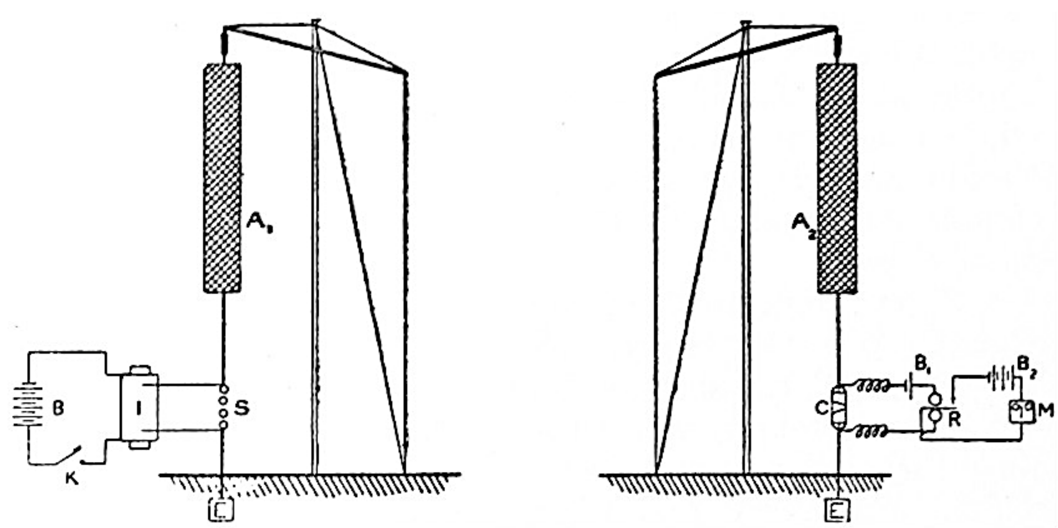


Figure 8 Marconi's vertical antenna (source: Attken)

The rectangular cross hatched object marked A_1 in Figure 8 is a metal sheet forming the upper plate of a capacitor. The lower plate being the ground itself. As illustrated in Figure 4, two relatively large capacitor plates were an integral part of early spark gap transmitters. The object marked A_2 in the receiving station performed the same function. These plates eventually disappeared as the technology evolved. Today's ground mounted vertical antennas date back to Marconi's 1st vertical in 1896.

Increasing the length of the vertical wire lowered Marconi's operating frequency according to the equation

$$f = \frac{c}{\lambda} = \frac{c}{4h}$$

where

c = the speed of light

$\lambda = wavelength$

h = the height of Marconi's vertical antennas which were roughly $1/4\ \lambda$ tall ($1\ \lambda = 4h$).

By using lower frequencies Marconi achieved longer ground wave propagation distances, particularly over salt water. Eventually Marconi wireless systems operated at frequencies from 1 MHz down to 300 KHz requiring very large antennas. The ionosphere at the time was of course unknown.

At his father's insistence, Marconi offered his wireless system to the Italian Navy but they were not interested. Marconi's father was not only rich, but his English born mother was part of a well to do well connected family back in England. In the spring of 1896 Marconi and his mother traveled to England. With access to the right people, the English Post Office became very interested in Marconi's wireless. By law the English Post Office was responsible for all electrical communications in England meaning that all telegraph business in England was under its control. To William Preece (chief engineer at the Post Office) the fledgling wireless technology would also become part of his jurisdiction. Others including Oliver Lodge were experimenting with wireless. But to Preece, young Marconi seemed to be a very promising prospect for developing a workable wireless communications system. Wishing to control the evolving wireless business, Preece threw his support and the prestige of the Post Office behind Marconi providing Marconi with considerable technical and political support.

Marconi's cousin, a very capable business man, strongly suggested that Marconi start a wireless communications company which Marconi did. It was a wholly own family business financed by his extended English family. Because of his well to do family, Marconi had opportunities that no one else involved in wireless had.

Marconi could not compete with the wired telegraph. The extensive world-wide telegraph networks, depicted in Figure 9, were owned and operated by large well-established companies. The wired telegraph networks were far more dependable than the wireless. In addition, morse code speeds were very fast over wired telegraph compared to the slow speed wireless.

Figure 9. World wide telegraph networks in 1891 (source: Wikipedia)

Throughout most of history, all knowledge of a ship was lost from the time it left port until, hopefully, it arrived at its destination. Once at sea, its location and its situation were completely unknown (Figure 10). Ship-to-Shore and Ship-to-Ship communications offered Marconi a golden opportunity which the telegraph business could not compete with. Ship owners were extremely interested in the wireless. But they did not want to buy hardware from Marconi. They wanted Marconi to provide the complete systems including shore-based radio stations plus Marconi build, installed, and operated radio equipment aboard ship.

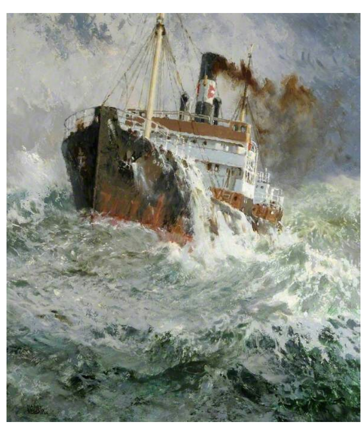


Figure 10 Situation of a ship at sea was unknown through most of history (source: transpress nz)

Marconi's new company did exactly that. For a while in the early 1900s Marconi achieved nearly a complete monopoly in maritime communications. Many ships were equipped with wireless communications by 1908. However, radios on passenger ships were more for the convenience of wealthy passengers wishing to conduct business while at sea than for safety. That all changed with the sinking of the Titanic. Old time maritime morse code radio operators described wireless nets between ships as one big "party line". A typical shipboard radio room in shown in Figure 11.

