RCA REVIEW

A Quarterly Journal of Radio Progress

Published in July, October, January and April of Each Year by

RCA INSTITUTES TECHNICAL PRESS
A Department of RCA Institutes, Inc.
75 Varick Street, New York, N. Y.

VOLUME V — October, 1940 — NUMBER 2

CONTENTS

SUBSCRIPTION:

United States, Canada and Postal Union: One Year \$1.50, Two Years \$2.50, Three Years \$3.50
Other Foreign Countries: One Year \$1.85, Two Years \$3.20, Three Years \$4.55
Single Copies: 50¢ each

Entered as second-class matter July 17, 1936, at the Post Office at New York, New York, under the Act of March 3, 1879.

Printed in U.S.A.

CATHODOLUMINESCENCE AS APPLIED IN TELEVISION*

BY

H. W. LEVERENZ

RCA Manufacturing Company, Inc., Research and Engineering Department, Harrison, New Jersey

Summary—The cathodoluminescent art is reviewed and some new data are presented not only as general information but also to correct some current misconceptions regarding solid luminescent materials (phosphors).

Synthetic luminescent materials have been known for 337 years, but most of the polychromatic efficient phosphors were painstakingly evolved during research of the past ten years, especially in television research laboratories. Luminescence research is becoming a valuable means of supplying and interpreting new information regarding the physics and chemistry of crystalline matter.

The constitutions and syntheses of the better phosphors are outlined and a simplified theoretical mechanism of phosphor luminescence is discussed in order to provide better understanding of phosphor properties and capabilities. There are eight important qualities, each of which must be possessed in superior degree by phosphors intended for television Kinescope use.

Unjustified restriction or over-emphasis of any one phosphor quality, such as phosphorescence (also known as persistence, retentivity, "after glow" or time-lag), would automatically eliminate most of the phosphors which are excellent in all eight. The choice of 30 frames/second and 60 fields/second is shown to be a minimum repetition rate, below which serious disadvantages are suffered by televiewers. The speculation of using unknown phosphors having concave-downward persistence characteristics to decrease frame and field frequencies, is demonstrated to be untenable.

I. INTRODUCTION

THIS article on the subject of cathodoluminescence is offered so that those in the radio and television art may have an outline of the historical, theoretical and practical features of the "last act" in television's complicated task of seeing at a distance. The "last act" comprises converting modulated electrical impulses and electron currents into visible images which give the sense of uninterrupted continuity and motion. The performers in the "last act" are tiny crystals of specially synthesized luminescent materials which have the unique property of being able to transform electron energy into light.

Certain foibles and fallacies have persisted in the luminescent art, largely due to its alchemical birth and upbringing, and it is hoped that

* A review, including hitherto unpublished data from the RCA Laboratories.

the ensuing factual presentation will assist in dispelling some of the prevalent, inaccurate notions.

The generic term "luminescence" connotes the act of energy absorption with subsequent re-emission as visible and near-visible radiation while the luminescing material maintains a temperature below that required for incandescence. In this respect, the term "cold light" is concisely descriptive. In its original usage, "luminescence" applied to visible radiation only, but for convenience its use has expanded to include the near-visible regions.

Luminescence has been sub-classified according to the types of energy used for excitation. Cathodoluminescence, for example, is light emission occasioned by cathode rays, i.e., electrons, impinging on matter.[1]

A further distinction is made with respect to duration of light emission after cessation of excitation. When the emission is completed within approximately 10^{-8} second, which is the normal interval for isolated excited atoms or ions to return to their ground states, the process is fluorescence. Emission continuing for a longer time than fluorescence is termed phosphorescence. Concomitance of fluorescence and phosphorescence in all but the gaseous state of matter requires the use of the more precise word "luminescence".

