A History of Some Foundations of Modern Radio-Electronic Technology*

J. H. HAMMOND, JR.†, SENIOR MEMBER, IRE, AND E. S. PURINGTON†, SENIOR MEMBER, IRE

Summary—**Experiments with high-speed boats at Gloucester, Mass, in 1910–12, resulted in sustained governmental interest in radiodynamic torpedoes from 1912 to 1931. This culminated in the successful radio control of the course of a modified standard naval torpedo throughout a 9000 yard run at an operating depth of 12 feet. Many developments by the Hammond Laboratory established basic principles used in modern airborne guided missiles, including the stabilizing, security, bat, proximity, and homing principles. Needs for improved devices and circuitries for control purposes stimulated the early development of principles used in modern communications. The Hammond Laboratory sponsored the first nondetector applications of the filamentary type triode for linear amplification and for transmitter and heterodyne purposes in 1911–12. The intermediate frequency principle for selectivity purposes was developed in 1912, stemming from the heterodyne principle of Fessenden and the two-channel security principle of Tesla. At a conference at Gloucester in October, 1912, disclosure of these developments led immediately to the Alexanderson development of the tuned radio frequency principle and to the accelerated development of the high-vacuum triode by Langmuir and White. The first military application of the intermediate-frequency principle was to the solution of a World War I problem of mitigating the interference of enemy spark type transmitters upon communications from front line infantry. The receiver of the system was structurally of the most general superheterodyne type for continuous-wave reception. With the advent of broadcasting, the early 1912 conference had paved the way for the commercialization of the modern receiver with its IF basis for selectivity, its trf basis for sensitivity, and its hard tube basis for reliable practicability. Additionally, the unicontrol feature for operational simplicity of the superheterodyne stemmed from military designs of 1917–18. In the early record, there are suggestions in the fields of automatic volume control, remote cutoff triodes, and feedback for improving signal fidelity. Special early applications of the IF principle were in the fields of multiplex signaling and of privacy systems for radio telegraphy and telephony. In the field of frequency modulation, the principle that fm and AM transmissions may be independently in the same wave band was established in 1912. The technique of gathering a complete signal element from both ends of a transmitted fm spectrum was practiced in a wide-band single-shot system of security control in 1914. Researches in 1921, established the all-electronic method of fm transmission for telephony; in 1922, led to the development of a wide-band noise-reduction system for telephony having fm properties and using dual-channel noise-cancelling dual-detection reception, in 1927, provided for the possible development of wide-band artificial fm transmission with a stabilized carrier, developed from quasi-phase modulation produced from AM by a 90° carrier phase shift.**

I. Introduction

DURING the period of the founding of The Institute of Radio Engineers, the radio-electronic art was in a transitional stage. This art had its beginnings with the utilization of the Edison effect in the Fleming diode detector, and a highly important advance was the invention of the de Forest triode which provided an electrically controlled impedance. But radio receivers for spark type transmitters usually comprised a coupled-circuit tuner with a simple detector and headphones; sometimes the detector was an audion triode which, because of the operating cost, was used only when the simpler devices were just insufficiently sensitive. Continuous wave receivers for arc and rotary type wave generators were of similar nature but with a device before the detector to produce a local modulation of the incoming signal; this was usually a rotary tikker to chop the continuous wave whenever present or an arc driven heterodyne to create a tonally modulated signal.

The Hammond Laboratory was privately and personally organized during this period with the immediate objective of developing radio and other remote control systems for waterborne missiles. The laboratory, Fig. 1,

Fig. 1—The Hammond Laboratory, 1912–1928.

was located on the westerly shore of the outer harbor at Gloucester, Mass., at the entrance to Freshwater Cove. Here from 1912 to 1928, many electronic and allied devices and circuitries, developed at first for radio control purposes, formed the basis of important developments in communications and other fields of application. Since most of these developments were with military and naval applications in mind, publication of technical and historical information was highly limited by governmental and self-imposed restrictions.

The purpose of the present paper, therefore, is to make available a better understanding of the background of some of the most important developments of radio-related electronics. For clarity, the report is not on an integrated chronological basis, but rather discusses different aspects in turn. After a discussion of

* Original manuscript received by the IRE, February 11, 1957; revised manuscript received, June 5, 1957.
† Hammond Research Corp., Gloucester, Mass.

radiodynamic torpedoes, attention is given to the triode tube and its fundamental nondetector applications, to the beginnings of the modern intermediate-frequency type transmitter and receiver systems, and to the development of frequency modulation and related transmitter and receiver circuitries.

II. The Radiodynamic Torpedo

A radiodynamic torpedo is a carrier of high explosives for a military target, the course of which is alterable dynamically by radiant energy from a distance. The radiant energy may be of any form: compressional wave, light beam, or from a radio antenna. The torpedo may have the form of a surface boat, an underwater craft, or an airborne guided missile. The history of such torpedoes began in 1897, when E. Wilson[1] and C. J. Evans controlled slowly moving boats by radio waves on the Thames river. Almost simultaneously, N. Tesla then became the pioneer in the United States[2] and in 1898, built a working model which he successfully demonstrated at the Auditorium in Chicago. In 1905, the Quintard Iron Works in Boston actually built four radio-controlled torpedoes according to designs by H. Shoemaker. These were hybrid devices, having the form of an underwater naval torpedo rigidly suspended from a surface float which supported the receiving antenna. In the United States, these were not considered to have military merit, and they were sold to the Japanese Navy. Preliminary work by Hammond[3] from 1910 to 1912, with the "Pioneer" and the "Radio" in Gloucester Harbor resulted in the development of the automatic course stabilization principle for high-speed waterborne and airborne torpedoes. This and contributions to the security of control firmly established U. S. governmental interest leading to the modern developments.

The U. S. Army had long since expressed a desire for a coast defense torpedo controllable from the shore and had experimented with a cable linkage from the control point to the torpedo. In the late fall of 1912, Brig. General E. M. Weaver and Col. R. P. Davis, the Chief and the Assistant Chief of Coast Artillery, came to Gloucester and witnessed the successful radio control of the 33-knot "Radio."[4] In May, 1913, General Weaver detailed Capt. F. J. Behr[5] to the Hammond Laboratory as an observer, with twelve technical sergeants to assist with 60-inch searchlights and other Coast Artillery materiel. After he had assisted in establishing the direction of the developments along desirable military lines, Capt. Behr was relieved by Lieut. S. M. Decker.[6]

General Weaver believed that high-speed motorboats with depth charges could constitute a serious menace when directed against enemy ships. That he was right was proven in World War II when the Italians sent in their high-speed explosive boats at Suda Bay in Crete, on March 25, 1941, and sank *H.M.S. York* and other British ships. And during the Philippine actions at Lingayen, the Japanese on January 10, 1945, attacked our LCI's and LST's by manned explosive motor boats with great effect. At Kwajalein, the U. S. Navy improvised radio-controlled boats called "Stingray" to attack Japanese positions. Better results were obtained by the American forces with radio-controlled drone boats against the Germans on the Riviera.

History has therefore vindicated the early judgment of the Coast Artillery and its chief. However, as is the case with most radically new weapons, the remote control torpedo was to encounter the usual diversities of opinions which are enjoyed by the different military services.

The principles of modern radio missile guidance developed in the total pioneering period of 1910 to 1914, were as follows.

Automatic Course Stabilization

In the absence of a control signal, the torpedo should be stabilized as to course by automatic mechanisms within itself. Course stabilization had been practiced in naval torpedoes by a gyroscope energized only at the start of a run. But in 1912, the Sperry Gyroscope Company and the U. S. Navy were developing a precise and reliable motor-driven gyrocompass with remote repeaters. Accepting this as a best possible basis,[7] the Hammond Laboratory engineered the modification of one of these devices so that the repeater controlled not a compass indicator, but the operation of a steering engine. The first navigational application of this automatic pilot principle was to the third boat, the *Natalia* of Fig. 2; the system was first put into long period operation on March 25, 1914, throughout a 60-mile run from Gloucester to Boston and return. Here, of course, the gyro setting was changed manually from time to time.

Radio Control of Gyro Setting

Directional changes of the missile should be accomplished by applying control signals to alter the course setting of the automatic stabilization device. The gyroscope for the *Natalia*, therefore, was fitted with electromagnetic solenoids to step the course setting backwards or forwards by a definite number of compass points in response to each control signal. Differentiation as to sense was by the timing pattern of a few dashes applied to a

[1] U. S. Patent 663,400 (1898–1900) to Wilson and Evans. Note: Filing and issue years are both given.

[2] U. S. Patent 613,809 (1898–1898) to N. Tesla.

[3] The early radio-control developments were popularly described by C. Moffett, *McClure's Mag.*, vol. 42, pp. 27–39; March, 1914, and vol. 44, pp. 21–31, March, 1915.

[4] "Fortifications Appropriation Bill, 1917," Hearings conducted by the Subcommittee (Swager Sherley, Chairman) of the Committee on Appropriations, House of Representatives, in charge of the Fortifications Appropriation Bill, pp. 288–298, 301–304; January 24 to February 10, 1916.

[5] *Ibid.*, pp. 306–314.

[6] *Ibid.*, pp. 314–324.

[7] U. S. Patents 1,418,788–789–790 (all 1913–1922) to Hammond. The original sketches were made during a transatlantic crossing on the "*President Grant*," on July 21, 1912.

Fig. 2—The *Natalia* leaving pier for tests, 1914.

primary control relay in response to suitably transmitted signals from one or from two antennas located somewhat southerly from the laboratory building.[8] Other dash patterns controlled various other operations, such as engine speed, searchlight shutter opening, and laying of mines.

Security of Radio Control

The control system should provide security both against control and against interference upon control by such countermeasures as are currently available to a potential enemy. This principle is of special radio-electronic interest and involves the application of several subprinciples.

First, the gyro radio principle of control contributed greatly to security since it reduced the number of control signals required during a run and made the analysis of the signal pattern much more difficult. Second, the nature of the control signals could be changed[9] even during a run, and irrelevant signals introduced, so that the proper control officer would have a far superior statistical advantage in any radiant energy battle for guidance. Third, after the missile had received its main course corrections, the gyro setting could be locked against all countermeasures, providing conventional stabilized course control during the final approach to the target. Fourth, as an alternative to locking the course, the missile at the end of the run could be converted to a target-seeking type to home upon the countermeasure interference, still with gyro stabilization. Fifth, undesired control could be minimized by adoption of a suitable timing pattern for different frequency radiations, obviating ready analysis and duplicative synthesis.

The remote control system used mainly with the *Natalia* was of a dual channel type. Tesla had already proposed a security system based upon the coincidental transmission on two channels: a forerunner of the "and" principle[10] of modern computers. Two receivers with rectifying detectors operated two corresponding output relays, the simultaneous closure of which operated a third relay for actuating the control pattern analyzing and distributing system. The Hammond system for the *Natalia*[11] transmitted dual channels not simultaneously but in quick sequence from the same transmitter. Thus, to produce a control dash in the receiver, a Poulsen arc transmitter[12] was made to send a first half-dash on one frequency and the second on another greatly different frequency. Control response was only after reception of both ends of the spectrum of this single-shot frequency modulated wave. The receiver was tuned normally to the frequency of the first half-dash; reception of this part of the spectrum retuned the receiver to the frequency for the second half-dash, with the first half-dash held in storage. If the second half-dash was received within a given time limit, coincidence between the second and the stored first parts registered a complete dash, for control operation. The system restored to normal, regardless of whether the first half was signal or interference, to await further signals. This was the first example of security systems using both time and frequency diversity.