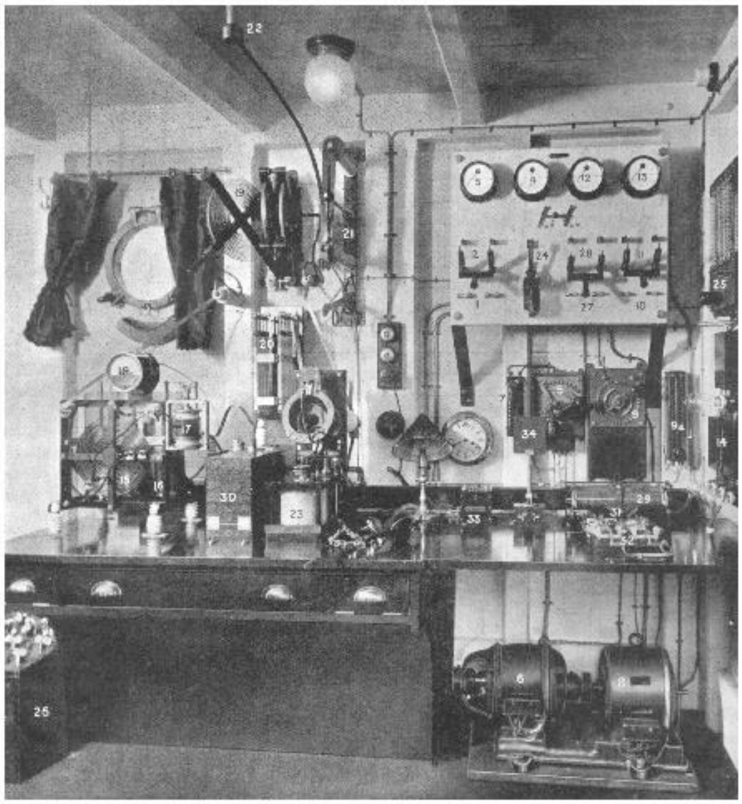
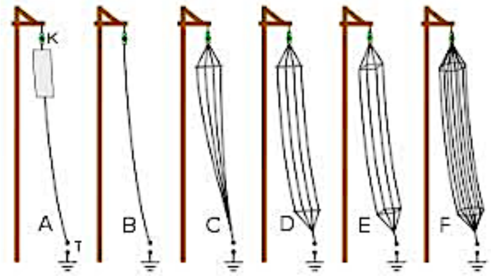


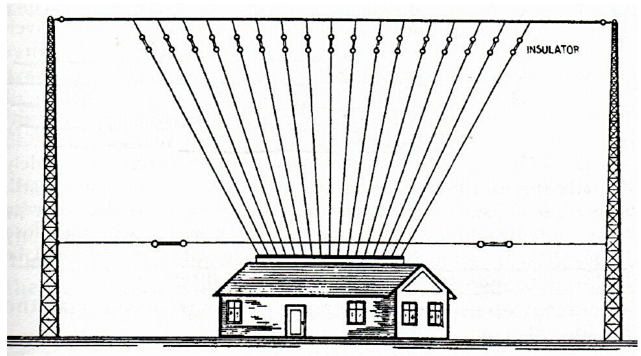
Figure 11 Radio room on the S.S. Chindwin (source: Brown's Signaling)

1.4 Antennas

Vertical antennas were the best long-distance antennas. While they were extensively used on shore, they were not practical for ship board operation because of their extensive height. Typical early vertical antennas are shown in Figure 12. Notice the height of Marconi's first 1901 Cornwall Antenna intended for his transatlantic experiment. The antenna is huge because of its low operating frequency. This antenna blew down in a heavy wind storm and was replaced by a sturdier fan vertical like that shown below.



Various vertical wire antennas



A vertical fan antenna

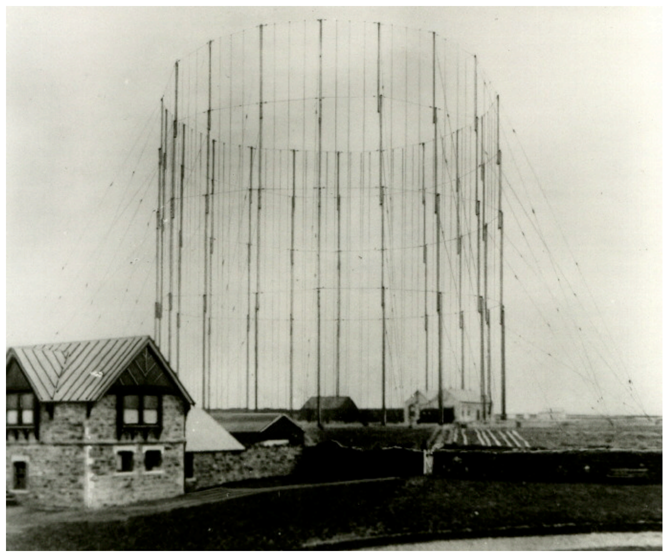
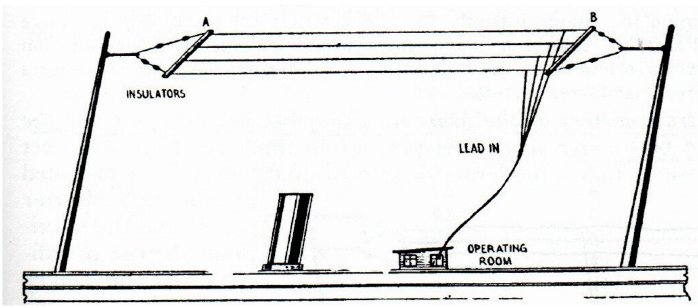

