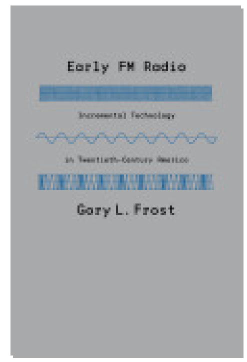
Early FM Radio

Frost, Gary L.



Published by Johns Hopkins University Press

Frost, L..
Early FM Radio: Incremental Technology in Twentieth-Century America.
Baltimore: Johns Hopkins University Press, 2010.
Project MUSE., https://muse.jhu.edu/.

CHAPTER ONE

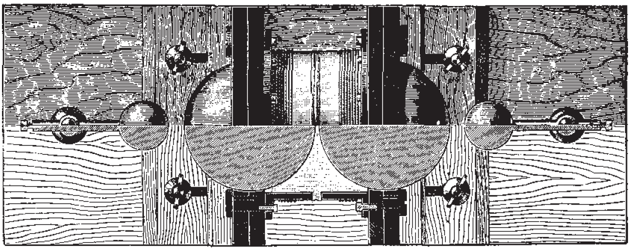
AM and FM Radio before 1920

> The process of altering the length of the emitted wave must be abandoned.
>
> *Valdemar Poulsen, inventor of FM radiotelegraphy and its first critic, 1906*

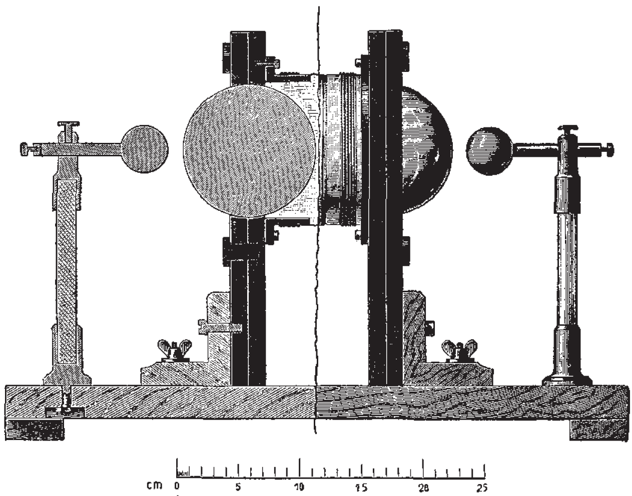
The Spark Gap and the Coherer

To understand why frequency-modulation radio first appeared in 1902, one must know something about the technological context of the radio at that time. Two devices—the spark gap, used in transmitters, and the coherer, the basis of almost all early wireless receivers (figs. 5 and 6)—had defined the possibilities and the limitations of the art since Guglielmo Marconi invented radio during the 1890s.[1] The simplest form of spark gap featured a pair of spherical brass electrodes separated by one or two inches of air. When a battery-and-coil circuit caused the electric potential (i.e., the voltage) between the electrodes to rise above a certain threshold, a spark leaped across the gap, discharging violently and emitting a train of *damped waves*—invisible electromagnetic (EM) waves that decrescendoed to nothing in a fraction of a second (fig. 7). The phenomenon resembled the dropping of a stone into a still pool of water, or a clapper striking a bell.

To send the dots and dashes of Morse code messages, Marconi borrowed a method of modulation from overland wire telegraphy. For decades telegraph operators had utilized a key—a hand-operated electrical switch—to signal either full-power *marks* when current was on or zero-power *spaces* when current was off. Though unnamed, this method could be described as a type of binary amplitude modulation—that is, transmission occurred either at full amplitude or



MARCONI'S SPARK-GENERATOR. FROM ABOVE.



MARCONI'S SPARK-GENERATOR. CROSS SECTION, SIDE VIEW

Fig. 5. Marconi Spark Gap Transmitter, 1898. Depicted are two gaps between metal spheres. Other versions of the device had one or several gaps. Adolph Slaby, "The New Telegraphy," *Century* 55 (April 1898): 880.

not at all, a short mark corresponding to a dot, a somewhat longer mark a dash. Binary amplitude modulation required substantial modifications before Marconi could adapt the method to wireless. A single train of damped waves sufficed to signal a dot, but several trains had to signify a dash, which necessitated rapid refiring on the part of the spark gap. Marconi therefore designed his transmitter to recharge the spark gap quickly and automatically, so that holding the key down caused the transmitter to sputter out continuous trains of damped waves, one closely ranked group after another.

In the receiver side of his system, he complemented the spark gap with a coherer. This was a hollow glass tube, approximately six inches long and packed with metal filings (Marconi preferred a mixture of nickel and silver). Metal plugs at each end compressed the filings and functioned as electrical terminals besides.

Fig. 6. Branley Coherer, 1902. Inside an evacuated glass tube are metal filings. Silver plugs act as terminals and enclose the filings by capping the open ends of the glass tube. Detail from figure in Ray Stannard Baker, "Marconi's Achievement: Telegraphing across the Ocean without Wires," *McClure's Magazine*, February 1902, 291.

If all went well—which rarely occurred for reasons discussed later—the coherer detected radio waves that caused the filings to transform from a nonconductive (off) state to a conductive one (on), much like an electric switch. An opposite transformation—from on to off—required more than merely removing the tube from the presence of EM waves, however, as the following description of an off-on-off sequence illustrates:

Step 1. When the amplitude of nearby electromagnetic waves remains below a certain threshold amplitude A_n, the coherer's filings are normally nonconductive (off).
Step 2. When electromagnetic waves rise above A_n, the coherer instantaneously transforms to a conductive state (on).
Step 3. When electromagnetic waves subsequently fall below A_n in amplitude, the coherer remains on—until some physical motion disturbs the arrangement of the metal filings.
Step 4. Tapping the side of the coherer, therefore, causes the device to revert to a nonconductive state (off).
Step 5. The newly nonconductive coherer now exists in the same nonconductive state as in step 1, but with a new threshold amplitude, A_{n+1}. Its value corresponds to the arbitrary physical rearrangement of its filings and might differ substantially from A_n.

It should be emphasized that the coherer could not detect messages by itself because, once the device turned on, it stayed in that state until two events took place. First, local electromagnetic waves had to subside in strength below the trigger threshold A_n; and, second, some material object had to jar the internal metal filings with force adequate to alter their physical arrangement, causing the coherer to revert to a state of nonconductivity. To meet the second requirement, Marconi fastened to the coherer a "vibrator," or "tikker"—a small electrically driven hammer mechanism that continuously rapped the outside of the glass