The target-seeking feature of the *Natalia* provided for homing both upon an enemy searchlight and upon enemy interference. Using the principle of the "Electric Dog," Fig. 3, which would follow a moving light, two selenium cells mounted on the foremast operated differentially to alter the gyro setting until the boat was headed toward the controling light source. Radio homing was by the same general principle,[13] using crossed loops both tuned to the wave length necessary for interference. Fig. 4 shows schematically the essentials of this guidance system, practiced with special duo-triode tubes to improve the permanence of differential balance. Technical developments, of course, led to better application of the differential output to the gyro system and to satisfactory freedom from course overshoot and

[8] The location of these antennas at Point Radio is marked by a bronze plaque on the grounds of the present estate at 160 Western Ave., Gloucester, Mass.

[9] U. S. Patent 1,420,257 (1910–1922) to Hammond.

[10] U. S. Patents 723,188 and 725,605 (both 1900–1903) to N. Tesla.

[11] U. S. Patent 1,486,885 (1915–1924) to Hammond.

[12] Poulsen arc equipment from the Federal Telegraph Co. was installed with the personal attention of L. de Dorest in 1913.

[13] U. S. Patents 1,370,688 (1914–21); 1,467,154 (1912–23); 1,513,108 (1914–24); re: 16,181 (1914–25) to Hammond.

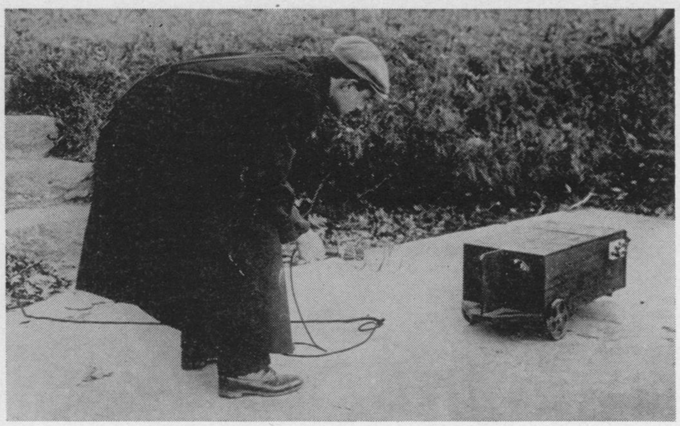

Fig. 3—The "Electric Dog" with light-controlled guidance, 1912

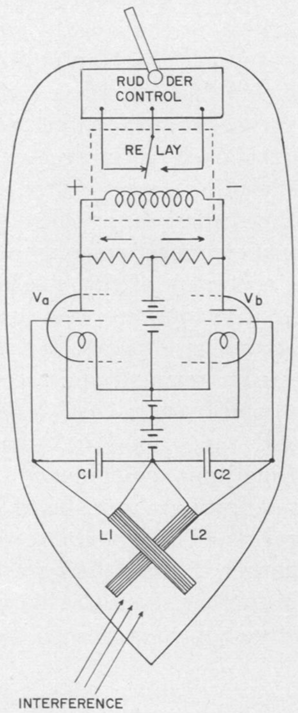

Fig. 4—Radio homing guidance system, 1914.

hunting. Homing by the differential control from two and even more channels has many modern military applications.

Practical interference tests were made with the cooperation of the *U.S.S. Dolphin*, fitted with the then latest types of communication equipment. These were made in Gloucester Harbor, October 6, 1914, during which the *Natalia* was controlled by the single-shot fm system using 118 and 1000-kc frequencies. The *Dolphin*, using in turn signals of 1000, 600, 400, and 300 kc, was at no time able to exercise control and could not block control by paralysis of the *Natalia*'s rectifying detectors until she had been worked to within 250 feet of the target. Thereupon, the *Natalia* could have been operated with the gyro locked for the final fixed course, or could have been converted automatically to a target-seeking torpedo under guidance from the interference.

After demonstrations to the Government on November 16 and 24, 1914, it was officially considered the work was through its experimental stage and steps were taken toward the development of a service weapon. The original plans were to establish the first radiodynamic torpedo base at Fisher's Island off New London, Conn.; this was submitted to Congress in 1915, too late for action. Extensive hearings[14] were held on the broad subject of "Radiodynamic Torpedoes," and on March 23, 1916, Hammond submitted a "Proposal Z" to the War Department. This in turn was presented to the 64th Congress, which[15] on July 6, 1916, provided for a special executive board of three Army and three Navy officers to be appointed by the President with a provisional appropriation of $1,167,000. Subject to the recommendation of the Board after a satisfactory demonstration of radiodynamic control of a torpedo, the main part of this fund was to be used for the development and procurement of a complete sample service equipment and for the acquisition of complete rights then expressed by patents and applications placed in the secret archives of the Patent Office, not open to disclosure even in cases of interference.

During its existence as a continuing body, the Torpedo Board made various changes in the requirements defining the nature of the demonstration necessary to result in a favorable recommendation. Initially, the Army members had one concept which was that of the Chief of Coast Artillery; the Navy members had another which was that of the Chief of Naval Operations. In the Congressional hearings, special interest had been expressed in the use of an airplane for controling the torpedo far from the shore as a weapon of offense. The pilot of the plane could simultaneously observe the course of the torpedo toward its target and correct the course by radio signals preferably sent directly from the plane to the torpedo. Proceeding on this basis with preliminary work in Gloucester Harbor, official demonstrations were made to the Board on August 23, 1918, using the fourth and final surface boat, the *H-4*, shown in Fig. 5. The control was exercised directly from a plane at altitudes up to 5000 feet and at distances up to 5 miles, with the *H-4* moving through the intricate wartime shipping of Hampton Roads off Norfolk, and Fortress Monroe, Va. Since it had appeared that the final service weapon might be a submarine-like torpedo to which radio waves might not be able to penetrate, demonstrations were also successfully made of control by underwater radiosonic compressional waves. For the possible nontorpedo use of radiodynamic control, demonstrations were made with the *H-4* functioning as an unattended remotely controlled mine layer.

But in 1919, further experiments were conducted by the Navy at New London, with the cooperation of the

[14] "Fortifications Appropriation Bill, 1917," *op. cit.*, pp. 258–357.
[15] Congressional Record, 64th Congress, First Session, pp. 10873–876, 10904–907, 11667–671, 11784–792; June 13–30, 1916, Acts of Congress, July 6, 1916; March 3, 1919; June 30, 1922.

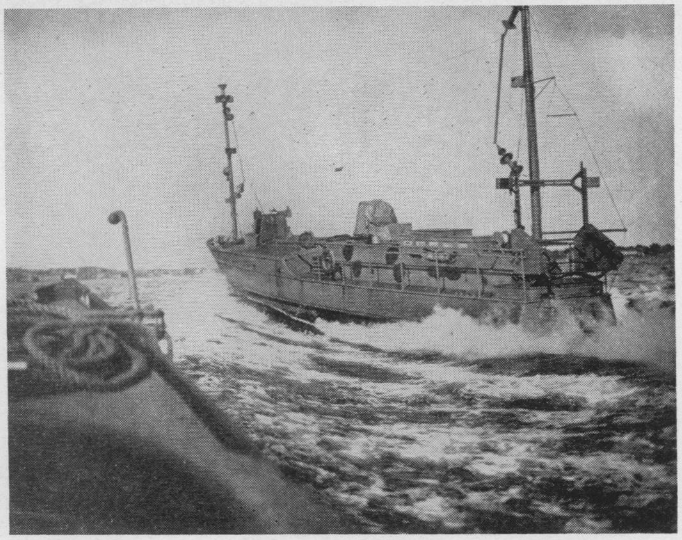

Fig. 5—The *H-4* under radio control from a plane, 1917.

Bureau of Standards, the Hammond Laboratory, and the Marconi Wireless Telegraph Company of America. The objective was to determine the essential facts of radio transmission to underwater craft. New techniques resulted as to reception on large conventional submarines, using the Bureau of Standards loop arrangement of Lowell and Willoughby with a large cross-sectional area to provide strong pickup from the magnetic component of the underwater radio field. But for small torpedo-like missiles, the antenna was required to be of the trailing wire type to minimize loss of speed and therefore of range. This antenna was to be a flexible rubber-covered wire with the trailing end sealed by a vulcanized and leakproof rubber tip and with the receiver end passing through the steel shell of the missile. Tests with a submerged submarine with such an antenna 300 feet long supported at a depth of 6 feet proved that radio control of an underwater torpedo was entirely practical with signals from a plane with 85 w continuous wave output at 188-kc frequency. It was therefore decided by the Board that the proper service weapon to meet the military requirements should be a standard naval torpedo[16] modified by the addition of a midsection to house the radio control gear.

This decision to discontinue development of the surface type radiodynamic torpedo was undoubtedly made in the belief that an underwater missile of superior military value could be readily developed. Therefore, interest in the surface type lay dormant for twenty years until reawakened by the exigencies of World War II.

In December, 1921, the Board further established the quantitative requirement that the radio control of the modified naval torpedo would be considered satisfactory if, throughout a run of 9000 yards, the torpedo was controlled at a depth of twelve feet. This increase of 6 feet over the test depth in the New London experiments was for making certain that the torpedo would be able to strike a battleship below the armor belt. After preliminary studies by the Hammond Laboratory with Navy cooperation, the Navy actively pressed the new phase of missile development when research funds appropriated by the 64th Congress became reavailable in 1925. The work was then transferred to the waters of Narragansett Bay under the direct supervision of the Naval Torpedo Station at Newport, R. I.

The problem of radio control under the stipulated conditions even from the standpoint of signal transmission was highly complex; it is along the same general lines as that of delivering energy purposely to an insulated dipole embedded in the walls of a modern waveguide. Ocean water has a conductivity of about 4.5 mho/m, a permeability the same as air, but a dielectric constant about 80 times that of air. The operating frequency must be high enough to permit a strong field to be established with E and H components parallel to the water surface at the boundary, but low enough to prevent excessive attenuation of the downwardly penetrating energy due to the conduction currents. The optimum frequency depends upon the depth at which the signal must be received, as well as the current status of the radio-electronic art; the radio band 150 to 200 kc was chosen for the purpose. At midband, the attenuation of ocean water computes to be 4.65 db per foot of depth, so that the extra 6 feet in the stipulated depth corresponded to a power ratio of about 600. Practical operation was achievable only by an increase of transmitter power in combination with improved design of the receiver system.