Marconi 1st Cornwall Antenna 1901

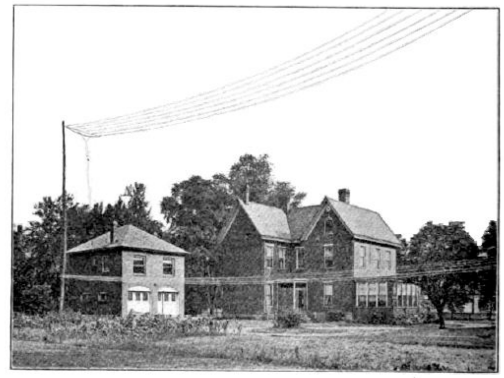
Figure 12 Early vertical antennas (source: Wikipedia)

Inverted L and T antennas were the most popular for ship board operation. Examples of early inverted L antennas are shown in Figure 13. Some land based inverted L antennas, such as the lower antenna in Figure 13, were as much as a mile long. Figure 14 depicts some early T antennas. The Vertical, inverted L and T antennas were all $1/4\ \lambda$ radiating elements requiring an extensive ground to completed the antenna. Aboard ship the steel hull provided the needed ground. On shore elaborate radial systems buried in the ground were installed. Nearly all of the early antennas used multiple parallel wire assuming that they would radiate more power and provide better capacitance between the antenna and ground than single wire antennas. Although that assumption is questionable, multiple parallel wires are often used on 80 meters to increase antenna bandwidth. One wavelength at the frequencies being used (1 MHz down to 300 KHz) ranged from 900 to over 3,000 feet. Inverted L and T antennas, even very tall ones, were only a small fraction of a wavelength $(0.06\ to\ 0.2\ \lambda)$ above ground. Consequently, these antennas were all Near Vertical Incident Skywave (NVIS) short range antennas with very high radiation angles as illustrated in

Figure 15. Inverted L and T antennas are still used today by amateur radio operators on 160 and 80 meters.



Ship board inverted L antenna



Land based inverted L antenna

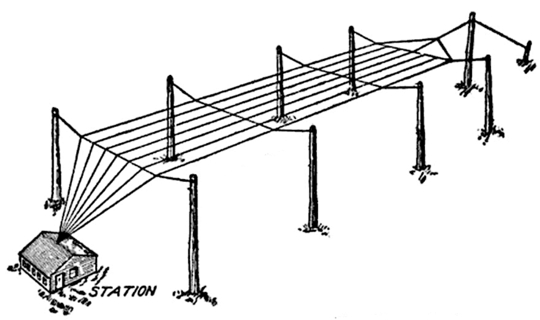
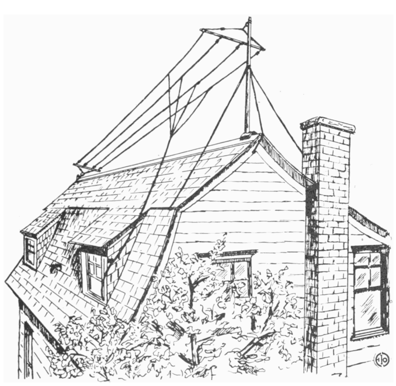


Figure 13 Early inverted L antennas (source: Bucher)



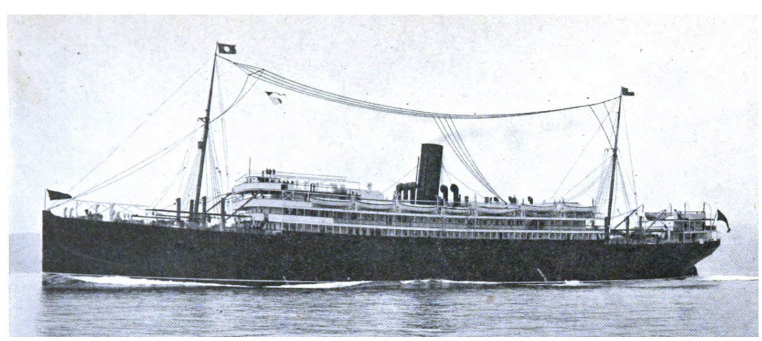


Figure 14 T antennas (source: Wikipedia)

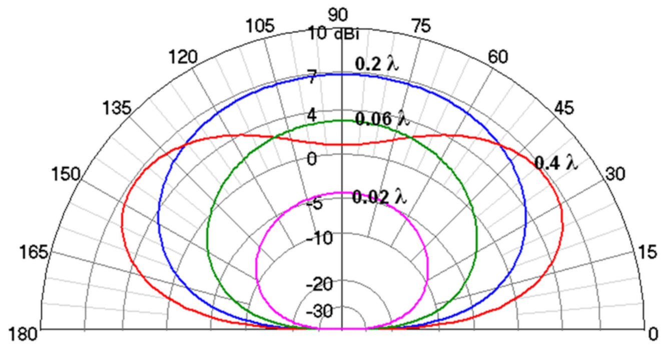


Figure 15 NVIS antenna vertical radiation patterns (source: Researchgate)

1.5 Better Receivers

The coherer was a miserable detector, but for a long time no one could come up with a better device. Many different designs were tried including an electrolytic detector. By 1906 the coherer had been replaced by the carborundum crystal detector which was simple, low cost, and very effective. The crystal detector was extensively used throughout the wireless industry. Crystal detectors allowed anyone with even minimal mechanical skills to build crystal radio sets dramatically increasing the number of amateur radio operators. Figure 16 shows the schematic of a typical crystal radio receiver and crystal detector.