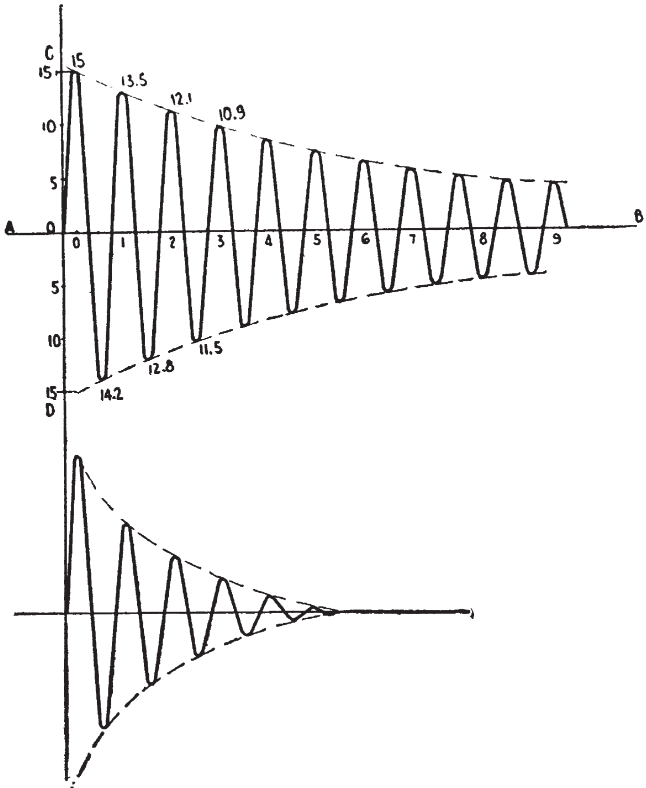


Fig. 7. Damped Waves. Two trains of waves illustrate "feebly damped" (*top*) and "strongly damped" (*bottom*) waves. Dotted lines trace the "envelope" of the wave train peaks. Elmer E. Bucher, *Practical Wireless Telegraphy: A Complete Text Book for Students of Radio Communications* (New York: Wireless Press, 1921), 1.

tube. Finally, to record messages, Marconi wired his coherer-switch to operate either a buzzer or a writing device, usually a paper-tape inker.

This was the theory of Marconian wireless telegraphy, but in practice the spark gap and coherer amounted to what Thomas Hughes has termed "reverse salients" and Edward Constant calls "presumptive anomalies."[2] That is, practitioners by and large knew that no matter how much the two devices could improve, no one would ever overcome their inherent limitations, and sustain progress in the art. The coherer's fickle responsiveness especially rankled operators. Each tap of the vibrator caused the filings to jump from one level of sensitivity to another with kaleidoscopic capriciousness. At one moment, for example, a coherer might fail to detect a transmitter situated only yards away, but a fraction of a second later electromagnetic waves from a source dozens of miles distant could trigger the filings into a conductive state. Further, the coherer, as a two-state binary switch, proved entirely inadequate for wireless telephony, because replication of speech required a continuously and proportionally responsive detector—that is, a detector capable of tracking variations corresponding to the instantaneous amplitude of sound.

As for the spark gap, the heart of the Marconi transmitter, the amplitudes of its damped waves *could* be varied—modulated, as practitioners say—which tempted some to try adapting the device to wireless telephony. No one succeeded entirely, but the Canadian-born inventor Reginald Fessenden came close. In December 1900 he transmitted speech on Cobb Island, Maryland, by amplitude-modulating a continuously triggered spark gap. Fessenden obtained only minimally intelligible reception, though, on account of static-like noise created by the spark.[3]

Nor could spark gaps and coherers overcome natural impediments to radio wave propagation. Electromagnetic radiation does not always move along predictable paths. Local weather, upper-atmospheric conditions, sunlight, and the lengths and amplitudes of the waves themselves chaotically affect the attenuation and refraction of electromagnetic waves. A large mass, such as a hill, office building, or forest, for example, can absorb electromagnetic energy or create a virtual mirror that causes waves to carom off in another direction. Although early wireless pioneers understood the laws of propagation poorly, if at all, experience quickly taught them two dismal facts: first, operating more than one transmitter on the same wavelength invited interstation interference; and, second, the strength of a signal at the receiver might fade, which multiplied the ill effects of the coherer's notoriously erratic sensitivity.

Tuning and the Resonant (*LC*) Circuit

Resonance, the principle behind tuning, also figured prominently in the technological context of early radio. All circuits, whether a piece of wire or a complete radio transmitter, possess the complementary reactive properties of inductance, symbolized by *L*, and capacitance, symbolized by *C*. (Circuits also contain resistance, which we neglect here for the sake of simplicity.) Stimulating a circuit electrically causes it to resonate naturally at a specific wavelength. In electrical circuits, this mathematical formula defines the length of the resonant wave

$$\lambda_i = \frac{c}{f_i} = 2\pi c\sqrt{L_i C_i}$$

where:

λ_i = instantaneous resonant wavelength (measured in meters),
f_i = instantaneous resonant frequency (measured in cps),
L_i = instantaneous circuit inductance (measured in henrys),
C_i = instantaneous circuit capacitance (measured in farads),
π = a constant, approximately 3.14, and
c = the speed of light, a constant equal to 299,800,000 meters per second.

This formula boils down to some simple concepts. Because c and π are constants, the wavelength λ_i is proportional to $\sqrt{L_i C_i}$. Thus, increasing either L_i or C_i will increase λ_i (and decrease the resonant frequency, f_i). Conversely, if one decreases either L_i or C_i, λ_i will also decrease.

Traditionally, practitioners have used the preceding resonance formula far more often to understand resonant circuits than to construct them. As the historian of radio Sungook Hong observes, tuning was primarily "not a mathematical principle but a *craft*."[4] Early wireless pioneers devoted much of their time to perfecting practical techniques for adding reactive components—either inductive or capacitive—to a circuit for the purpose of adjusting it to resonate at a precise wavelength. Amateur radio operators usually accomplished this by making their own components. Coiling wire around a cylindrical oatmeal box or an iron core, for instance, provided inductance. A stack of metal plates, each sandwiched between layers of air, oil, or paper, made up a condenser, the component that supplied capacitance. Tuning to a particular station was commonly accomplished by making the inductive element of an *LC* circuit variable and adjusting the inductance *L* until the circuit resonated at the wavelength of the station's frequency. One could do the same thing with the capacitive element by mounting on an axle half of several interleaving plates constituting a condenser. Rotating the axle changed the capacitance *C* and thus the wavelength. Until a few years ago all consumer radio receivers employed fixed-value inductance and a mechanically variable condenser for tuning purposes. A knob on the front panel of a home radio set was fastened to the axle of a variable condenser. Listeners tuned their radios to a station by adjusting this knob to a number on a dial that corresponded to the station's carrier frequency—a method still employed in cheaper radio receivers. Today, electronic circuits have replaced this mechanical arrangement to do the same thing.

Cornelius Ehret and the Invention of Frequency-Modulation Radio

Although most histories of frequency-modulation radio state that Armstrong invented the technology in December 1933, FM appears more than thirty years earlier in American and Danish patent records. On 10 February 1902 a Philadelphian named Cornelius Ehret filed a patent application for a frequency-modulation system.[5] Seven months later, Valdemar Poulsen applied for a Danish patent for a radio-frequency "arc oscillator" that also employed FM.[6] These two men shared little beyond being contemporaries, however. Ehret began as an unknown amateur and, despite his invention, remained so, but Poulsen had already achieved international recognition by inventing magnetic recording. Further, Ehret explicitly claimed to have invented a frequency-modulation system of wireless telegraphy and telephony, though he failed to make a functional prototype. By contrast, Poulsen eventually made a radiotelegraph system using a method of frequency modulation that he would renounce, but which practitioners would copy for more than two decades.