Then too, the field establishable parallel to the air-ocean boundary for the start of the downward flow of energy is itself small in comparison with that available for the antenna of a surface craft. Ocean water has an intrinsic impedance of about (0.001) $(1+j)$ times that of air at 175 kc operating frequency; therefore, there is about 24 db of mismatch loss for a wave directed vertically downward from a plane, just in passing from the air to the water medium. There is even greater difficulty if the energy is originally propagated horizontally from a control ship or a shore station with the H field tangential but the E field in the air medium nearly perpendicular to the air-water boundary. By the theory established by Zenneck,[17] with a counterpart in modern waveguide theory, the E field tangential to the surface for the start of the downward wave computes to be at 175 kc, about (0.001) $(1+j)$ times the E field of the air wave flowing parallel to the surface. This factor is the same as the water-to-air impedance ratio, since at the boundary the H field pertains to both the air and the

[16] Torpedo mechanisms were popularly described by E. F. Chandler, "The modern automobile torpedo," *Sci. Amer.*, vol. 113, p. 112; August 7, 1915.

[17] The Zenneck theory is available in J. A. Fleming, "The Principles of Electric Wave Telegraphy and Telephony," Longmans, Green & Co., London, Eng., 2nd ed., pp. 729–744; 1910.

water wave. Therefore, the power flux for the downward wave even at the surface is 30 db less than that of the air wave above the surface available for antennas of surface craft. There is in effect a 30-db loss of signal energy as the wave makes a right angled turn to enter the water. Since the *H* field does not suffer this turning loss, it is probable that a loop antenna, if practicable, would not have the same disadvantage as a trailing wire antenna in being converted for use in the ocean.

Fig. 6, derived from a Navy release in 1930, shows the "Hammond Radio Controlled Torpedo" just after ejection from a standard tube mounted on a test barge. The addition of the radio section, bracketed by white bands, was not detrimental to the proper launching of the torpedo. When the torpedo hit the water, the filament circuits of the receiver were closed by an inertia type switch, and a reel carrying the antenna was forced from its mount behind the propellors. This reel floated to the surface after the antenna had unwound and trailed to its full operating length of 150 feet. The gyro was given a preliminary setting before launching, and the course of travel was changeable by a specific number of compass points in response to each control signal from a transmitter located at Melville, R. I. This transmitter[18] was a General Electric continuous-wave type with 10-kw output rating in the 150–200 kc band, but modified as to its output circuits to provide efficient generation of two waves with as much as 10-kc separation. Control was in accordance with the time-frequency pattern fed into the power amplifier from a switchable master oscillator. Reliable operation was achieved with only occasional lapses when during a run the line from the transmitter to the torpedo antenna was at right angles to the line of torpedo travel. Final tests, meeting the requirements of the Board, were made in the early winter of 1930–31. Thereupon, by the "Hammond Patents Purchase Agreement of July 30, 1932," designation NOd 393, the Government acquired rights for radiodynamic purposes in over a hundred of the Hammond patents, including many of importance not covered by the original listing in Proposal *Z*.

It is believed that underwater radio-controlled torpedoes were not used in World War II; perhaps this was because of difficulties with production torpedoes in the earlier part of the war. Magnetic detonators and contact exploders were unreliable, and the depth regulation was undependable. It was not until 1944, that these difficulties were fully overcome. In the latter part of the war, torpedoes were set mostly for a 6-foot running depth to make certain that they would not underrun important targets. Under such conditions, as proven by the New London tests of 1919 and the Newport developments of 1925–31, the radio control of underwater torpedoes from a plane at a safe distance from the target would have been highly practicable.

Fig. 6—The "Hammond radio-controlled torpedo," 1930.

An important application of the radio control of surface craft was in the contests between bombing planes and battleships. The *U.S.S. Iowa*, Fig. 7 (opposite), scheduled for scrapping in accordance with international agreements, was the first to be fitted out by the Navy as a target ship, with the General Electric Company and the Hammond Laboratory cooperating. In tests of June 29, 1921, at sea about fifty miles off the outer Maryland shore, eighty dummy bombs made of concrete were dropped in 200 minutes from 20 planes at 5000 feet elevation. It is probable that the number of hits did not exceed two near misses, and that the otherwise defenseless target ship, even with reduced speed, escaped certain theoretical destruction by the zigzag course set up by radio control from the *U.S.S. Ohio* at a safe five-mile distance.

Hammond Laboratory patents and electronic developments in the general field of remote guidance, restricted neither to the water medium nor to radio links, have been made available. Early patents necessary for application of the stabilization principle in radio control of aircraft[19] were made available in 1925, to the Sperry Company for governmental developments. The government was given free use of patents and ideas with important applications in World War II and later developments of airborne missiles involving radio-electronics, although worked out originally for sound control in the water medium. Thus the Hammond statement[20] of the "Bat" principle is as follows: "In combination with a self-propelled body, means for automatically steering said body on a predetermined course,

[18] After the completion of the Newport work in the field of radiodynamic torpedoes, the three phase power rectifier of this transmitter was incorporated into the Harvard University cyclotron, later sent to Los Alamos for research, leading to the development of warheads for guided missiles.

[19] U. S. Patents 1,568,972 (1914–26); 1,568,974 (1915–26), 1,625,252 (1919–27) to Hammond and 1,772,343 (1920–30) to Dorsey and Trenor. See also U. S. Patent 1,568,973 (1915–26) at that time, in effect, optioned to the U. S. Government.

[20] U. S. Patent 1,892,431 (1928–32) to Hammond; claim 28.

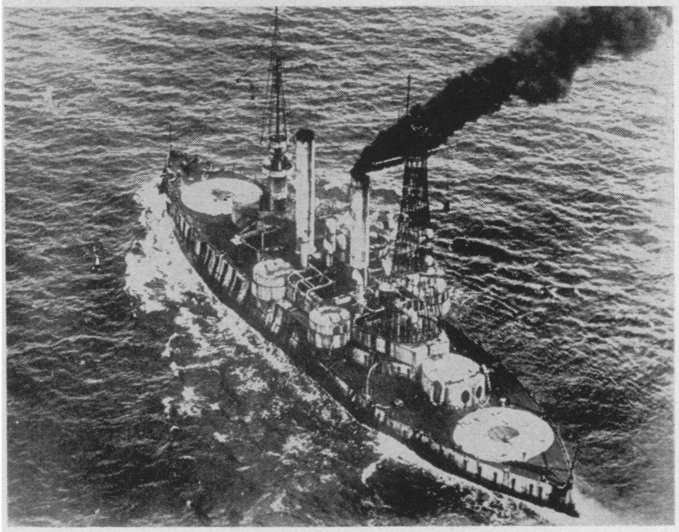

Fig. 7—The *USS Iowa* under radio control, 1921.

means carried by said body for propagating energy waves, means associated therewith for receiving said waves when reflected from a distant object and means operable in response to said reflected waves for altering the course of said body." And the statement of the "Proximity" principle[21] is: "In a torpedo having a warhead, means for producing and radiating energy, means for receiving said energy after reflection from an outside object, and means responsive to the received energy to detonate said warhead." The problems of the radio control of rockets[22] were given inventive thought. Special research was conducted[23] in the field of security control, especially for the design of radio-guided drop bombs with partially publishable results.[24] And finally, a special report upon security design principles was made to the governmental agencies concerned with advanced developments to make certain that the results of the Hammond Laboratory studies in this field were fully available.

III. The Triode Tube

The Institute of Radio Engineers was created by merging the interests of two competing societies, largely through the efforts of Robert H. Marriott, Alfred N. Goldsmith, and John V. L. Hogan, Jr. Marriott was serving as the president of the Wireless Institute which was the junior society of the merger. The president of the senior Society of Wireless Telegraph Engineers was Fritz Lowenstein. Marriott and Lowenstein became the president and the vice-president, respectively, of the newly formed Institute.

In an historical paper,[25] Marriott has indicated that the "Audion" three-electrode form of detector "was used to some extent as early as 1906, but apparently in very small numbers until about 1912 when the amateurs became active in its use. . . ." In early 1912, Lee de Forest indicated[26] that Dr. L. W. Austin of the Navy had "found the sensitiveness of the Audion 1.5 that of the electrolytic detector, which ranked more sensitive than any other including the magnetic detector and the crystals." The nature of the potential figure of merit of the de Forest triode is establishable from curves given by J. H. Morecroft[27] for a tube that was later well evacuated and baked to remove the gaseous irregularities. For a central working condition of 45 v on the anode with 0.8 ma current at 4 v grid bias, the amplification constant was about $\mu = 2.3$ and the internal impedance was about $r_p = 36{,}000$ ohms. Commercially, the de Forest audion detector circuit in a cabinet with three spare bulbs retailed at $125.00; replacement bulbs were $5.00 each with a rated but unguaranteed life expectancy of 40 to 60 days.

Thus it would appear that at the time of the founding of the Institute of Radio Engineers, de Forest triodes were applicable only in the limited field of receiver circuitry and in the even more limited field of detectors to cause amplitude variations of radio signals to be revealed audibly. After five years of public availability, there were no other established applications in the electronic art, nor had the wire communication or other nonradio electrical companies in the United States initiated researches for the purposes of developing the potentialities.

In the preliminary work in radio control, the first relays were operated from the output of conventional detectors, with dc rectified outputs of the order of 0.1 ma. Lowenstein, as a power engineer, had been interested in the development of mercury vapor and other ionic devices for power applications. He had recognized[28] that while transmitters with Duddell and Poulsen arcs depended for operation upon the negative resistance characteristics of two-terminal devices, similar results might be possible with positive resistance devices through magnetic or electrical control of the current stream. As a consultant of the Hammond Laboratory, Lowenstein on May 11, 1911, undertook in New York the development of the three element "ion controller" for relay-operating rectifier-detector purposes, and also for the nondetector purposes more obviously related to amplification. Starting with power applications with a motor generator of 1000 v, $\frac{1}{2}$ kw dc rating shortly thereafter replaced by a battery, Lowenstein reported by a letter of September 19, 1911: "So far the experiments . . . were without results, but . . . will advise you immediately upon obtaining the repetition of experiments which I made some years ago. The probable cause of

[21] U. S. Patent 2,060,198 (1932–36) to Hammond; claim 18.
[22] U. S. Patent 2,413,621 (1944–46) to Hammond.
[23] Partly under NDRC-OSRD Contract OEMSR-694.
[24] U. S. Patents 2,424,900 (1944–47); 2,449,819 (1944–48); 2,465,925 (1944–49); 2,480,338 (1944–49); 2,510,139 (1944–50); 2,522,893 (1945–50); 2,635,228 (1948–53) to Purington.
[25] R. H. Marriott, "United States radio development," Proc. IRE, vol. 5, p. 184; June, 1917.
[26] Letter of L. de Forest to Hammond, January 6, 1912.
[27] J. H. Morecroft, "Principles of Radio Communication," John Wiley & Sons, New York, N. Y., p. 402; 1921.
[28] Lowenstein Notebook entry on October 5, 1909.

failure so far seems to lie in the degree of vacuum. . . . I have ordered from the same glassblower who made my tubes five years ago a tube of the same dimensions, and I trust to good luck that he will get as good vacuum now as he did then." But on October 7, 1911, "I have concentrated my efforts on reproducing the telephone tests of last winter, but have not succeeded as yet. . . . I am convinced that only a systematic investigation of the influence of the vacuum, the brightness of the cathode and of the screen potential will assure permanent success." Finally on November 13, 1911, Lowenstein reported upon the first application of the Class A triode amplifier: "At last a test over actual long distances. When I heard your voice I fairly jumped in delight; it came in so clear with every shade of its personal characteristics. . . . Your low voice spoken one foot from the transmitter came in as loud as conversation carried on between two extension phones on the same switchboard."