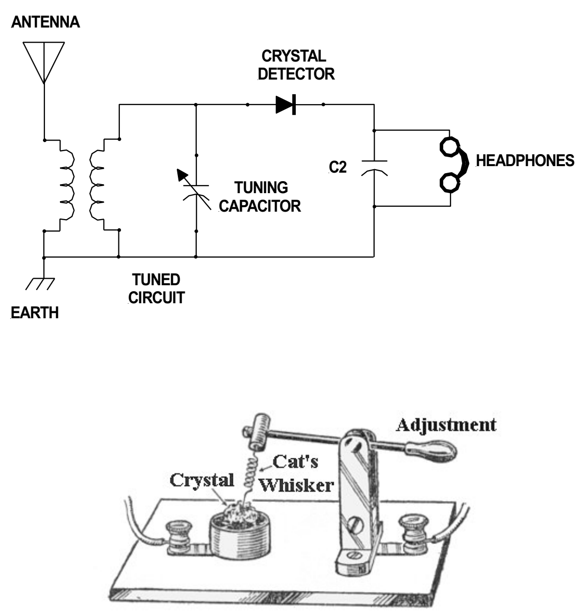
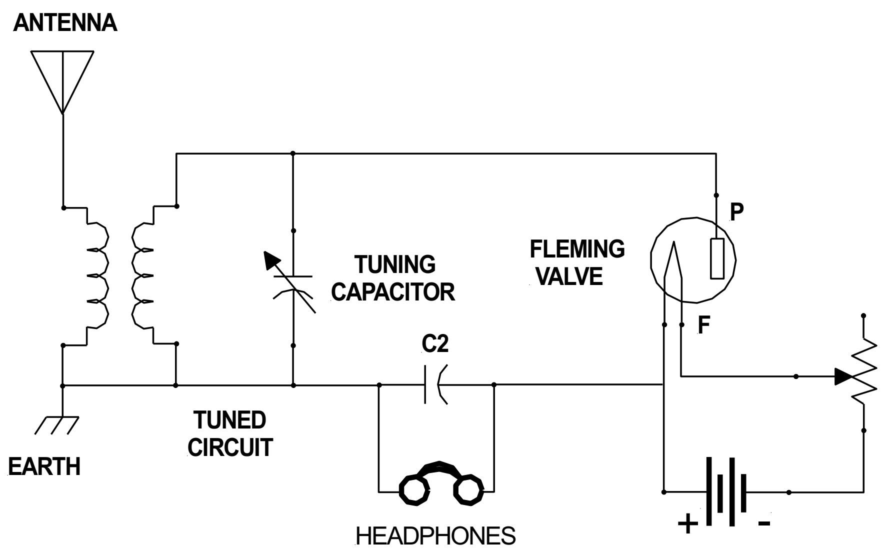


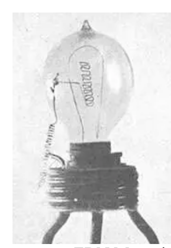
Figure 16 Crystal radio receiver (source: Wikipedia)

A wire (cat's whisker) in contact with a piece of carborundum crystal formed a diode which in series with the capacitor C2 demodulated spark gap wireless signals allowing them to be heard in the headphones. The operator probed around the surface of the crystal until a point was found at which radio signals could be heard. The variable tunning capacitor allowed the operator to tune to various radio stations.

In 1904 John Fleming invented the vacuum tube diode while working on various wireless detectors at the Marconi Company. The invention was based in part on work he had done on light bulbs in the 1880s at the Edison Company. The Fleming valve consisted of two elements enclosed in a glass bulb evacuated of air, a filament F and a plate P. An electrical current flowing from a battery through the filament heated the filament to a high temperature. The high temperature caused energetic electrons to escape from the filament and flow through the vacuum to the positively charged plate. Electrons could not flow in the opposite direction. Consequently, the Fleming valve performed the same diode function as the carborundum crystal.

A schematic of a Fleming Valve receiver is shown in Figure 17. Conceptually, it is the same as the crystal radio set shown in Figure 16 with the carborundum crystal replaced by the Fleming valve. The vacuum tube diode worked well as a detector, but the crystal detector worked as well and was much less expensive.





EDN Magazine

Figure 17. Fleming valve diode receiver

1.6 Beginning of the Electronic Age

In 1906 Lee DeForest invented the triode vacuum tube which he called the Audion. Many claimed that he had simply added a 3^{rd} element to Fleming's diode. The courts finally concluded that DeForest's work was original. Fleming stated several years later that the idea of adding a third element to his diode had never occurred to him.

The three elements of an Audion vacuum tube were the filament F, plate P, and grid G enclosed in a glass bulb evacuated of air, as illustrated in Figure 18.

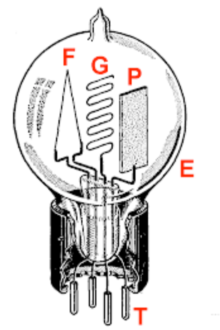


Figure 18 DeForest audion 3-element vacuum tube. (source: Wikipedia)

The audion worked the same as Fleming's valve with the exception that the grid controlled the flow of electrons from the filament to the plate. A small voltage change between the filament and grid, caused by a weak input signal, resulted in a large voltage change between the plate and the filament producing an amplified version of the input signal as illustrated in Figure 19.

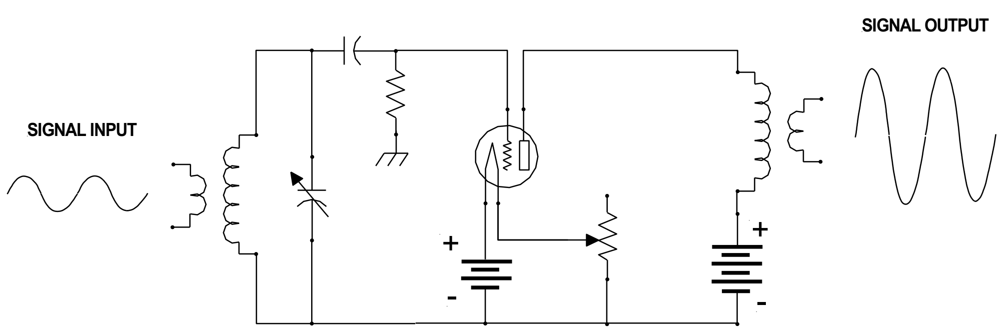


Figure 19 Audion amplifier circuit (source: author)

The Electronic Age as we know it today began on October 26, 1906 when DeForest applied for a patent on "a device for amplifying feeble electrical currents". However, despite its enormous importance, the audion was basically ignored for five years or so.