As the first American holder of a frequency-modulation radio patent, Ehret ranks among the most obscure inventors in the history of wireless. Except for a short article that appeared nearly seventy years ago in *Communications* magazine, virtually no twentieth-century history of radio mentioned him.[7] He does appear in patent court records, however. In 1959, New York's Southern District Judge Edmund L. Palmieri decided in favor of Armstrong's patent infringement suit against Emerson Radio, which had cited Ehret's system to dispute the novelty of wideband FM. Palmieri acknowledged that "the Ehret patent was one of the earliest patents in which it was proposed to transmit and receive intelligence by varying the frequency of a radio wave," but he dismissed outright Ehret's influence on modern FM radio by declaring that "the Ehret patent did not teach anything at all concerning the problem of reducing static and noise in radio signaling. It did not refer to, or suggest anything concerning, the bandwidth to be employed in frequency modulation or the extent of variations in frequency to be employed. It did not refer to or suggest limiting in a frequency modulation receiver." In other words, the fact that Ehret neglected to specify a channel width, employ a limiter circuit, and claim a reduction in "static and noise"—all features that Judge Palmieri attributed to Armstrong FM— banished Ehret to the backwaters of history. "There is no evidence," Palmieri concluded, "that the Ehret patent had any impact upon the art."[8]

The conclusion is fair enough insofar as the law goes, and this study does not challenge Palmieri's assertion that Ehret left no impression on the art of radio design. But historians should not leave the evaluation of Ehret's historical significance to lawyers. Patent courts make winner-take-all decisions chiefly by weighing competing claims of priority, and whether an invention "works," criteria that ignore crucial questions of historical interest that go beyond, say, which individual should be given all the credit for inventing a particular technology. For what purpose, for example, did Ehret envisage his invention? What did his invention reveal about the state of the art of wireless—what some would call the "culture" of wireless—during the first decade of the twentieth century? And what was the relationship, if any, between his FM and the kinds of FMs that followed? Did Ehret discover anything inevitable about frequency modulation?

Because Ehret's patents constitute the entire record of his career, answering these questions presents difficulties. Nevertheless, as the historian Eugene Ferguson has similarly demonstrated for mechanical engineering drawings and architectural plans, Ehret's patents, when carefully decoded in context, reveal far more information than the fact that their inventor made an impractical invention that had little effect on later work.[9] They illuminate, for example, the tacit knowledge of wireless engineering during a period that paved the way for modern FM radio. Virtually all wireless pioneers, and most electricians in 1902, would have understood implicitly the symbolic language of Ehret's circuit schematics, and why he connected spark gaps, condensers, hand-wound inductors and transformers, "air-gaps," tuned circuits, wires, and "telephone-receivers" in the ways he did.

That Ehret invented not only frequency-modulation telegraphy but also a radiotelephony system is indisputable. On 28 March 1905 the U.S. Patent Office, which had divided his original application, issued to the Philadelphian a pair of almost identically worded patents for a system that transmitted and received "the reproduction of speech and other signals through the agency of means responsive to changes or variations in the frequency of the received energy."[10] Although Ehret never explicitly articulated the motivations behind his invention, clearly he sought to overcome two difficulties associated with wireless. One was fading, which still plagues electromagnetic communications. Discarding the skittish coherer, whose electrical properties transformed with every tap of the tikker, Ehret combated fading by designing instead a wireless telegraph receiver with rock-steady sensitivity (figs. 8 and 9). Moreover, unlike the coherer, which detected only the presence and absence of waves, his receiver contained a resonant *LC* circuit that attenuated incoming EM waves roughly in proportion to their length. For instance, suppose capacitance 39 and inductance 40 in figure 8 are chosen so that the circuit resonates with waves 340 meters in length. A wave with

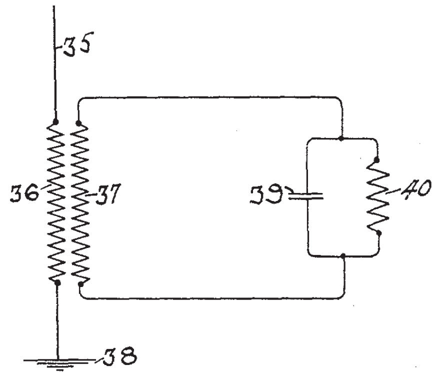


Fig. 8. Detail from Ehret Slope Detector, 1902. In this Ehret frequency-modulation receiver, 36, 37, and 40 specify inductors, and 39 is a condenser. Cornelius D. Ehret, "Art of Transmitting Intelligence," U.S. Patent No. 785,803, application date: 10 February 1902, issue date: 28 March 1905.

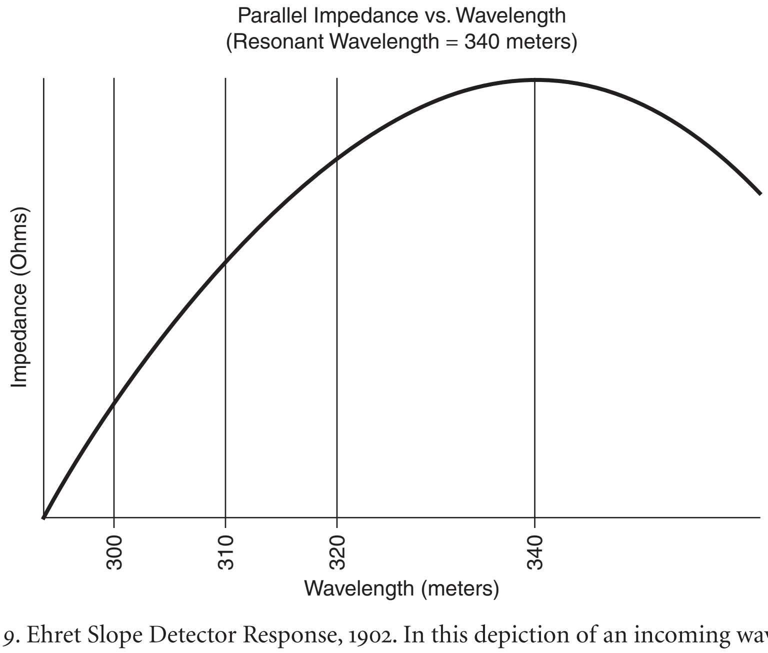


Fig. 9. Ehret Slope Detector Response, 1902. In this depiction of an incoming wave amplitude across a parallel *LC* circuit modeled on Ehret's frequency-modulation slope detector, the circuit is tuned to a resonant wavelength of 340 meters, at which the relative amplitude is maximum. Amplitude decreases approximately linearly from 320 to 300 meters. For telegraphy, a transmitter radiated 300-meter waves to indicate a mark; 310-meter waves indicated a space. For telephony, the instantaneous audio amplitude is proportional to the instantaneous positive-or-negative deviation from a center (i.e., the reference) wavelength of 310 meters.