Prior to securing protection by a patent application filed in April, 1912, Lowenstein demonstrated his amplifier invention to the American Telephone and Telegraph Company, with the circuitry hidden in a "black box"; even subsequently he was reluctant to make a complete technical disclosure. In Gloucester, in late October, de Forest disclosed the achievement of an amplifier gain of 120 times, using three audions in cascade; that his system worked both ways for telephony and was highly regarded by John Stone-Stone; that he was asking the Bell people $50,000 for his invention as applied to telephone work alone. Established histories indicate that including the stated amount for repeater work, the Telephone Company paid de Forest a total of $250,000 for their use of his triode inventions. And Lowenstein, after the issuance of his patent,[29] in 1918 sold the entire Class A grid-bias invention to the Telephone Company for $150,000; the validity of the patent was sustained in later infringement procedures. After the telephone repeater demonstrations of 1912, by Lowenstein, de Forest, and Stone (the three presidents of the Bostonian Society of Wireless Telegraph Engineers), the Telephone Company carried forward its own developments under the competent direction of Dr. H. D. Arnold.

The creation of interest in the triode by a hitherto noncommunications company of development and production competence was on a more simple and direct basis. In September, 1912, the Hammond Laboratory obtained for radio control work seven of the latest designed de Forest triodes. But these were deficient as to reliability, uniformity, and amount of dc change of rectified output in a given test setup.

But already in early 1912, Dr. Irving Langmuir of the General Electric Company, had a very active interest[30] in the flow of current in two-electrode tubes due to the Edison effect; the immediate objective was improvement of incandescent lamps and the potential objective was the development of high-voltage X-ray tubes. There were those who believed that a gaseous content was necessary and desirable for space flow applications. But attracted by the simplicity of the Richardson theory of 1903, as to the source of the current-carrying charges, Langmuir and Sweetser, by August 23, 1912, were already obtaining currents in a thermionic two-electrode tube of 3 to 5 ma at 200 to 250 v with the current limited by space charge and unaffected by gaseous ionization. By November, Dr. Langmuir had completed the work upon which the Coolidge X-ray tube development was based and announced commercially a year later. Langmuir's experiments of November 19, 1912, had shown that pure electron flow in two-terminal tubes could be influenced by electric charges placed upon the glass walls; he immediately recognized that large amounts of power could be controlled by the application of the small amount required for establishing an electric field. Thus the systematic investigation of the influence of the vacuum and the brightness of the filamentary cathode, indicated by Lowenstein as necessary for permanent success in October, 1911, had been independently made for two-electrode tubes by the end of 1912.

At the same time, Dr. Ernst F. W. Alexanderson, also of the General Electric Company, was developing rotary machines[31] for use by operating companies in continuous-wave transmitters. In October, 1912, Dr. Alexanderson was at Gloucester discussing special designs of alternators for radio-control purposes. He was then familiarized with the difficulties with the de Forest triodes, and was briefed as to the Lowenstein developments. It appears[32] that the General Electric Company received one of its first de Forest triodes by February, 1913, and by February 7, 1913, had received the views of the Hammond group as to the proper triode design. These were expressed in four sheets of sketches prepared by Dr. G. W. Pierce of Harvard University, as a consultant to the Hammond Laboratory. The first sheet of specifications, as shown in Fig. 8, based upon the originals, called for a straight tungsten filament held taut by a spring, a nickel spiral grid, and a nickel cylindrical plate. Other sheets referred to a tube with a conical spiral grid, a tube with a lime-coated platinum strip cathode at a dull red heat, and a tube with the de Forest linear flow structure but with grids and plates on both sides of the filament.

By the middle of February, 1913, W. C. White was transferred to assist Dr. Langmuir in the development

[29] U. S. Patent 1,231,764 (1912–17) to Lowenstein.

[30] U. S. Patent Office Interference 40,380, Arnold vs Langmuir; Subject: Electron Tubes; Langmuir's Record.

[31] U. S. Patents 1,008,577 (1909–11); 1,042,069 (1911–12) to E. F. W. Alexanderson.

[32] Hearings before the Committee on Interstate Commerce, U. S. Senate, 71st Congress, 1st Session, on S6, a Bill to provide for the regulation of the transmission of intelligence by wire or wireless, pp. 1360–1371.

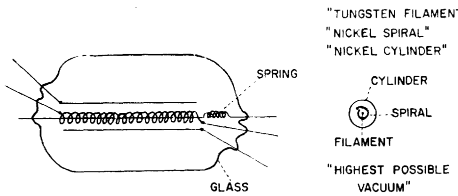

Fig. 8—Hammond-Pierce design for a high vacuum triode, 1913.

of the triode, to test the current de Forest designs, and to follow through with designs according to Langmuir sketches. In his earlier experiments, Langmuir had studied the flow between a cold and a hot tungsten filament, both having been previously heated in the process of eliminating gases and vapors before pumping. Therefore, the early General Electric triodes[33] followed neither the de Forest nor the Hammond-Pierce structural designs, but all three electrodes were of a filamentary nature. Comparative tests between the prototype General Electric and de Forest types for radio control applications were made by Langmuir and Pierce at the Hammond Laboratory on May 19, 1913. The test circuit as recorded by Capt. F. J. Behr of the Coast Artillery, and redrawn in Fig. 9, shows the Lowenstein in-

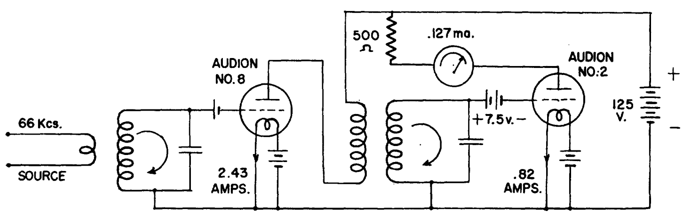

Fig. 9—Langmuir-Pierce test circuit for audion triodes, 1913.

fluence with the negative bias of the first tube as a repeater, the Lowenstein-Hammond influence of the ultranegative bias of the second tube as a potentially operated rectifier detector, and the Alexanderson influence with the tuned cascaded circuits with electronic isolation. For this setup, the de Forest design was considered to be more suitable for control work, but the General Electric tube operated at 200 v with greater future promise. With researches on the relations between triode structure and performance, and with development of new methods of heat treatment for anode plates, Langmuir and White then developed the modern triodes for transmitter and receiver purposes, more closely along the design of the Hammond-Pierce structural specifications.

There is ample evidence from separate sources that Lowenstein developed a radio-frequency oscillator using the low-power triodes of 1911–12. Thus along with discussion of other matters,[34] the Federal Telegraph Company was advised: "Our method is far more reliable and simpler than the Poulsen arc method or the high-frequency alternator method as used by Fessenden and others. . . . In experiments we have found that our method is highly suitable for wireless telephony, as there is absolutely no sound produced whatsoever as in the arc or H.F. alternator. . . . I am quite familiar with the art in Europe and during my recent trip to Germany found that most of the companies had abandoned the arc method of oscillation production. It is for this reason that I believe there is quite a future in the development of the work which we are carrying on."

Possibly the first recorded application of the de Forest triode as an oscillator was in the development of the final circuits of a special security control system,[35] as shown in Fig. 10. This is based upon a report of B. F.

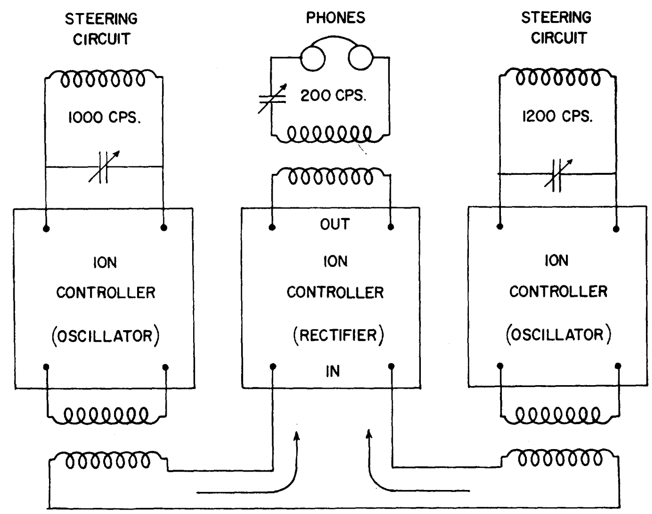

Fig. 10—Hammond selective beat system, with Lowenstein oscillating triodes, 1912.

Miessner on the work at the Lowenstein Laboratory, dated February 18, 1912. Here two triodes, or "ion controllers" by the Lowenstein designation, were fitted with "steering circuits" to generate currents with frequencies of 1000 and 1200 cps. The exact nature of the ionic devices and the manner of operation are not known. The joint outputs were fed to a third ion controller operating as a rectifier, and the resulting 200-cycle output was selectively applied to headphones in lieu of a second rectifier and a control relay.

It is possible that because he was unable to produce a triode transmitter to take the contemplated 500-watt input, or because of previous experience with triodes with mercury pool cathodes, Lowenstein did not consider his triode oscillator work of inventive importance. At any rate, in February, 1912, he turned to the development of his well-known 5-kw spark transmitters for the Navy, but cooperated in making his special knowledge and opinions available to the General Electric Company.

While awaiting the development of hard tubes with

[33] J. A. Fleming, *op. cit.*, 3rd ed., p. 873; 1916, Fig. 7.
[34] Letter of Hammond to B. Thompson, January 25, 1912.
[35] U. S. Patent 1,491,772 (1912–24) to Hammond, Fig. 5.

sufficient sensitivity and power output, the control work of the Hammond Laboratory continued with de Forest triodes and with a mercury triode later commercialized in thyratron circuitry. Such triodes would be biased just below the threshold of firing and would be triggered by the incoming signal. For de Forest types, the tube would be restored several times during a signal dash by a plate or a filament chopper[36] with power delivered to the output relay mainly during the triggered blue-glow condition of gaseous conduction. Restoration of the Pierce mercury type circuit[37] was automatic by use of an alternating plate supply. With close adjustments, the single tube circuit of Fig. 11 provided powerful relay operation even from transatlantic signals. Improvements by Dr. E. L. Chaffee,[38] with special care in aging the interior of the tube during the pumping process and in properly thermostating the tube with an oil bath at 70° C, permitted the device to be used in improved types of radio control circuits.[39] The General Electric Company acquired rights to the Pierce tube and circuitry by way of Cooper-Hewitt and Hammond. When hard tubes for power amplification purposes became available, they were of course utilized. Thus, in the final torpedo work, the control signal was built up to a high level so that with a special ac-dc converter system, the first electromagnetic mechanism in the chain operated from a triode with 10-w dc output. The development of the triode for transmitter and high level purposes was the step necessary for providing reliable circuitry in the radio-control field in which the need for improved electronic devices was early appreciated.