Initially, no one, not evening DeForest, appreciated the importance of the audion as an amplifier. In 1914 AT&T (the telephone company) was committed to providing transcontinental telephone service and needed some type of audio amplifier. An audion modified for telephone operation was the answer. AT&T's success with the audion radically changed the wireless industry. Electronic circuits based on the audion quickly replaced the old electro-mechanical devices producing new high performance electronic receivers and transmitters

Edwin Armstrong realized that considerable amplification could be achieved by feeding part of the amplified output signal back to the audion's input causing the signal to be amplified over and over again. Armstrong used this concept to produce a simple highly affective regenerative receiver which he patented in 1914. An early regenerative receiver is shown in Figure 20. The regenerative receiver was extensively used by amateur radio operators for many years. However, if not carefully adjust, it would go into oscillation becoming a transmitter. At least some of the interfering noise on the amateur radio bands were caused by improperly tuned regenerative receivers.

Figure 20 Regenerative receiver (source: Wikipedia)

The problems of oscillation were not restricted to regenerative receiver. Audion amplifiers frequently broke into oscillation. However, it was quickly realized that with appropriate circuitry an audion could intentionally be caused to oscillate turning it into an RF oscillator capable of generating highly desirable continuous radio waves.

1.7 Long Distance Low Frequency Communications

Spark gap technology improved from 1900 through 1915. Other electro-mechanical transmitters also evolved. Arc transmitters were the most important alternate technology. Unlike the spark gap, an arc transmitter, illustrated in Figure 21, transmitted a continuous wave (CW). Spark gap transmitters, with their highly damped output signals (Figure 22) were excellent for wireless telegraph, but they could not transmit voice. Arc transmitters with their continuous wave outputs could, and by 1910, were transmitting voice as well as wireless telegraph. But there was no market for wireless telephone until the 1920s when commercial broadcast radio stations emerged.

By 1915 arc transmitters were the most powerful transmitters in the world. Built under contract to the Navy, the Federal Company in Palo Alto, California built arc transmitters (Figure 23) ranging in power from 100 to 1,000 KW and operating, like all commercial transmitters of the day at 300 to 1,000 meters (1,000 KHz down to 300 KHz). Arlington, Virginia (Figure 24) was the hub of the

Navy's network. Arlington could contact Panama and San Francisco at night and sometimes during the day. Hawaii could be reached by San Francisco and sometimes directly from Arlington. Guam and Hawaii talked regularly, all before amateur radio operators discovered short wave radio.

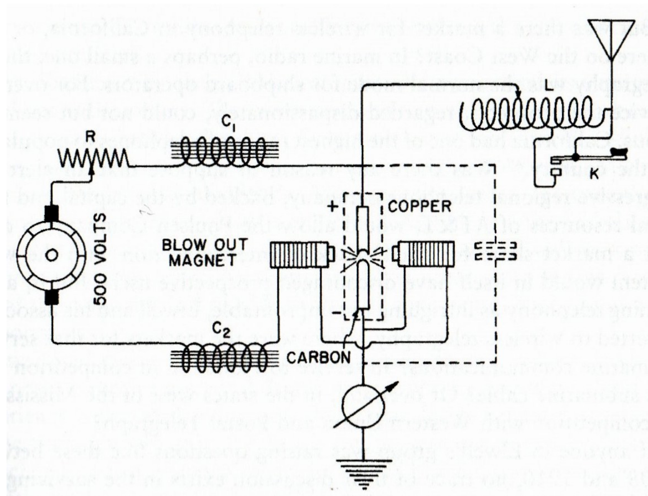


Figure 21 Arc Transmitter (source: Bucher)

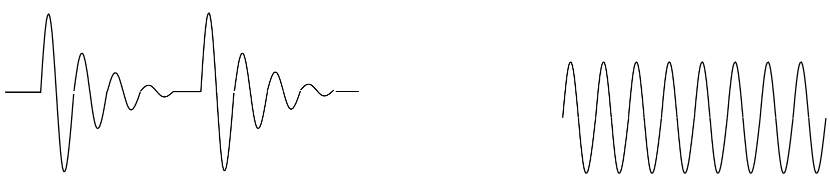


Figure 22 Highly damped and continuous waves (source: author)

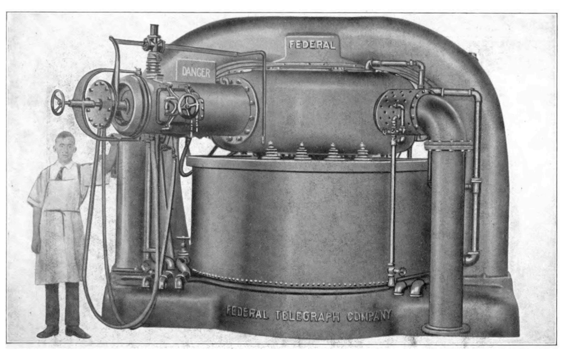


Figure 23 Federal 100 KW Arc transmitter weighing around 60 tons (source: SlideShare)

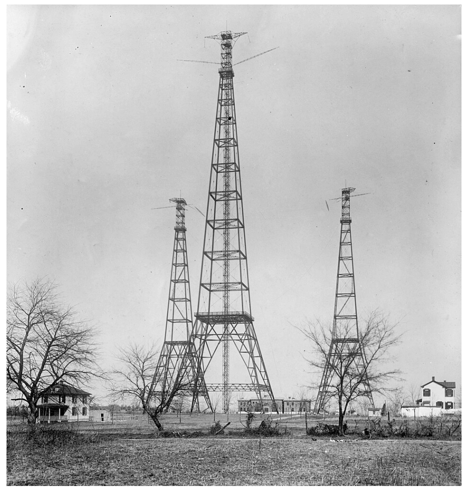


Figure 24 Navy's Arlington, Va. radio station antennas (source: Pinterest)

1.8 Propagation Theory Did Not Match Results

In 1923, despite all that had been accomplished, how radio waves propagated was still unknown. From the mid 1990s through 1923 it was believed that radio waves traveled along the ground, or across the ocean surface. Complex theories were developed attempting to explain the long-distance propagation of radio waves. But the theories did not match the results of carefully conducted experiments.