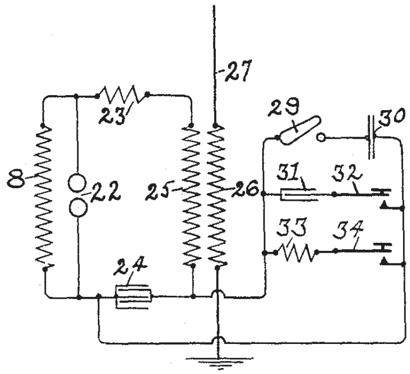


Fig. 10. Detail from Ehret Transmitter, 1902. Drawing of an Ehret frequency-modulation transmitter illustrates two methods of telegraphy and one method of telephony: 22 is a continuously triggered spark gap (trigger mechanism not shown); inductors 8, 23, and 25 constitute the nominal circuit inductance; 24 is the nominal circuit capacitance, which together with the inductance causes the circuit to radiate a wave with a fixed length; and 32 and 34 are telegraph keys that switch condenser 31 and inductor 33 respectively in and out of the circuit, thereby shifting the wavelength of the circuit slightly. This is FSK modulation. For radiotelephony, element 29 switches telephone transmitter (i.e., the microphone) 30 in and out of the circuit. If the microphone is of the inductance type, speaking into it causes the overall inductance of the tuned circuit to alter with the instantaneous amplitude. Thus, the instantaneous frequency of the tuned circuit varies with the amplitude of the speech. Cornelius D. Ehret, "Art of Transmitting Intelligence," U.S. Patent No. 785,803, application date: 10 February 1902, issue date: 28 March 1905.

a wavelength of 340 meters is minimally attenuated, and therefore its amplitude as measured across 39 and 40 is maximum. If the wave decreases in length, say to 300 meters, the circuit will attenuate the wave more, causing the amplitude across 39 and 40 to decrease as well.

Ehret's second goal was to design a system that transmitted and received wireless telegraph and telephone messages. In doing so, he borrowed from Marconian technology, even as he tried to overcome its limitations. Although his telegraph retained the spark gap, his transmitter (fig. 10) radiated damped waves nonstop, as opposed to Marconi's system, which sparked only during keying. Further, Ehret's telegraph radiated waves with either of two lengths, in contrast to the single-wavelength trains of a Marconi transmitter. The patents neglected to mention specific values of capacitance and inductance, so we cannot determine even the range of wavelengths Ehret had in mind, but this time, again for the sake of illustration, let us assume that 300 meters represented a space and

320 meters a mark. To achieve this, the aggregate effects of inductors 8, 23, 25, and 26, and condenser 24 constituted a resonant circuit that radiated 300-meter waves. Switching another inductor or condenser into the circuit caused the circuit's resonant wavelength to increase to 320 meters. Ehret employed telegraph key switches to do this: switch 32 connected and disconnected the condenser 31, and switch 34 caused the wavelength to toggle between 300 meters (space) and 320 meters (mark) by electrically removing and adding inductance 33.

The telephone transmitter resembled the telegraph but with an important difference. Instead of jumping between two fixed wavelengths, Ehret's transmitter instantaneously stretched and compressed the length of its outgoing wave to correspond proportionally to rapid variations of sound. Opening switches 32 and 34 (fig. 10) removed capacitance and inductance used only for telegraphy. Then, closing switch 29 connected to the circuit element 30, a microphone whose diaphragm deflected in proportion to the instantaneous amplitude of the sound. Because the diaphragm was mechanically coupled to a variable inductor (or variable condenser, depending on a designer's preference), the *L* (or *C*) of a tuned *LC* circuit was proportional to the sound's instantaneous amplitude. As an illustration, if the microphone detected no sound, the transmitter radiated waves with a length of, say, 310 meters, the nominal length for the transmitter's tuned circuit. At maximum amplitude, the microphone's inductance wobbled between two extremes and thus pushed and pulled the wavelength to maximum and minimum values of, say, 300 and 320 meters. Similarly, a midlevel amplitude would cause the wavelength to wobble between 305 and 315 meters.

Although no evidence exists that anyone knowingly copied his circuits, Ehret anticipated much that appeared in frequency-modulation systems of several later decades. His idea to link mechanically a microphone to a condenser or inductor would be replicated in most FM radiotelephony patents through the mid-1920s. And even after electronic amplification revolutionized radio circuit design after 1920, FM detectors that resembled Ehret's receiver—called "slope detectors" by then—appeared in systems well into the 1940s.[11] More impressively, Ehret was the first to modulate the length of a transmitted wave by altering the inductance or capacitance in a resonant *LC* circuit, a practice that survives today in radiotelegraphy as "frequency-shift keying" (FSK). Because marks and spaces correspond to different wavelengths, a receiving station operator can distinguish a transmitting station that has gone off the air from one that has simply halted transmission temporarily. A receiving station operator who detects a space wave that lasts several minutes can be certain that the sending station's signal has not faded away and that the transmitter's operator has stopped keying. Ehret never pointed out this advantage, though, possibly because he never realized it.

Despite its novelty, the Ehret system exemplifies how even the seemingly most innovative technological innovations draw primarily on traditional ideas. Marconi's first wireless telegraph—an invention that wrought radical changes on the world if one ever did—borrowed liberally from the decades-old practices of electrical engineering and overland telegraphy. The very ordinariness of the Ehret patents also shows how inventors lean far more toward the evolutionary than the revolutionary. Ehret worked well within the normal practice of electrical technology, using the already-venerable resonant circuit, a device that will probably continue to survive for several decades, if not centuries. Anyone familiar with the visual language of electrical engineering in 1902 would have found no basic device in the Ehret patents that had not been previously used elsewhere. Moreover, the staying power of Ehret's circuits also confirms how an innate conservatism characterizes technological innovation. Not until at least the 1950s did FM Ehret's slope detector and reactance microphone fall out of normal practice.

Ehret's FM also exemplifies the fact that historical and technological contexts shape how problems that technologies purport to solve can wax and wane in importance. What seems an urgent issue at one point can recede in significance a short time later, thus causing innovators to abandon technological paths that they had previously hoped would lead to a solution. Or, perhaps, other technologies alleviate the same problem more effectively. In 1902, when wireless communication rarely extended beyond a few miles, Ehret undertook to defeat with frequency modulation the exasperating tendency of radio signals to fade in strength. Conceptually he was on the right track. Modern FM compared with AM resists fading extraordinarily well because of its inherent insensitivity to amplitude variations, and perhaps someone might have used Ehret's ideas as a stepping-stone toward a practical system of frequency modulation. But only a few years after those patents were filed, other technological improvements, such as electronic amplification and directional antennas, made for stronger signals and mitigated fading sufficiently to cause the problem to decline in importance.

Ironically, one feature that made Ehret's inventions exceptional also accounts for their fundamental impracticability. Ehret used an *LC* circuit for a receiver because, unlike the coherer, a resonant circuit responds to changes in the wavelength of the incoming wave. But *LC* circuits detect amplitude fluctuations as well as frequency swings. In other words, his slope detector was both frequency- *and* amplitude-responsive, making it as vulnerable to fading as any other detector.