IV. Modern Intermediate Frequency Circuitry

Intermediate-frequency circuits carry power of a frequency range intermediate between that used in conveying the information to the receiver and the audio frequency used to operate the receiver indicator. They were developed from consideration of the phenomena of beats, long established in the physics of sound, but first applied for radio purposes by Fessenden in heterodyning a keyed continuous wave to produce an audio current for signal indication purposes. As previously indicated, Tesla had proposed security of radio control by simultaneous transmission of two high-frequency signals and coincidental operation of relays after individual reception and rectification. If this Tesla method were practiced with continuous waves, beats would exist in the signal medium, but nowhere would there be any physical current of the beat frequency. In 1912, it was therefore proposed[40] that, as before, the two continuous

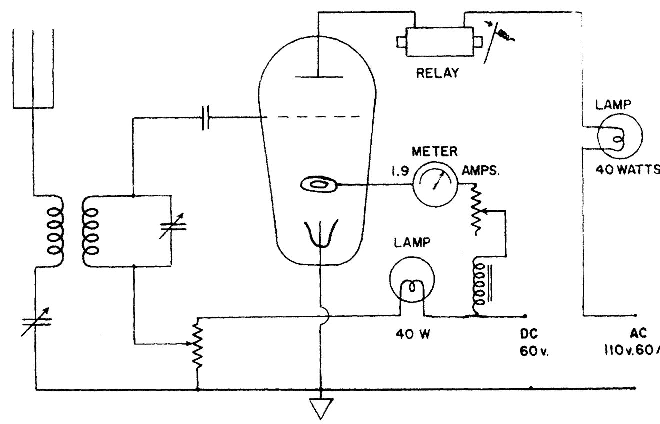

Fig. 11—Pierce mercury triode receiving circuit, 1914.

waves should be individually received to yield some degree of security, but that this security then be increased by applying the two tuner outputs to a common detecting circuit and tuning the detected output to the difference or beat frequency. For control purposes, this new ac signal in turn, could be applied to a rectifier to drive a dc relay. Or for communication purposes, if the beat frequency also was too high for direct operation of headphones, it could be applied to a conventional continuous wave type receiver of the Fessenden type. Or if the system was conveying telephony with the beat frequency itself modulated, the signal would be recoverable by a second detection, as by the condenser type headphones of the period that served both as detectors and as indicators.

Even greater security was proposed[41] in which the two radio waves were amplitude modulated at different frequencies. In this manner, the transmitter to exercise control was required to have not two but four frequency generators of precision. It was in the working out of this system that the triode oscillators of Fig. 10 were used to simulate the two modulating frequency signals as recovered by first rectifier detections.

It was planned to practice these inventions for radio control by use of two Alexanderson alternators, using both of the Point Radio antennas; for the second method, the alternators were to be modulated by currents applied to the field windings.[42] But the applications of beat selectivity in communications were also of importance. Quickly considered were double modulation with the intermediate frequency approximately the geometric mean of the radio and the signal frequencies,[43] and the securing of selectivity in radio telephonic communications.[44] All these advanced ideas for increased selectivity were delayed in patent prosecution and in publication during World War I.

When Dr. Alexanderson conferred in Gloucester in October, 1912, regarding the detail design of the alter-

[36] U. S. Patent 1,610,371 (1914–26) to Hammond.
[37] U. S. Patents 1,087,180; 1,112,549 (both 1913–14) to G. W. Pierce.
[38] U. S. Patents 1,550,877 (1916–25) and 1,627,231 (1915–27) to Chaffee.
[39] U. S. Patent 1,491,775 (1916–24) to Hammond, Figs. 8, 9.
[40] U. S. Patent 1,522,882 (1912–25) to Hammond.

[41] U. S. Patent 1,491,772 (1912–24) to Hammond, Figs. 1, 5.
[42] U. S. Patent 996,445 (1909–11) to E. F. W. Alexanderson.
[43] U. S. Patent 1,491,773 (1912–24) to Hammond, claim 4.
[44] U. S. Patent 1,491,774 (1912–24) to Hammond, claims 72–79.

nators, he was shown the general concept of selectivity in reception with a radio tuner, a first detector, an intermediate frequency selector, a second detector and an audio output circuit. Audions were shown as detecting devices, and their use as amplifiers by Lowenstein was known. Therefore, Dr. Alexanderson considered that for immediate engineering purposes,[45] the first detector should be changed to be a radio amplifier and the intermediate frequency circuit to a second radio tuner. Disturbed by the suggestion that the de Forest tubes were not sufficiently reliable and perhaps too sluggish for radio amplification, Alexanderson later proposed[46] a highly interesting diversity system, Fig. 12. Here two groups of triodes were driven from two transmission line-like arrangements. All tubes of a group were driven in phase because of the nature of the line construction. Therefore, their outputs were additive, and without blue-glow gaseous conduction the combined gain could be greatly increased. If the tubes were all detectors, then the system would be of the intermediate frequency type, but if the first were amplifiers and the second were detectors, then it would be of the new tuned-radio-frequency type with both lines of the same high frequency. Thus the October, 1912 conference not only disclosed the ultimate selective receiver with double detection but, in addition to stimulating the development of the hard triode, initiated the invention that mainly supplies the sensitivity features of modern receiver systems. This conference was, therefore, of special historical significance.

While the Alexanderson alternators, each with 2 kw, 100-kc rating, were not often used in control work, experiences in their operation were undoubtedly of value in the well-known future developments leading to the formation of the Radio Corporation of America by the General Electric Company as a result of governmental policy fostered by the Navy Department.

The intermediate frequency principle was first applied outside the laboratory[47] to the solution of a World War I communication problem of high military importance.[48] Radio-telegraphic equipment was desired for signaling from front-line infantry to barrage-laying artillery in the face of powerful interference from enemy spark-type transmitters. The Hammond solution involved a superaudible amplitude modulation of a gap-type radio transmitter with a suitable intermediate frequency receiver for producing an audio signal tone. The transmitter schematic, Fig. 13, shows[49] how the radio or "A" radiation was created by charging, at a superaudible "B" frequency rate, a capacitor which produced "A" oscillations by discharging across a special

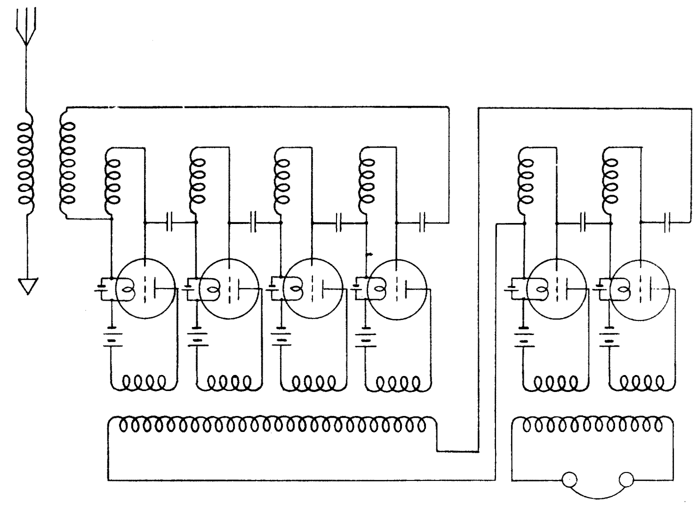

Fig. 12—Alexanderson multiple tube diversity circuit, 1912.

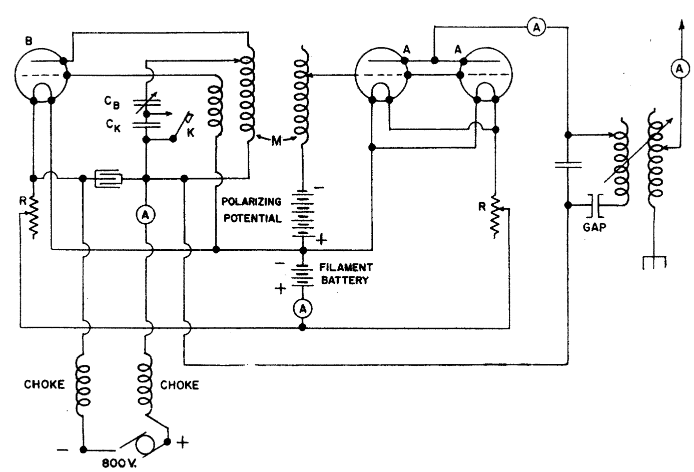

Fig. 13—Hammond-Chaffee tube-gap military transmitter, 1917.

type of spark gap. The system was of the constant energy type with signaling by key shift of the modulating frequency. While the prime purpose was selectivity, the method of keying in combination with the irregularity of the phase of the modulation at which the gap fired made the system highly secure against reception by conventional receivers. The receiver schematic, Fig. 14, shows[50] a regenerative radio tuner and first detector *X*, followed by an intermediate-frequency selector circuit *B* with an oscillatory detector *Y*, followed by an audio tuned circuit *D* with an amplifier *Z* providing audio regeneration but without causing excessive ringing. In normal operation, the first detector was nonoscillatory, but the feedback was such that the receiver was usable for single-tube continuous-wave reception. Structurally, therefore, the receiver was of the most general "superheterodyne" variety, since both detectors could be, and during adjustment often were, of an oscillatory nature. The final received signal due to the desired message was

[45] U. S. Patent 1,173,079 (1913–16) to E. F. W. Alexanderson.
[46] Letter of Alexanderson to Hammond, October 21, 1912.
[47] Signal Corps Order No. 40,105 of 1917.
[48] Even as late as October 9, 1918, this problem was recognized as "one of the most important matters connected with the war." See letter of that date, L. N. Scott, War Comm. of Tech. Soc. under Naval Consulting Board heading.
[49] U. S. Patent 1,610,425 (1918–26) to Chaffee.
[50] U. S. Patent 1,681,293 (1917–28) to Hammond; and 1,469,889 (1918–23) to Chaffee.

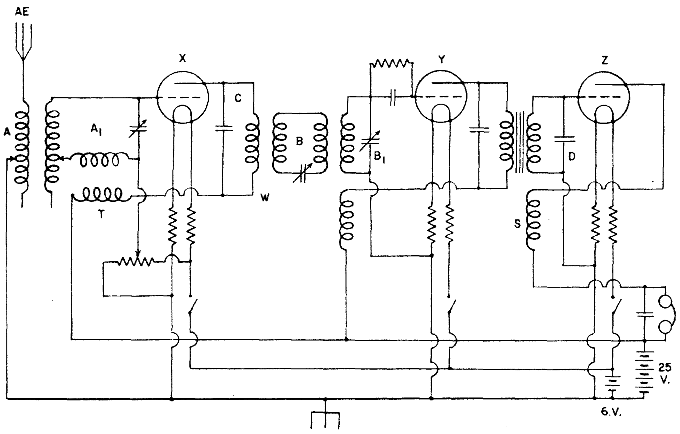

Fig. 14—Hammond-Chaffee intermediate frequency receiver, 1917.

highly free of disturbance from spark interference for a number of reasons. From a frequency domain viewpoint, the intermediate frequency circuit tuned to the fundamental of the envelope of the desired signal was not appreciably energized by the weak harmonics in the envelope of the interference. In the time domain there were lapses at the ends of the spark trains of interference during which the only signal getting to the first detector was the desired signal. Additionally, a limiter-type soft detector greatly reduced the ratio of the interference to the signal as they were both applied to the intermediate frequency circuit. The net result was that desired signals were often readily received with the interference to signal ratio at the tuner output corresponding to as much as 20 db.