In 1902 Oliver Heaviside (English) and Arthur Kennelly (United States) independently theorized that a conducting layer in the upper atmosphere could account for long range wireless transmissions. They suggested that wireless signals traveled back and forth between the conducting layer and earth's surface, first reflecting from the conducting layer, then reflecting from the earth as illustrated in Figure 25.

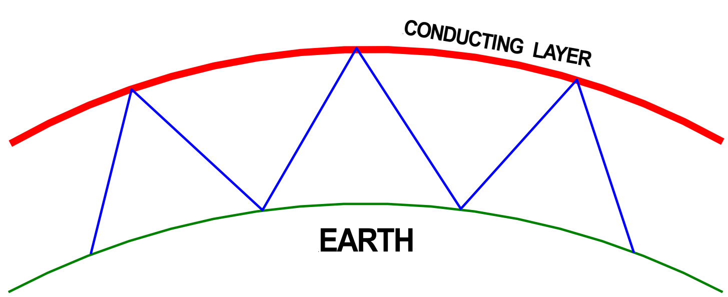


Figure 25 Presumed reflection of HF radio waves in Earth's ionosphere (source: ResearchGate)

In 1910 English engineer and scientist William Eccles theorized that a body of electrons and positive ions existed in the upper atmosphere resulting from sunlight decomposing the air. He proposed that the level of ionization gradually increased as sunlight entered the upper atmosphere, reached a peak, and then slowly declined until it finally disappeared as illustrated in Figure 26. In this figure the block dots illustrate the level of ionization.

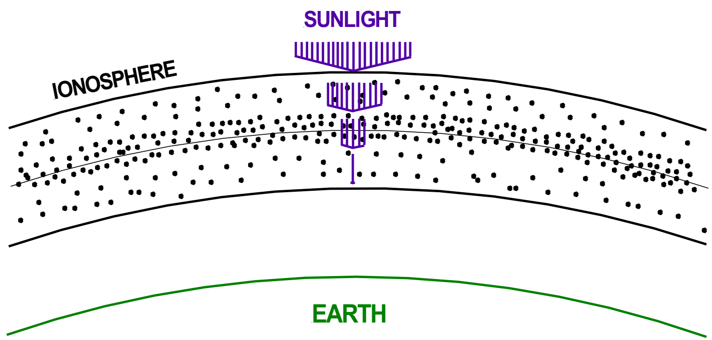


Figure 26 Ionization of Earth's upper atmosphere (source: author)

Based on this concept, radio waves traveling upward into the atmosphere were refracted, bent back toward Earth, as they encountered increasing levels of ionization (Figure 27). The refraction of light at the boundary between two transparent medium was well understood. Eccles presumed that radio waves acted similarly. We are all familiar with refraction. Refraction creates rainbows and separates sunlight into its various red through purple colors as it passes through a prism (Figure 28).

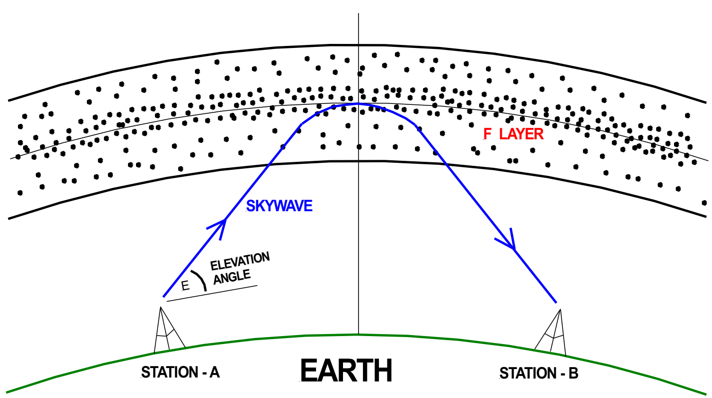
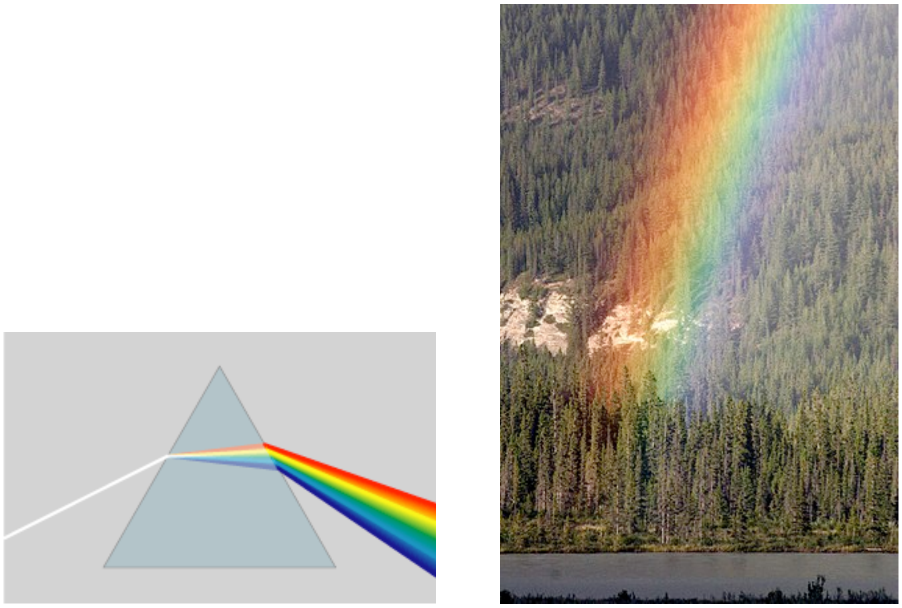