Modern practitioners will find it difficult to imagine how Ehret could have overlooked this flaw, and soon after electronic vacuum tubes became widely available during the 1920s, FM researchers found a solution. They compensated for amplitude fluctuations with an electronic circuit that automatically raised and lowered the amplitude of incoming radio waves to a fixed voltage—the very same "limiter" that Judge Palmieri in 1959 found wanting in Ehret's patents and which Palmieri mistakenly implied that Armstrong had invented for the first time in 1933.[12]

Valdemar Poulsen and Frequency-Modulation Radiotelegraphy

Historians of radio have recognized the other earliest inventor of FM, the Danish engineer Valdemar Poulsen, for his pioneering work with the arc oscillator, one of the most important devices of the early wireless era, but almost no one has mentioned that Poulsen incorporated frequency modulation into his invention.[13] The arc earned its prominence because it emitted relatively low-distortion continuous-wave radio frequencies at previously unattainable levels of wattage. An ideal continuous wave is perfectly sinusoidal, and by 1902 engineers had essentially met that standard with electromechanical alternators that delivered 50 and 60 cps electrical power. But wireless communications required at least a thousand times those frequencies, which presented the daunting problem of making an alternator spin fast enough without flying to pieces. Eventually the General Electric Company (GE) manufactured high-frequency alternators that achieved 200,000 cps.[14] But until the wide use of electronic vacuum tube oscillators during the twenties, only the arc created close-to-sinusoidal radio waves in frequency ranges above 500,000 cycles per second.

Poulsen borrowed both the arc and frequency modulation from the field of electrical music, which itself descended from the arc light of the nineteenth century. First developed by the English physicist Humphry Davy about 1808, arc lights illuminated vast areas by forcing a large continuous electrical current to flow across a gap of air separating two carbon electrodes. (The current appeared to follow a curved path, which accounted for the device's name.) The arc was characterized by not only its blinding brilliance but also an audible hiss, which, as the English physicist William Du Bois Duddell apparently realized, indicated the production of a mishmash of audio-frequency waves. In 1899 Duddell discovered that placing a condenser in the air gap circuit caused the arc to hum at a more or less constant pitch. In effect, the condenser completed a resonant circuit because arcs already contained inductive choke coils to stabilize the heavy current flow. Eventually, Duddell found a way to control the pitch precisely enough to warble "God Save the Queen" on what he called his "singing," or "musical" arc.[15] One could plausibly argue that because each musical tone corresponded to a different wavelength, Duddell invented FSK and therefore a kind of frequency modulation. But he never concerned himself with telegraphy and, moreover, his instrument oscillated below 30,000 cps, well under the minimum threshold required for electromagnetic communications.

In September 1902 Valdemar Poulsen and P. O. Pedersen improvised three modifications that dramatically elevated the device's oscillation frequency: substituting water-cooled copper "beaks" for the electrodes; burning the arc in an atmosphere of compressed hydrogen (or a hydrogen-compound gas); and placing the arc in a strong magnetic field. To be sure, not even Poulsen understood why these changes caused his arc to radiate at radio frequencies, and a residual shushing sound betrayed the arc's imperfections as a sinusoidal generator.[16] But well after the advent of vacuum tube oscillators in 1913, the arc reigned as the best high-wattage emitter of continuous radio-frequency waves.

Poulsen saw frequency modulation not as a solution to a problem but only as a loathsome expedient—a necessary evil to tolerate until he could work out a means to amplitude-modulate his invention. The arc's heavy current was the chief obstacle to this goal. Starting up and keeping it going required a vigilant human operator to maintain as constant an amperage as possible in the antenna circuit. Dips and surges from amplitude-modulating the device risked causing the arc to shift its waves to another frequency or multiple simultaneous frequencies, or even to shut down. Attempts to change the antenna current also often resulted in dangerous and destructive high-amperage "secondary arcs" across the telegraph key's open terminals. Consequently, something as simple as sending Morse code by abruptly starting and stopping the current proved exceedingly difficult with even small arcs and impossible with large ones.

To get around this problem, Poulsen essentially replicated Ehret's FSK method. Rather than modulate the transmitter wave's amplitude, Poulsen alternated its length slightly by cutting a small value of inductance or capacitance in and out of an *LC* circuit. Unlike Ehret's invention, the Poulsen arc worked splendidly, but Poulsen emphatically objected to FSK on grounds of its profligacy with radio waves. Indeed, his condemnation of frequency modulation partly accounts for his lack of recognition for inventing the method. "The process of altering the length of the emitted wave," he insisted in 1906, "must be abandoned fundamentally, since this implies that each sending station would be characterized by two waves, and thus the number of stations which can work on the same service would be reduced to one half."[17]

His aversion to FSK notwithstanding, Poulsen used the technique at several Danish arc stations—again, as a temporary measure until someone worked out how to modulate the current's amplitude.[18] Naturally, Poulsen tried his own hand at this challenge. On one occasion, he claimed to have "a good method [where] the telegraph key throws the antenna and its balancing capacity in and out of connection with the other parts of the system, in which the oscillations are allowed to pass uninterruptedly."[19] He intended with this complicated scheme to isolate the high-amperage parts of the arc from the transmitter antenna, but one can scarcely imagine how the circuit could have accomplished this without interrupting the continuous wave and pitching the oscillator into an unusable state. Not surprisingly, no evidence exists for the widespread use of Poulsen's "good method" of amplitude modulation.

Poulsen's distress from using FSK was rendered moot in 1909, when he sold the arc's American patent rights to a recent Stanford University graduate from Australia named Cyril F. Elwell, who founded the Federal Telegraph Company in San Francisco a year later. Elwell and his engineers harbored no qualms about using FSK, which they justified on pragmatic grounds. The arc's "persistency," as Elwell flatly explained, makes it "irresponsive to rapid variations of current."[20] By and large, all of Federal's transmitters used circuits that resembled Poulsen's and Ehret's: a telegraph key switched an inductance in and out of the resonant circuit, which caused the circuit to alternate between two resonant wavelengths. Federal Telegraph receivers were perhaps even simpler than Ehret's slope detector, for they detected only the longer of two transmitted wavelengths. This method amounted to binary amplitude modulation of the longer "mark" wave, because a receiver wastefully ignored the shorter wave—the "space." Lee de Forest, who worked for Federal in 1913, spun this method as a security feature, because an eavesdropper who tuned to the redundant shorter wave heard instead of the normal pattern a confusing signal of transposed marks and spaces. Noting that amateur radio operators had complained about the difficulty of copying Morse code from such a "reversed signal," de Forest quipped that "we feel responsible for [their] state of thorough disgust."[21]

The experience of Ehret, Poulsen, and the Federal Telegraph Company engineers raises two questions about technological options. First, does their use of FSK—and the fact that Ehret and Poulsen invented FSK independently of each other—prove that the method was inevitable? Or did other choices exist? The inability of Poulsen and Elwell to amplitude-modulate arcs indicate they did not. So does the failure of Reginald Fessenden, who in a long-term quest to perfect the arc attempted to forge an entirely different path. In 1893 Fessenden's friend, and one of the founders of General Electric, Elihu Thomson, patented an arc that he claimed oscillated at "ten thousand, twenty thousand, thirty thousand, fifty thousand per second, or more"—barely above the minimum threshold of radio frequencies.[22] Nine years later, Fessenden reported that he "had by his experiments verified" Thomson's claim, and indeed in 1907 Fessenden, who apparently knew nothing of Poulsen's recent work, hailed "the genius of Professor Elihu Thomson for practically every device of any importance in this art."[23] But Fessenden never tried to frequency-modulate the arc and instead strove to make amplitude-modulation arcs work. Fessenden's employer, the National Electric Signaling Company, paid for his loyalty to AM by selling no more than a few low-power arcs—all amplitude-modulated radio telegraphs—and thus the firm failed even to approach Federal's success in that field. Partly for this reason, no historian to date has ever mentioned Thomson's and Fessenden's arcs in print.