This equipment was designed in accordance with military requirements as to size, weight, number of tubes, etc., and constructed by the Hammond Laboratory in 1917–18, with E. R. Cram serving as field representative of the Signal Corps. Delivered to the U. S. Army at Tours, France, by Dr. Chaffee of the Hammond group on October 10, 1918, the equipment created great interest. The performance claimed was verified by official tests directed by Capt. E. H. Armstrong of the Paris Laboratories of the Signal Corps, October 22 to November 11, 1918. Post-armistice tests were made by the British Army from whose report by Major A. C. Fuller of the Royal Engineers the schematics have been taken. This report of December 3, 1918, included the following: "The Hammond set and a 500-watt spark set were simultaneously operated on the same wavelength —710 meters—at Woolwich Common. Aerial current from Hammond set about 0.7 amp., from 500-watt set, 4.5 amps. The Hammond receiver was set up on Blackheath. The interference due to the 500-watt set was insignificant and did not prejudice reception."

In this equipment, the mean frequency of the transmitted band was 500 kc, and the bandwidth was at least twice the modulation frequency of a mean value 25 kc. While such a 10 per cent bandwidth would be intolerable for peacetime signaling merely to produce an audio tone in a receiver, this early application of intermediate frequency circuitry provided perhaps the only possible solution of an urgent wartime problem.

The value of superaudible modulation of a gap transmitter for obtaining selectivity, sensitivity, and security continued to be demonstrated to the Army and Navy in equipments for intrafleet and plane-to-ground communications and in radio control of aircraft. But with the increase of radio services without a corresponding opening up of the higher frequency bands, the interferences by superaudibly modulated spark transmitters upon conventional communication channels became of compelling importance. Therefore, the Hammond Laboratory developed[51] all-tube equipment for radiating only the minimum spectrum necessary for operating the system. Two continuous-wave equivalents of high purity were created by push-pull modulation with carrier suppression, the key-down modulation frequency determining the intermediate frequency of the receiver. Security of communications was obtained by means of an artificial frequency modulation of the carrier oscillator, using a variable capacitor wobbler in combination with key shifts with several excursions of wobble per telegraphic dot element. The total frequency swing was usually a compromise value. On the one hand, a wider swing reduced the amount of time that a given narrow-band disturbance could affect the intermediate-frequency circuit of the receiver; on the other hand, a narrower swing permitted better radio selectivity, as by coupled-circuit tuning with primary and secondary regeneration. Experiments in 1921–22, with these systems conducted by Dr. A. Hoyt Taylor and Dr. Chaffee in the Navy laboratories established points of interest regarding information theory when the interference greatly exceeds the signal. These tests proved that the nature of the detector system is of high importance in security systems based upon simple modulations. In the steady state condition, the square-law detector alone has a sensitivity for the desired signal that is not disturbed by the presence of the interference. Material pertinent to these matters has been recorded by Aiken;[52] the need for square-law detection instead of linear is not necessarily as great in pulse systems for security control.

With the increased crowding of channels, it became evident that the important field of application of superaudible modulation was in multiplex signaling. That is, one single carrier could serve a number of sidebands, each carrying its own signal message. Laboratory and field demonstrations of August, 1925, used a carrier of 33 mc with one fixed and one variable superaudible frequency of amplitude modulation. Keying for sending messages simultaneously was by the constant energy method, with frequency shift of the amplitude modulating frequency values for minimizing cross signaling at high carrier modulation levels. These tests showed

[51] U. S. Patent 1,690,719 (1922–28) to Chaffee and Purington.
[52] C. B. Aiken, "The detection of two modulated waves which differ slightly in carrier frequency," Proc. IRE., vol. 19, p. 120; January, 1931.

the practicability of sending as many as eight independent messages with the same carrier, with AM first modulations and fm second modulations. While this combination is widely used in modern practice, much radiotelephonic multiplex uses AM for both modulations with special feedback methods to linearize the modulation characteristics of the system.

While observing the performance of the two-wave system in 1922, Gen. David Sarnoff of the Radio Corporation suggested consideration be given to private radio telephony by the same general principles. This was quickly developed[53] with an artificially wobbled carrier that was amplitude modulated by an audio band after conversion to the superaudible range by heterodyne methods and filtering. The receiver was of the same general type as in the telegraphic system but with the intermediate-frequency channel of speech bandwidth. The receiver heterodyne was required to be set correctly within 5 to 10 cycles for good speech quality, but more precision was required for music. This system of telephony was practiced with a transmitter on the roof of the Ministero del Interno in Rome, in 1928, to provide coverage up to 30 km with simple receivers. A later development[54] of a narrow-band system in which the speech band was not appreciably increased in width during the conversion was proposed for police work, and improvements with greater complexity were made available[55] through OSRD-NDRC for consideration in transoceanic telephony for World War II.

When the high carrier-frequency bands commenced to open up, interest again developed in the use of superaudible modulation in simplex communications. In 1932, the Hammond Laboratory illustrated the system by transmissions at the 1-kw level from Gloucester to an Army group assembled in Washington in what is now the French Embassy. While the best monitoring facilities of the Government tuned to the Hammond band reported the radiations resembled some new kind of man-made static, the messages were received on the proper equipment with high telegraphic quality at loudspeaker level. In these tests, three-wave supersonic amplitude modulation was used with sufficient artificial frequency modulation of the carrier to yield a fairly smooth spectral energy distribution. It was noted that the signals were relatively free from selective fading, and this was considered to be due to the frequency modulation.[56] That is, during the shortest signal element not three but hundreds of spectral lines were involved in the transmission; thus the effect of selective fading was reduced as well as perhaps some noise.

Similar results were obtained in the higher frequency ranges in plane-to-ground transmissions at the 50-watt level. These transmissions were received at Washington throughout runs to Aberdeen, Md., to Martinsburg, W. Va., and to Norfolk, Va. These tests perhaps contributed to getting military aircraft radio out of the broadcast band and to establishing better monitoring facilities for examining static-like signals coming into the Washington area.

In the field of intermediate-frequency receivers, there are four species classifiable by the different combinations of the beat and nonbeat natures of the two detectors. With both of beat nature, as is possible with Fig. 14, the receiver is applicable to continuous-wave reception; with both of the nonbeat type, the receiver is used for doubly modulated waves either in simplex or multiplex systems. For the military applications just discussed, the first detector alone is of the nonbeat type; the remaining combination with the first of the beat type and the second nonbeat is that of the familiar "superheterodyne" for reception of radio telephony as in broadcast radio.

Considering the commercial potentialities and the rigid publication restrictions upon the Hammond developments, it was probably inevitable that many others should develop parallel lines of thought. With the release from secrecy at the end of World War I, an important interference developed[57] in the patent office between R. A. Heising,[58] Hammond[59] and L. Levy[60] in the broad field of intermediate-frequency circuitry. After thorough studies of the early United States and foreign art and a clarification of the distinctions between detection and rectification, the broad subject matter in controversy was awarded Hammond, giving rights[61] for the exploitation of the following word combination: "A carrier wave transmission system comprising means for receiving and detecting the energy of a modulated wave, means for selecting a component of said detected energy, and means for detecting said selected component." The entire principle of IF selectivity is expressed by the words "selecting a component" regardless of whether the unselected components were to be utilized otherwise as in multiplex reception, or were to be discarded as in simplex telephonic reception. Patent claims more specific to the superheterodyne structure for telephonic reception were awarded to Hammond in a coissued patent.[44]

The commercialization of the IF principle for broadcast reception involved a somewhat different approach at first involving sensitivity rather than selectivity considerations. Observing the difficulty of amplifying short-wave signals in comparison with those of long-wave, E. H. Armstrong[62] developed the idea of applying heterodyne conversion of the incoming signals to a lower

[53] U. S. Patent 1,642,663 (1922–27) to Chaffee.
[54] U. S. Patent 2,204,050 (1938–40) to Purington.
[55] U. S. Patent 2,400,950 (1942–46) to Purington.
[56] U. S. Patent 1,761,118 (1924–30) to A. N. Goldsmith.

[57] U. S. Patent Office Interference No. 43,858.
[58] U. S. Patent Application Ser. No. 81,980, filed in 1916.
[59] U. S. Patent Application Ser. No. 175,134, filed in 1917.
[60] U. S. Patent Application Ser. No. 249,572, filed in 1918.
[61] U. S. Patent 1,491,772 (1912–24) to Hammond, claim 46.
[62] E. H. Armstrong, "A new system of short wave amplification"; Proc. IRE, vol. 9, p. 3; February, 1921.

frequency and more readily amplifiable band, then amplifying this band before detection to produce an audio signal. Armstrong filed for a French patent on December 30, 1918. Possibly because of official knowledge of the Hammond development of IF selectivity, Armstrong discussed only the sensitivity features in his patents and technical papers. His U. S. patent claims[63] were later awarded to the other claimants; those pertaining to the radio-telephonic superheterodyne went to L. Levy,[64] those pertaining to the continuous-wave superheterodyne went to Alexanderson,[65] but one of these also relating to radio rebroadcast repeaters later was awarded to B. W. Kendall.[66]

With the growth of broadcasting, the Alexanderson TRF system of receiver design gave way to the Hammond IF system mainly because of the requirements for superselectivity.[67] But the TRF system of cascaded selective amplification continues in preamplifiers and in IF circuits to be one of the most important elements of modern receivers.

It is noteworthy that the Hammond group additionally made other important but less basic contributions to the details of modern broadcast and receiver techniques. The principle of radio relaying by change of the carrier frequency was an early contribution.[68] An early form of automatic volume control[69] used the grid capacitor method of detection with an electronic shunting triode having a resistance that was a decreasing function of the capacitor voltage. Remote cutoff action was inherent in the Pierce proposal of 1913 to build a tube with a conical spiral grid; another solution was the use of three triodes with grids in parallel, plates in parallel, but with the cathodes at different dc potentials. There is even a suggestion in the radio-control records of the modern idea of feedback to improve fidelity of output by having the output relay of a detector rectifier discharge the grid capacitor in its input. But more positively, the Hammond group contributed to the adoption of the unicontrol superheterodyne for broadcast reception. In early models, separate controls for the radio tuner and the heterodyne oscillator provided a technically desirable flexibility. But the existence usually of two and sometimes three or more different heterodyne settings for developing the proper intermediate frequency was confusing and gave the home user a sense that the system was not selective. And in comparison with the TRF designs, the added heterodyne dial was an undesirable complication. The Hammond group, in 1917, had developed the military simplification[70] of a single-knobbed switch connected to fixed capacitors such that the difference between the heterodyning frequency and the frequency of the selective circuit tuned to the incoming signal was independent of the switch setting. With this background, the Hammond group developed and, on April 13, 1925, was among the first to demonstrate a continuously variable unicontrol superheterodyne and to urge the adoption of this technique now almost universally used in home-instrument type receivers.