Figure 27 Refraction of radio waves in the upper atmosphere (source: author)



Light refracting through a prism and through water droplets forming a rainbow of colors

Figure 28 Refraction of light (source: Wikipedia)

But all of these ideas were simply theories. There was no proof that any of it was actually true.

In the early days of radio, amateur radio operators, with their home built wireless equipment, crowded the commercial wireless frequency bands. This was a serious problem from the perspective of commercial and military radio operators. To solve the problem legislation was enacted restricting the frequencies at which amateurs could operate. From 1912 on amateurs were exiled to the seemingly worthless short-wave frequencies of 200 meters and down, better known today as the 160, 80, 40, 20, and 10 meter frequency bands. While most amateurs operated at 200 meters (around 1.5 MHz), a few ventured down into the lower wavelength frequency bands. In the process they discovered the long-distance capabilities of “short wave” radio. Amateurs began communicating from the United States to Europe, New Zealand, and the Pacific rim on a regular basis using transmitters operating at relatively low power levels of 1,000 watts or less and small 20 meter antennas. In contrast, military and commercial companies operating in the 300 to 1,000 meter frequency bands required transmitters ranging in power from a hundred thousand to a million watts (Figure 23) and huge antennas (Figure 24) to cover the same distances.

Amateur radio operators contributed significantly to the understanding of long-distance radio communications. They discovered that short wave radio disobeyed the accepted theories of radio propagation. Short wave propagation, unlike long wave, seemed to involve only the atmosphere. Strange things happened at short wavelength frequencies. For example, it was possible to communicate with distant radio stations while closer radio stations could not be heard. Radio signals seemed to skip over the close by radio station creating what became known as the skip zone shown in Figure 29. We know today that a signal transmitted at a low elevation angle E travels further than a signal transmitted at high angles. It follows that relatively high elevation angles are required to communicate with close by stations. However, if the elevation angle is too high the signal will penetrate the ionosphere and be lost to outer space as is the case with the pink trace in Figure 29. The elevation angle at which penetration begins to occur is called the Maximum Usable Angle (MUA) illustrated by the red trace in Figure 29. Consequently, Station-B is the closest station that Station-A can communicate with. Stations between A and B can not be heard. They are skipped over. Why this occurred was not understood in the 1920s. At the time the skip zone was a complete mystery.

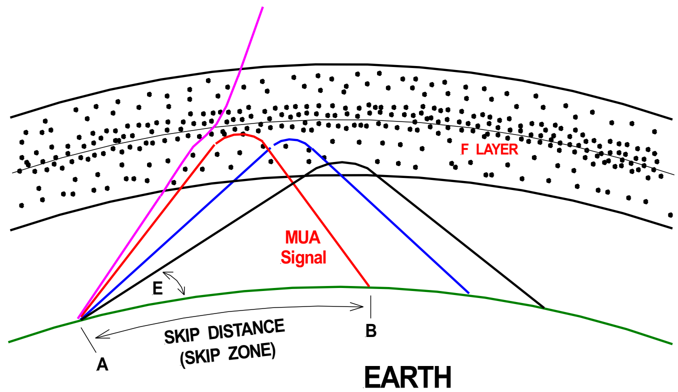


Figure 29 Skip distance (source: author)

In April 1924 Edward Appleton studied the strength of radio signals received at Kings Collage Cambridge from the London BBC (British Broadcasting Company) radio station approximately 65 miles away. He discovered signal strength was constant during the day but varied at night. He theorized that at night he was receiving two signals, one traveled along the ground the other reflected from the upper atmosphere. Further investigations showed that the radio waves were reflecting from atmospheric ionization about 60 miles above the Earth from what he called the electrified or E-layer. He later discovered that short wave radio signals passed through the E-layer and were reflected about 100 to 200 miles above the Earth. He named this electrified region the F-layer as illustrated in Figure 30. The signal absorbing D-layer was discovered later. Appleton's experiments proved that there were in fact ionized regions in the upper atmosphere that refract radio waves back to Earth.

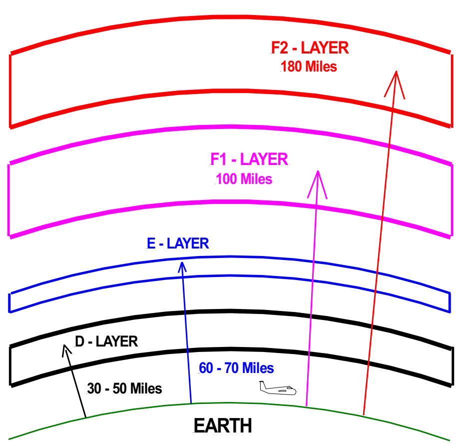
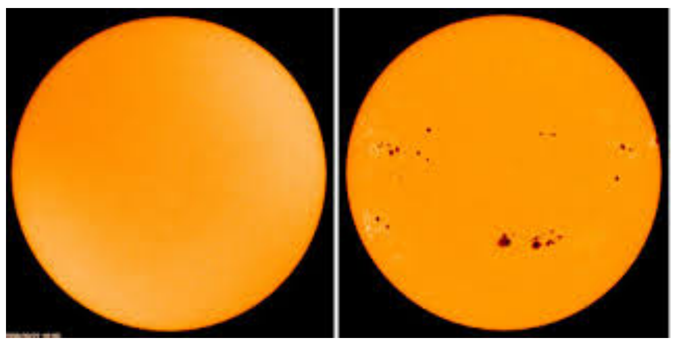