By the end of World War I, radio practitioners recognized FSK as the de facto standard for all but small systems. Elmer E. Bucher, an RCA engineer who published a widely read radio engineering textbook in 1921, rejected binary amplitude modulation out of hand as impractical for arcs, and he saw no alternative to FSK: "It is obvious," he declared, "that a telegraph key cannot be placed in series with the arc gap for signalling [with amplitude modulation] and, in consequence, the formation of the Morse characters is usually effected by changing the inductance of the antenna circuit."[24] Moreover, FSK survives today. If Poulsen were alive today, he might be astonished to discover that the method of modulation he dismissed as a wasteful workaround remains in common usage, having outlived the arc itself by nearly a century.

One could also ask why no one before 1920 took the next logical step of placing Ehret's reactive microphone modulator in an arc to make an FM radio transmitter. Possibly the arc's characteristic hiss explains this inaction, but otherwise the technology was feasible, and Reginald Fessenden even came close to making it. In 1901 he mechanically coupled a microphone to reactive components in a resonant arc circuit, which produced "a change in the frequency or the natural period of vibration"—exactly the same technique for frequency modulation that Ehret and Poulsen independently invented a year later.[25] But Fessenden used his circuit to drive indirectly an experimental amplitude-modulation transmitter, not a frequency-modulation one. Some might interpret this as failure of imagination on Fessenden's part, but in 1901 he had no reason yet to give up on amplitude-modulation radio, which he was in the early stages of developing. After all, wireless telegraphy itself was only five years old.[26]

In July 1920 Alexander Nyman, an employee of Westinghouse, applied for the only patent ever issued for an arc-based FM radiotelephone system (fig. 11). In terms of modulation, Nyman invented nothing new. He fitted a Poulsen-like arc with an Ehret-style microphone modulator. The receiver was updated, though; it was a vacuum-tube circuit that his patent tersely described as a "simple receiving station" and illustrated with a drawing of something resembling an Ehret detector of 1902, showing that the slope detector was part of normal practice by 1920.[27] Indeed, it is difficult to see what Nyman hoped to achieve when one notes that the arc was already fast approaching obsolescence because of newly available, cheaper, and more reliable vacuum-tube oscillators. Nor did he disclose any advantage of FM radiotelephony over AM. Although Nyman's electronic invention would likely have worked far better than Ehret's radiotelephone, Westinghouse engineers likely perceived no urgency to develop an alternative to AM radio. Thus, Nyman's invention had almost as negligible an effect as Ehret's, and its chief historical significance is to reveal—again—that a heavy layer of conservatism can often underlie even unorthodox ideas.

FM radiotelephony descended into a state of moribundity from 1902 until 1920, but thanks to the Poulsen arc, FM radiotelegraphy thrived as normal practice during the same period. As the widest-used continuous-wave radiator until after World War I, the FSK-modulated arc demonstrated the practicality of altering the wavelengths of radiators with virtually unlimited wattage. In the final analysis, Poulsen's and Elwell's radio*telegraphs* cleared a far wider and more direct path to FM radiotelephony than did Ehret's and Nyman's radio*telephones*. Howard Armstrong once said as much to Cyril Elwell himself. In October 1940, when Armstrong was giving a speech about modern wideband FM before the American Institute of Electrical Engineers, he recognized Elwell sitting in the audience. Armstrong pointed him out and introduced him as "one of the first users of frequency modulation in the days of the mark-and-space keyed arc transmitters."[28]

The Crystal Detector and the Rise of Amateur Radio Clubs

Nothing in the social history of radio accelerated the development of FM radio more than the invention of the crystal detector in 1906. Before that year, the expense of apparatus like the coherer made even the reception of wireless messages chiefly the province of well-funded corporations, entrepreneurs, and

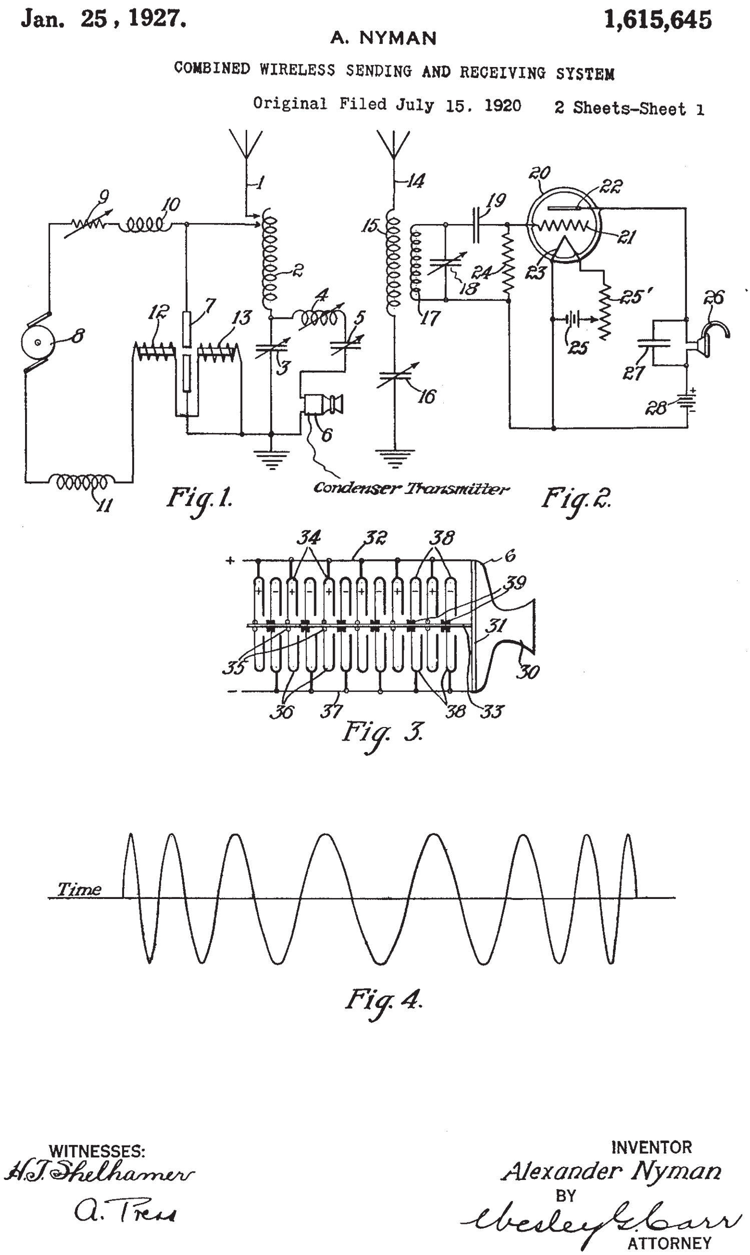


Fig. 11. Nyman FM Radiotelephone Patent, 1920. Nyman FM is based on the oscillating arc. Alexander Nyman, "Combined Wireless Sending and Receiving System," U.S. Patent No. 1,615,645, application date: 15 July 1920, issue date: 25 January 1927, assigned to Westinghouse.

governments. Afterward, the price of radio receivers plummeted, swelling the ranks of wireless practitioners with hundreds of thousands of hobbyists. Many were boys who grew up to be radio engineers; virtually every FM inventor of any importance started out as a young amateur radio operator.[29]

The crystal detector transformed the technological context of wireless communications in a matter of months. In late November and early December 1906, two Americans, Greenleaf Whittier Pickard and Henry C. Dunwoody, were independently issued patents for essentially the same device: a circuit that took advantage of the peculiar electrical properties of ordinary crystalline minerals such as silicon, galena, and carborundum.[30] Although they had not worked out the physics behind their inventions, Pickard and Dunwoody understood that crystals rectified—that is, filtered out all but the positive halves—of a high-frequency radio wave. From the rectified wave was extracted the lower-frequency components; namely the sound waves and telegraph signals superimposed on the high-frequency radio waves. Moreover, the crystal detector exhibited a sensitivity and a stability far superior to the much more expensive coherer, required no external source of power, and never wore out.