V. Frequency Modulation and Related Systems

Even before the development of radio communications, some of the very basic principles of frequency modulation were discussed by Helmholtz[71] in the field of sound. He indicated that when beats are formed from two unequal but substantially pure tones with slightly different pitch frequencies, "a little fluctuation in the pitch of the beating tone may be remarked." That is, a musician can hear such tones as a variation in strength and a variation of pitch of a single tone. The mathematical explanation appears to have been due to G. Gueroult[72] who had translated Helmholtz into French, while a corresponding graphical type of explanation was provided by Taylor.[73] Thus the velocity of a particle vibrating under the influence of two tones was:

$$v = C \sin (mt - \epsilon)$$

where C and ϵ were slowly varying amplitude and phase functions of time. Moreover, "the pitch number of the variable tone multiplied by 2π is $\cdot \cdot \cdot (m - d\epsilon/dt)$." Thus it was recognized that two unequal waves add up to the equivalent of a single wave modulated both as to amplitude and as to phase, and that the instantaneous frequency was determinable from the time derivative of the instantaneous phase. It was further shown that with the two tones unequal in strength, the instantaneous frequency would swing from within to outside the spectral limits.

In the radio field of 1912, frequency modulation was used commercially in the constant amplitude method of continuous wave transmission, without requiring a violent keyed change of the energy content of the oscillatory system. And in the Hammond Laboratory, it has been established that two independent communications could be sent in the same wave band, one by AM for telephony and one by fm for telegraphy.[74] In modern practice, the transmission of the two chrominance signals in color television is a refined example of this multiplex principle. So also to a lesser degree, in television receivers where the intermediate frequency for sound is established by the beating of the video and sound car-

[63] U. S. Patent 1,342,885 (1919–20) to E. H. Armstrong.
[64] U. S. Patent 1,734,038 (1918–29) to L. Levy. (See footnote reference 60.)
[65] U. S. Patent 1,508,151 (1916–24) to Alexanderson.
[66] U. S. Patent 1,734,132 (1916–29) to Kendall.
[67] See footnote reference 32, p. 278.
[68] U. S. Patent 1,313,860 (1912–19) to Hammond.
[69] U. S. Patent 1,649,778 (1917–27) to Hammond and Chaffee.
[70] U. S. Patent 1,484,605 (1917–24) to Hammond; see also U. S. Patent 1,849,651 (1924–32) to S. E. Anderson.

[71] Helmholtz, "Sensations of Tone," Peter Smith, Gloucester, Mass., 6th ed., p. 165 and pp. 414–415; 1948.
[72] *Ibid.*, footnote, p. 165.
[73] S. Taylor, "On variations of pitch in beats," *Phil. Mag.*, vol. 44, pp. 56–64; July, 1872.
[74] U. S. Patent 1,320,685 (1912–19) to Hammond.

riers, cross signaling between the picture tube and the speaker is minimized by this principle.

During the early development of radio, many engineers of good repute believed in the frequency modulation method of telephony, and electromechanical methods for transmission had been proposed.[75] Perhaps to a large degree such opinions were due to a then current belief that less bandwidth was required for fm than AM. As discussed under Section IV, the Hammond Laboratory was seeking means for conveying a superaudible frequency signal with a minimum of disturbance upon other channels. It was known that if a continuous-wave generator was changed periodically in frequency at a slow rate, then the tone resulting from suitable heterodyne reception was of the continuously varying siren type. In discussions, the question arose as to what happened when the rate and the extent of the frequency variation were pushed up in value. Since it was not practical to use the rotary capacitor method, the first all-electronic method of frequency modulation[76] was worked out in January, 1921. This involved setting up a triode oscillator with a plate tank circuit of high L/C ratio and with a high ratio of plate to grid feedback to yield poor frequency stability. By varying the plate and the grid dc supply voltages in phase with a proper ratio, it was found possible to produce fm with negligible AM; the oscillator could be varied about 14 kc either way from its mean 580-kc value. After checking at a slow 60-cycle modulation frequency and obtaining what was expected in a wavemeter varied as to setting, the modulation rate was pushed up to 22.5 kc. First order side frequencies were found as expected by the Fourier theory of recurrent wave forms, both by wavemeter and by beat oscillator tests. Lowering the modulation frequency to 9.5 kc, the two types of response as the swing was increased in steps from conditions 1 to 4 are as shown in Fig. 15. For the greatest swing used and the selected frequency, the carrier by both tests became smaller than the first order side frequencies. The shifting of the carrier toward the higher wavelengths with increased amount of modulation was probably due to the slight nonlinearity of the modulation characteristics.

These experiments proved that a frequency modulated signal would cause more interference upon other channels than an amplitude modulated signal for the same purpose, and the two-wave method of modulation was adopted for the immediate purposes of design. It was soon realized that a general expression for any modulated voltage wave was

$$e = A_t \cos(\omega t + \phi_t)$$

where A_t and ϕ_t are two slowly varying functions of time, and that for pure sinewave frequency modulation, the expression becomes

$$e = A \cos(\omega t + \phi_m \sin pt)$$

[75] U. S. Patents 785,803–804 (1902–05) to C. D. Ehret.
[76] U. S. Patent 1,599,586 (1922–26) to Purington.

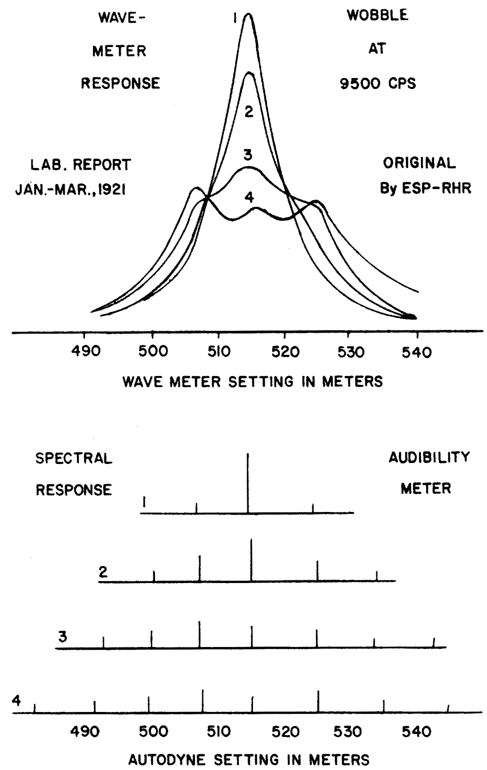

Fig. 15—Purington frequency modulator performance, 1921.

where A is a fixed amplitude value, ω and p are the angular frequencies of the carrier and modulating waves, and ϕ_m is the maximum departure of the carrier phase from mean. After the classical solution of the rotary capacitor frequency modulator by Carson,[77] expansion of the above expression by trigonometric methods confirmed the identification of the Fourier spectral amplitudes with Bessel function values. This expansion, also known to have been made by others, was first openly published by Roder.[78]

But as of 1921, although consideration was given to the fm method of telephony by use of a double winding modulation transformer to provide both plate and grid voltage variation for an oscillator, it was realized that demands for channels made it imperative that broadcasting be developed first on an AM basis. Nevertheless there was some willingness to consider wider-than-necessary systems for speech telephony in the interests of reduction of noise. Fig. 16 shows one of several arrangements due to Chaffee,[79] experimentally constructed to permit examination also of another fm related idea. Here two radio carriers F_1 and F_2 were amplitude modulated in an out-of-phase manner from the same speech source, making the radiated energy at one end of the spectrum a maximum when it was a minimum at the other end. By making use of the phase

[77] J. R. Carson, "Notes on the theory of modulation," Proc. IRE., vol. 10, p. 57; February, 1922.
[78] Hans Roder, "Amplitude, phase and frequency modulation," Proc. IRE., vol. 19, p. 2145; December, 1931.
[79] U. S. Patent 1,776,065 (1922–30) to Chaffee.

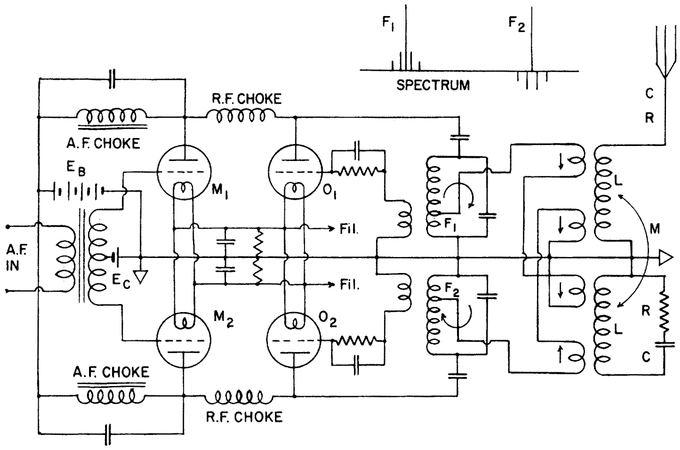

Fig. 16—Chaffee noise-reduction system transmitter, 1922.

relations of coupled circuits,[80] both modulated oscillators fed the same antenna without reactions of either source upon the other.[81] This was an early form of diplexing, but with half the power from each source lost in the dummy tank circuit. The same phase principles, but with direct coupling circuitry, are used in modern frequency discriminators to perform the opposite function of distributing the two ends of a spectrum to different use circuits such as diode detectors. The receiver for this Chaffee system used individual tuners as in the Tesla two-wave procedure, but with two output transformers with secondaries connected in series in a manner to cause the two detected signals to add coherently and to provide cancellation of the noises that were common to both channels. An alternative transformerless method would develop from the differential circuit of Fig. 4 with the loops tuned to frequencies F_1 and F_2, respectively and with the relay replaced by headphones. Patentwise, the principle was expressed in part as follows: "In a receiving system for radiant energy, a plurality of receiving channels tuned to the energy of different frequencies, respectively, means for producing currents of like frequencies but of different phases from the received energy, an indicating device, and means interposed between said channels and the indicating device for causing said currents to combine additively and to simultaneously actuate said device." Experimentally this system greatly reduced noise effects such as filament hum that were equally present as amplitude modulations in both channels. Since the Hammond group at that time was primarily concerned with high noise-to-signal ratios, no thought was given to the diversity properties of the system by which the desired signals added coherently and the random noises built up incoherently.

The relation of this system to that of modern frequency modulation is obvious. There can be no question but that the radiation is properly receivable by an fm type receiver with a suitable line-up of the IF and discriminator circuits. But to point up the comparison more clearly, the amplitude, phase and frequency variations of a single voltage vector representing the entire radiation have been evaluated for a special case (Fig. 17, opposite). Positive and negative lines in the spectrum are representations of various cosinusoidal waves of different frequencies. For positively shown lines, the phase is zero at $t=0$, at which time the phase of a negatively shown line is π radians. The main lines of amplitude E are offset an angular frequency value δ from the reference angular frequency ω at the center of the spectrum; the signal created lines of amplitude $kE/2$ are offset by the signal angular frequency p from the main lines. The expression for the instantaneous totalized voltage is recorded. From this expression, the amplitude function A_t, the phase variation function ϕ_t, and the angular frequency deviation function $d\phi_t/dt$ are readily developed in terms of the frequency parameters δ and p and the signal modulation parameter k; the routine procedure has been exemplified in the Helmholtz reference. With the separation parameter δ an integral multiple of the signal modulation angular frequency parameter p, these functions are recurrent in one signal cycle. Curves are shown in Fig. 17 for the conditions $\delta=3p$, $k=1$. The absolute value of the amplitude function is plotted positively to show the wave form recoverable by applying the spectrum to an ideal aperiodic linear rectifying system. The instantaneous phase and frequency curves correspondingly show the wave forms producible by applying the spectrum to ideal aperiodic phase and frequency modulation detector systems.