Figure 30 Layers of the ionosphere (source: author)

Early on it was known that signal strength and quality varied throughout the day, and seasonally. Around the world short wave communications had been excellent and exciting since its discovery in the mid 1920s. Then in the early 1930s, for some unexplained reason, excellent long-distance communications disappeared. It was difficult to make any long-distance short-wave contacts. In the late 1930s long distance propagation conditions returned. It soon became apparent that signal strength and quality also varied with the 11 year solar cycle. Solar maximum (large numbers of visible Sun spots Figure 31) occurred in the latter half of the 1920s (Figure 32) producing excellent long-distance short-wave communication conditions. By 1935 sunspots had mostly disappeared creating solar minimum and poor long-distance communications.



Solar Minimum Solar Maximum

Figure 31 Sun at solar minimum and maximum (credit: springer.com)

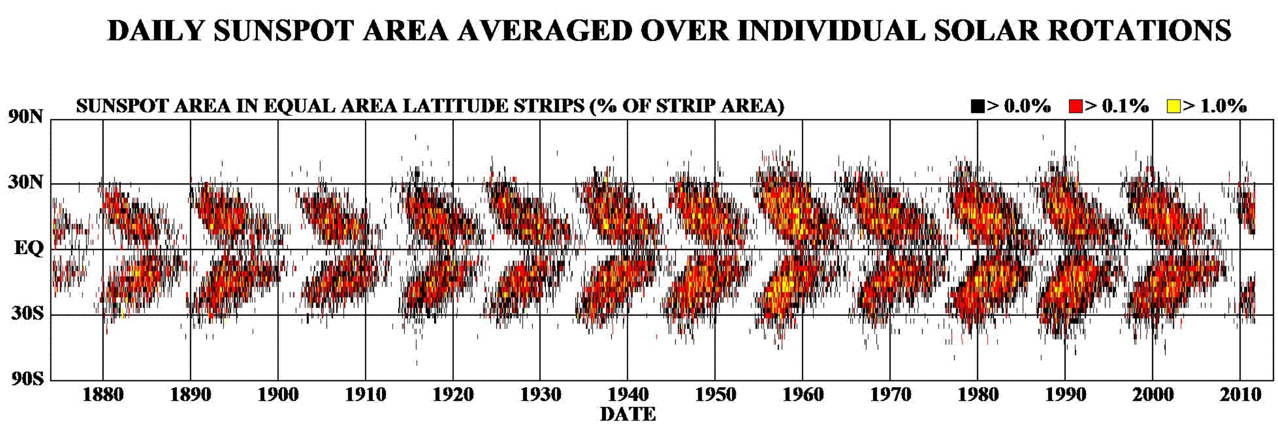


Figure 32 Solar sunspot cycle (source: D. Hathaway)

By 1940 it was clear that the Sun created the conditions needed for long distance short wave (high frequency) radio communications. It was also apparent that disruption of that communications was directly or indirectly the result of violent events occurring on the Sun. An understanding of the Sun and solar dynamics are thus necessary to understand and utilize the ionosphere.

The structure of the Sun and its violent nature, its interaction with the Earth's magnetic field, plus formation and disruption of the ionosphere are the subject of the following chapters. The last chapters discuss utilization of the ionosphere for HF communication and the problems that we must deal with in doing so.

References

Yeang, Chen-Pang; "Probing The Sky With Radio Waves"; The University of Chicago Press, 2013

DeSoto, Clinton B.; "200 Meters & Down" The American Radio Relay League, 1936

McNamara, Leo F.; "The Ionosphere: Communications, Surveillance, and Direction Finding"; Krieger Publishing Company, 1991

Davies, Kenneth; "Ionospheric Radio"; Peter Peregrinus Ltd.,1990

Hunsucker R. D.; Hargreaves, J. K.; "The High-Latitude Ionosphere and its Effects on Radio Propagation"; Cambridge University Press 2003

Nichols, Eric P.; "Propagation and Radio Science"; The American Radio Relay League, Inc. 2015

Aitken, Hugh G. J.; "Syntony and Spark, The Origins of Radio" Princeton University Press, 1985

Aitken, Hugh G. J.; "The Continuous Wave Technology and American Radio, 1900 - 1932" Princeton University Press, 1985

Bucher, Elmer Eustice; "Practical Wireless Telegraphy" Wireless Press, Inc, 1917

Collins, Archie Frederick; "The Radio Amateur's Hand Book" 1922

Collins, Archie Frederick; "Wireless Telegraphy" McGraw Publishing Company 1905

Coe, Lewis; "Wireless Radio A History" McFarland & Company Inc, 1996

Trowbridge Jr., Calvin D.; "Marconi" Calvin D Trowbridge Jr., 2009

Devoldere, John; "Low-Band DXing" fourth edition; ARRL, 2005

Levis, Curt A. ; Johnson, Joel T.; and Teixeira, Fernando L.; "Radiowave Propagation Physics and Applications"; John Wiley & Sons, Inc., 2010

Goodman, John M.; "Space Weather & Telecommunications"; Springer Science+Business Media Inc. 2005

Forbes, Nancy and Mahon, Basil; "Faraday, Maxwell and the Electromagnetic Field" Prometheus Books 2014

Nahin, Paul J.; "Oliver Heaviside" The John Hopkins University Press 2002

Stross, Randall; "The Wizard of Menlo Park" Three Rivers Press 2007

Moyer, Albert E.; “Joseph Henry” Smithsonian Press 1997