The replacement of the coherer with the crystal detector created social effects on a scale that dwarfed the technical ones. Because crystals cost about a dollar, the expense and complexity of wireless receivers declined precipitously.[31] A community of amateur radio operators—hams, they called themselves—arose and established a tradition of camaraderie and technological enthusiasm unmatched until the advent of the personal computer. That wireless fascinated so many boys, and not girls, was no coincidence, for Pickard's and Dunwoody's technological transformation meshed with a cultural shift that was already in motion. Susan Douglas has written that wireless allowed a boy to "straddle old and new definitions of masculinity." The older "primitive" ideal valued physical strength, a "commanding personality," and direct contact with nature. In contrast, Douglas says, a new masculinity emerged from the recent urbanization, corporatization, and mechanization of American society. Opportunities for outdoor experiences—especially for city youths—diminished, and traditional manly values were devalued. More and more, intelligence, education, and specialized technical knowledge were seen as opening paths to successful lives and careers. Although building and operating a radio station seldom called for outdoor or strenuous activity (except, perhaps, during the often-precarious job of erecting an antenna mast), by manipulating electromagnetic waves, hams demonstrated mastery over one of the most mysterious phenomena of nature.[32]

From the beginning, no institution shaped and spread the technology and cultural values of amateur radio more than did local, regional, and national radio clubs. Although neither the number of clubs nor their total membership is known precisely, contemporary sources indicate considerable early growth and numbers. In early 1908 the magazine *Electrician and Mechanic* founded "The Wireless Club," and by September that organization boasted chapters in 114 cities and towns in the United States and Canada.[33] A year later the energetic editor, publisher, science fiction author, and mail-order entrepreneur Hugo Gernsback created the Wireless Association of America, which by February 1913 claimed 230 affiliates.[34] In January 1910 the *Outlook* magazine estimated that more than 4,000 amateur radio operators lived in the United States.[35] Three years later, the Radio League of America counted some 350 local clubs and more than 300,000 radio amateurs in the United States altogether.[36]

Radio technology fostered especially well what has been seen as a culture of "brotherhood" or "fraternity" among radio amateurs, principally by connecting them with distant fellow hobbyists. Along with magazines, mail-order stores played a major role in expanding the worldview of many a young amateur. Because few firms before the 1920s marketed a completely assembled receiver, virtually all listeners built their own sets with mail-order parts. Hugo Gernsback, in New York City, operated the most important store in America, the Electro Importing Company, for this purpose. Boys all over America relied on Gernsback for how-to articles and radio parts, advertised in his *Modern Electrics* magazine. Harold Beverage, an engineer who participated in RCA's earliest experiments with frequency modulation during the 1920s at the company's Riverhead, New York, laboratory, credited Gernsback with his initial exposure to the field of radio. In a 1968 interview, Beverage recalled that, as a youth in rural Maine, "I got interested in a magazine called *Modern Electrics*. It was put out by one Hugo Gernsback. . . . That was quite interesting to me, fascinating, so I sent away and got a catalog from [Gernsback's] Electro Importing Company. . . . I bought [a condenser] and made my own coils. I swiped a piece of galena [crystal] from the high-school laboratory."[37]

The excitement of ham radio ruined Beverage for farming. "I used to copy a lot of news from a station on Cape Cod . . . which was sending out news to the ships at 10 o'clock at night," he recalled. "Back on the farm I thought it a lot more fun to be messing around with wireless than it would be pitching hay." Harold Peterson, a Nebraskan who partnered with Beverage at RCA during the 1920s, also credited Gernsback with his introduction to radio. *The Electro Import Catalogue*, he said in an interview, "had a nice little description of radio, how it works and what it could do. I remember reading that over and over again, and got started that way." Peterson, whose family also lived on a farm, "DX'd" stations as far away as Washington, D.C.[38] Countless radio engineers during the twentieth century could tell similar stories.

Among amateur organizations, the Manhattan-based Radio Club of America was the most closely associated with the origins of modern FM radio. The group began to take shape in 1907, when "three small boys"—George Eltz Jr., Frank King, and W. E. D. "Weddy" Stokes Jr.—met to fly model airplanes. Initially they christened their gang "The Junior Aero Club of America" and elected eleven-year-old Weddy president. Discussions at gatherings turned more and more to wireless, however, so the boys briefly called themselves "The Wireless Club of America" before permanently settling on "The Radio Club of America."[39] Today, a full century later, the Radio Club of America still exists; no other organization in the world has dedicated itself solely to radio for a longer continuous period of time.

The earliest members of the Radio Club of America by and large belonged to middle-class or relatively prosperous families, which enabled them to purchase more expensive apparatus. Armstrong, who joined about 1912, was the son of the American representative of the Oxford University Press. Weddy Stokes descended from a wealthy family of shippers, and his father was a successful race-horse breeder and entrepreneur who built and owned the deluxe Ansonia Hotel. The elder Stokes encouraged his son's interest in radio and allowed Weddy to host club meetings and to install in the Ansonia a wireless station with a 10,000-watt transmitter. Because the rig required an immense amount of power—more than all but a handful of military and commercial stations possessed at the time—keying the transmitter overloaded the hotel's in-house electrical generator, provoking guests to complain about flickering lights in their rooms.[40]

During its earliest years, the Radio Club forged three hallmark amateur radio traditions that would shape FM radio technology and the way it would be promoted: public demonstrations that showed off the social usefulness of radio, political activism, and a fellowship among amateurs that often transcended commercial interests. One of the club's first demonstrations occurred in 1913, when charter members Frank King and George Eltz constructed a small, doubtlessly low-fidelity arc radiotelephone transmitter in King's home on West 107th Street. Ten years later a reporter cited this project as "one of the first radio telephone broadcasting stations in the United States." It might also have been among the most dangerous to operate. Because the arc burned in a pressurized chamber of inflammable vapor, the boys had to synthesize a supply of gas by heating alcohol over a flame. Sometimes the "all home made, and naturally crude" apparatus spontaneously ignited. "Several amusing incidents occurred when the mixture in the arc chamber became explosive," the same reporter wrote, "and the operators were forced to beat a hasty retreat." Frank and George escaped these and other perils to broadcast phonograph records for several navy "battleships swinging at anchor a short distance away in the Hudson River."[41] Two years later, the club's Ansonia Hotel station relayed more than a thousand telegraph messages for the navy during another port of call. The club pulled off its grandest demonstration in 1921, when its members exchanged the first transatlantic messages using short waves. Howard Armstrong was one of five stateside radio operators in this experiment, and his friend Paul Godley, who would help publicize the Armstrong system of high-fidelity FM during the 1930s, operated the club's station in Scotland.[42]