In general, the amplitude function departs mostly from uniformity at the period of the modulation cycle when the two main spectral lines and likewise the two lines of each set of signal produced lines are momentarily of opposite phase to produce cancellation. At this time, the frequency deviation is also passing through zero. While the signal distortion producible by phase modulation reception is small, that for frequency modulation reception is of course much greater. With a properly lined up fm type receiver, the Chaffee type signal would be receivable without distortion; such a receiver would of course distort a pure sinusoidal pm or fm signal. The noise reduction merits of the system in comparison with an fm system of comparable width and signal carrying capacity have not been evaluated.

The Hammond group soon developed the ultimate of spectral compactness for a two-channel system most like fm except in bandwidth. This was based upon the phase reversal of one of the sidebands of an amplitude modulated signal, or as a practical alternative, a ninety degree phase shift[82] of the carrier with respect to the

[80] U. S. Patent 1,601,109 (1922–26) to Chaffee.

[81] The circuitry is discussed in E. S. Purington, "Single and coupled-circuit systems," Proc. IRE., vol. 18, pp. 996–998; June, 1930.

[82] U. S. Patents 1,935,776 (1929–33) and 1,976,393 (1929–34), to Hammond, Fig. 8.

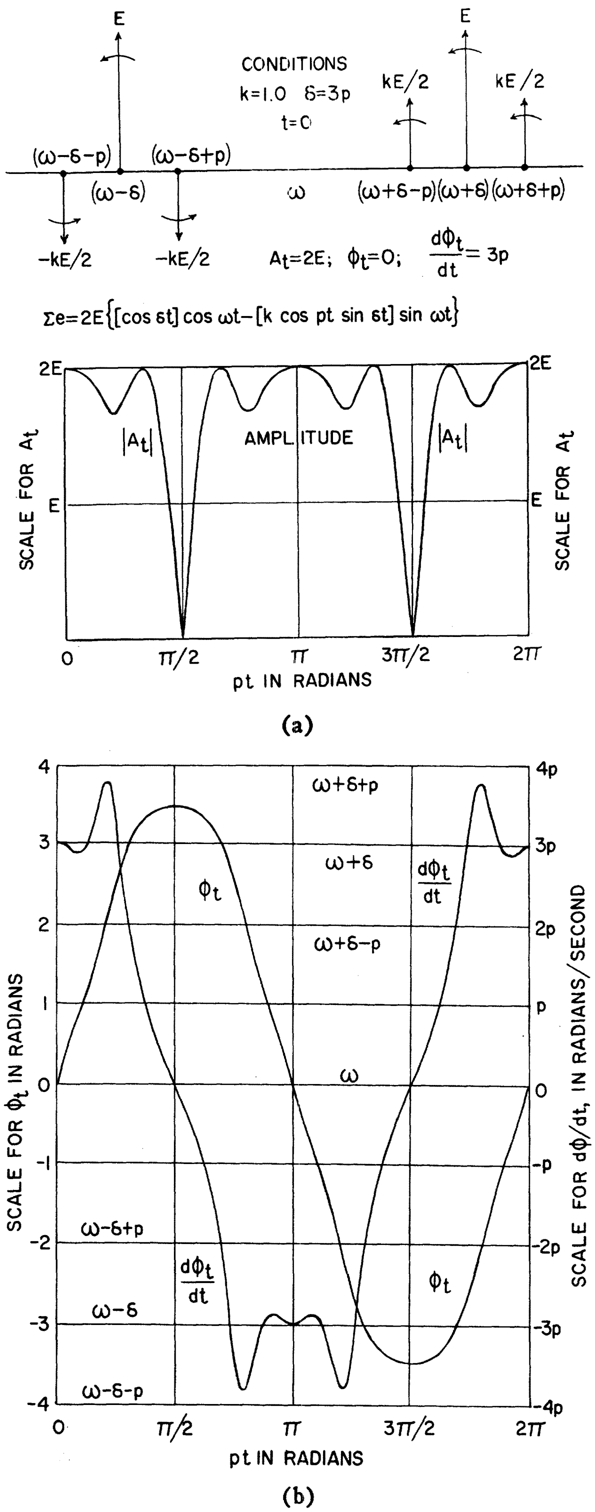

Fig. 17—Amplitude, phase, and frequency characteristics, Fig. 16.

condition for which the carrier and sidebands would represent pure amplitude modulation.[83] Fig. 18 shows how the sidebands were created by a push-pull amplitude modulation, with the output of which the carrier could be combined in any desired phase in accordance

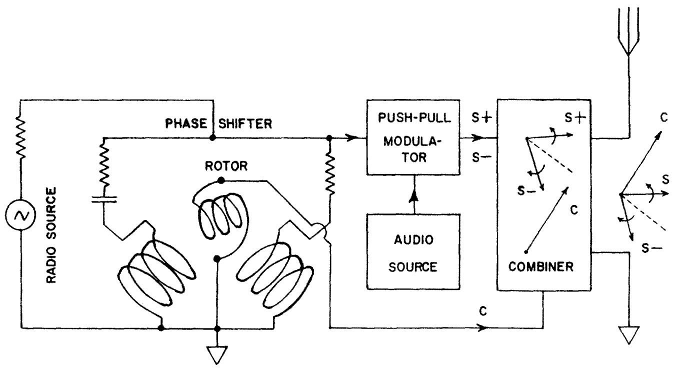

Fig. 18—Hammond "Quasi-phase" modulator, 1927.

with the setting of a rotary phase shifter. The desired radiation was produced when the rotor was so set that with a 1000-cycle signal tone, an amplitude detector of the radiation produced a 2000-cycle tone with no 1000-cycle residual. As in the reception of the previously discussed radiation, differential detection resulted in recovery of the desired signal and the reduction of noises common to both channels. This involved diverting one sideband and half the carrier to one detector and the other sideband and half the carrier to another detector.

It was quickly realized this "quasi-phase" signal could serve as the basis of an artificial frequency modulated signal with a stabilized carrier. This involved frequency multiplication and amplitude limiting together with tapering the audio signal to convert from phase modulation to frequency modulation characteristics. This process was outlined in a tutorial part of a patent[84] for consideration in commercial design when, as, and if the establishing of fm sound transmission was deemed to be in the public interest.

It has been recorded above how fm had early applications in security systems for radio control and for telegraphic and telephonic communications, and had been proposed for fading reduction in telephonic transmission. With the development of radio facsimile, fm was again proposed for fading reduction, using the principle that when the signal faded but still existed, the instantaneous frequency value could indicate the picture element tonal value desired to be recorded. Mertz[85] is considered to be the first to show this principle but in the field of wire communications. Wright and Smith[86] in Great Britain proposed it for radio-facsimile, and in the United States Hammond[87] further contributed to the art. The technique was to apply severe limiting action by clipping the received and detected fm signal to a fixed level, to feed the clipped signal through a slope filter to eliminate harmonics created by the clipping, and also to produce a signal of strength dependent solely upon the instantaneous frequency transmitted,

[83] The vector-tensor method of representing and handling an amplitude modulated wave is shown in E. S. Purington, "The Operation of the Modulator Tube in Radio Telephone Sets," Bureau of Standards, Scientific Paper No. 423; 1922.

[84] U. S. Patent 2,020,327 (1930–35) to Purington.

[85] U. S. Patent Office Interference 61,606; U. S. Patent 1,548,895 (1923–25) to P. Mertz.

[86] U. S. Patent 1,964,375 (1926–34) to Wright and Smith.

[87] U. S. Patents 1,977,438 (1929–34) and 2,036,869 (1929–36) to Hammond.

and finally to apply the filter output to a light producing indicator. This application of fm in facsimile systems was made in the Hammond Laboratory on May 25, 1927.

The Westinghouse group was probably the first to actively press the commercialization of fm for voice transmission. A. Nyman[88] had considered improved methods of electro-mechanical frequency modulation of an oscillator. E. H. Armstrong[89] in 1927, filed upon "a new method of transmission in which the frequency of the transmitted wave (not its amplitude) is varied in accordance with the voice frequency to be transmitted." His patent claims, however, were restricted to receiver circuitry, including the combination of limiter action and dual channel detection. In interferences, the two-channel idea was credited to Conrad[90] also of the Westinghouse group, but the basic limiter idea was credited to Mertz above mentioned. However, it appears that Armstrong and later Hammond[91] were among the first to realize that the distortions due to clipping the IF signal for minimizing AM effects could be mitigated by the harmonic rejection discriminator before the second detection.

As of 1927, Armstrong was uncertain as to what bandwidth should be used for fm transmission of speech and music. Thus in the patent application he indicated: "The band may be made any width desired depending on the particular conditions and the distance over which it is desired to operate. This can only be determined by experiment. In general, however, the narrower the band, the less the effect of atmospheric disturbances." This doctrine to favor narrow-band operation because of atmospherics is in accord with the accepted beliefs of the times, although irrelevant when the interferences as in military communications and radio control were sometimes far above the atmospheric level. But in 1933, after further consideration and experimentation as a free-lance inventor, Armstrong indicated:[92] "I have discovered that by imparting a greater swing to the frequency of the transmitted wave . . . a very great improvement in transmission can be produced." These discoveries were in a wave region where in general natural atmospheric disturbances were of a lessened importance, and where greater swing and channel width would by governmental protection not result in interferences from or upon most man-made radiations.

Under these conditions, fm provided a good engineering solution to the problem of providing high-quality, high-fidelity, disturbance-free music transmission at a relatively low transmitter cost. Experts in information theory revised their rules relating the possible amount of signal information to power, distance, wave frequency, bandwidth, time, and signal-to-interference ratio. Engineering developments were stimulated, resulting in improved methods of transmission and reception and the application of fm in special purpose voice communications. With the advent of television, the adoption of the highly developed fm system of sound transmission provided for the best possible cooperation between a sound channel and the basically AM system of video transmission.

VI. Acknowledgment

We wish to acknowledge the assistance given by Dr. E. F. W. Alexanderson, Dr. I. Langmuir, and Mr. W. Dubilier in supplying early data referred to in this history. We further express appreciation for the contributions of the many laboratory and field engineers of the Hammond staff who have participated in the little-known but highly important developments here for the first time openly recorded.

[88] U. S. Patent 1,615,645 (1920–27) to A. Nyman.
[89] U. S. Patent 1,941,447 (1927–33) to E. H. Armstrong
[90] U. S. Patent Office Interference 69,406; U. S. Patent 2,057,640 (1927–36) to Conrad.
[91] U. S. Patent 1,977,439 (1929–34) to Hammond.
[92] U. S. Patent 1,941,069 (1933–33) to Armstrong.