The Radio Club also excelled in political activism. Indeed, the amateur movement as a whole began largely as a populist uprising. Gernsback founded the Wireless Association of America in 1909 "in order to guard against unfair legislation as far as the wireless amateur was concerned."[43] In 1910 Gernsback marshaled the collective force of amateurs to resist the first proposed legislation that was harmful to the interests of hams. "The association had no sooner become a national body," he reminded his readers in 1913,

> than the first wireless bill made its appearance. It was the famous Roberts Bill, put up by the since defunct wireless "trust." The writer [Gernsback] single handedly, fought this bill, tooth and nail. He had representatives in Washington, and was the direct cause of having some 8,000 wireless amateurs send protesting letters and telegrams to their congressmen in Washington. The writer's Editorial which inspired the thousands of amateurs, appeared in the January, 1910, issue of *Modern Electrics*. It was the only Editorial during this time that fought the Roberts Bill. No other electrical periodical seemed to care a whoop whether the amateur should be muzzled or not. If the Roberts Bill had become a law there would be no wireless amateurs to-day.[44]

No individual local club defended the interests of hams more than the Radio Club of America did. In April 1910 George Eltz, Frank King, Weddy Stokes, and Ernest Amy traveled to Washington to lobby against another piece of proposed legislation known as the Depew Bill, which if enacted, the boys feared, would prohibit amateur transmissions. Fourteen-year-old Weddy testified before the Senate Commerce Sub-Committee, delivering, by one account, an appealing call to alarm. "Clad in knickerbockers," the *New York Herald* reported, "he captured the hearts if not the judgment of the Senators." Weddy denounced the Depew Bill as a "stock-jobbing scheme," and warned that "soon a great trust will be organized to corner the very air we breathe."[45] Two years later the Radio Club dispatched a second delegation to Congress to lobby against a similar bill.[46] In 1922 the club's president, Howard Armstrong, represented American amateur radio operators at the first of four annual National Radio Conferences, hosted by Secretary of Commerce Herbert Hoover. These meetings marked Armstrong's debut as a "public engineer" whom government officials consulted about radio-related issues. During the remainder of his life he testified several times as an expert witness before congressional committees and the FCC, most often during the 1930s and 1940s as an advocate of wideband FM radio.

One of the greatest strengths of the Radio Club of America was its tradition of welcoming all radio practitioners, both amateur and professional. Club policy encouraged members to share technical and other information, a practice that blurred proprietary boundaries and facilitated transfers of knowledge. This was not a unique role for the club; in 1913 the all-professional Institute of Radio Engineers was established for similar purposes, and the two groups shared many members. But no other organization merged so seamlessly the camaraderie of ham radio with a professional-like seriousness of purpose that sometimes led to first-rate research. Admittedly, papers that appeared in the *Proceedings of the Radio Club of America* as a whole lacked the theoretical rigor one could expect to find in the *Proceedings of the IRE*, but the Radio Club possessed an unrivaled atmosphere of "fellowship." As *Radio Broadcast* magazine explained in 1923,

> A club is a place for good fellowship, true; and that describes the Radio Club of America, which has already stimulated good fellowship in radio and more specifically among its members. In that sense, the word stands.
>
> But in the case of this group of young men, there has been something more than a club atmosphere. With the serious intentions of its members, the thoroughness of the papers and discussions marking its meetings, and the scientific value of its experiments and tests, the word "club" is almost a misnomer. This organization might well call itself a scientific society, although it does retain that spirit of fellowship which goes with the usual meaning of club.[47]

By the mid-1920s, according to club historian George Burghard in 1934, most of "the original small boys had grown to be full fledged men of affairs" in the radio industry. "Naturally," Burghard explained, "the character of the membership of the club as-well as that of the papers, underwent a similar change. The club had now all the earmarks of a genuine scientific body. The spirit of the organization, however, never changed. These men, now engineers, executives or scientists were still amateurs at heart."[48] Such a description fit Armstrong perfectly. In his entire life, he held only two salaried jobs: his military service during World War I, and a dollar-a-year research professorship, which Columbia University gave him during the late 1920s.

Traditions of technological expertise, openness, and friendship among practitioners largely accounted not only for how widely word about experiments with FM radio spread through the radio industry during the 1920s but also for why both amateurs and professionals participated in developing, testing, and promoting Armstrong's wideband FM during the 1930s. Five friends of Howard Armstrong—Tom Styles, Jack Shaughnessy, Paul Godley, George Burghard, and especially Carman Runyon Jr.—all long-standing members of the Radio Club—helped test and sell the Armstrong system to the public, sometimes for no remuneration. Styles worked full time as Armstrong's financial manager and secretary, Shaughnessy as his assistant; in 1936 Godley published the first lengthy article explaining Armstrong FM to the broadcasting industry; and in 1934 Armstrong installed a prototype receiver in Burghard's Westhampton Beach, Long Island, home during RCA's initial field tests of the Armstrong system.[49] Runyon, a founding member who managed a coal-delivery company in Yonkers full time, witnessed Armstrong's secret work with FM as early as 1932 and for several years afterward routinely took part in public demonstrations of broadcast FM radio.[50] Other members involved themselves with experimental broadcast FM, sometimes in connection with Armstrong's RCA trials, sometimes with the FM work of other companies, sometimes as pioneer FM broadcasters. Harold Beverage, who pulled strings to obtain permission for Armstrong to test his FM system in RCA's Empire State Building transmitter in 1934, joined both the Radio Club and the Institute of Radio Engineers soon after World War I. Harry Sadenwater, an RCA Manufacturing Company engineer, permitted Armstrong to install prototype receivers in the basement bar of his home in Haddonfield, New Jersey, in 1934. John V. L. Hogan, a veteran wireless pioneer who had worked for both Fessenden and de Forest, founded WQXR, the first broadcast FM radio station in Manhattan, in 1939. Albert Goldsmith, who was also a former president of the IRE, evaluated Armstrong's FM system for RCA in 1934. Beverage's colleague at RCA Communications Company, Murray Crosby, patented several FM-related inventions before Armstrong did, published important technical papers about frequency modulation, and attended Radio Club meetings as a guest from time to time. One man, Howard Armstrong, patented wideband FM in 1933, but a community of both professional and amateur practitioners also contributed to the development of the technology.

Such was the technological and social context of radio from 1902 until the early 1920s, a context out of which emerged both AM and FM broadcast radio. Certainly, frequency modulation was nothing new by the time Armstrong's patents were issued in 1933; FM thrived in radiotelegraphy, and minimal imagination was required to begin adapting the method for sound. Even more important to the progress of FM was the culture surrounding radio technology that formed years before 1920, for the final shape of radio owed much to a radio engineering profession that an amateur tradition deeply influenced.