The first reported crystalline, inorganic luminescent materials, also called "phosphors", were accidentally prepared over 337 years ago, in 1603.[2] For 283 years subsequent to 1603, the alchemists synthesized phosphors by crude methods such as by heating oyster shells with sulphur to give feebly violet-phosphorescing alkaline-earth sulphide phosphors which were socially ostracised because they decomposed in moist air, evolving hydrogen sulphide. The first efficient (and, incidentally, non-odorous) synthetic phosphor was blue-green luminescing copper-activated zinc sulphide, prepared by Sidot in 1886.[3] Zinc sulphide was used extensively in the first practical application of phosphors as detectors of invisible radiations such as ultraviolet, cathode ray, and X-ray as these new energy manifestations were discovered during the course of the 19th century. Radioactivity's discovery was an accidental by-product of Becquerel's work in unsuccessfully testing a theory of Poincaire who had postulated an intimate connection between X-rays and luminescence.[4]

---

[1] H. Pender and K. McIlwain, Electrical Engineers Handbook, Vol. V, Communications, 2-(49-53), J. Wiley, 1936.

[2] W. Wien and F. Harms, Handbuch der Experimentalphysik, XXIII, Part 1, page 1, Akademische Verlag., Leipzig, 1928.

[3] T. Sidot, "Sur les proprietes de la blende hexagonale", Comptes rendus, *63*, 188-189, 1886.

[4] T. A. Boyd, "Research—The Pathfinder of Science and Industry", p. 164, D. Appleton-Century Co., 1935.

Despite accelerated research on luminescence during the past half-century, phosphor applications remained chiefly of the detector variety and phosphors were usually associated with very low brilliancy values, requiring scotopic or dark-adapted vision for observation of the well-known radium watch dials, X-ray fluoroscope screens and theatrical "black magic".

With the vigorous inception of electronic television, approximately eleven years ago, an urgent need was felt for greatly increasing the capabilities of luminescent materials used in Kinescopes (television cathode-ray [TCR] tubes). The first researchers had but two phos-

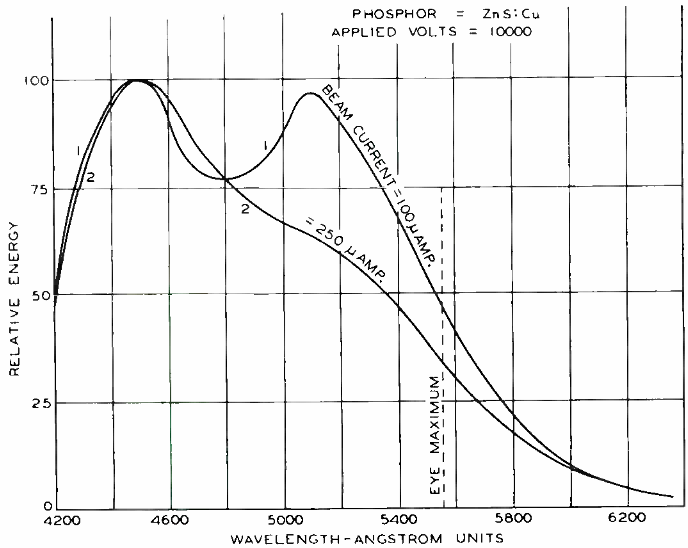


Fig. 1—Emission spectrum of a copper-activated zinc sulphide phosphor, showing the spectral variation with different excitation densities.

phors sufficiently efficient for television purposes: (1) the previously mentioned zinc sulphide, and (2) willemite, a zinc silicate mineral containing about one per cent of manganese silicate. Willemite was first discovered in 1830 and named after King Willem I of the Netherlands.[5] Both phosphors emitted preponderantly green light, giving early television viewers considerable aesthetic dissatisfaction with the reproduced images.

Figure 1 shows the emission spectrum of a commercial luminescent zinc sulphide such as was used in early TCR tubes. The two curves are for the same material under different electron beam current densities, and show the spectral variation which caused an undesirable color

[5] J. W. Mellor, "A Comprehensive Treatise on Inorganic and Theoretical Chemistry", Vol. VI, p. 438, Longmans, Green & Co. Ltd., London, 1930.

change from green to blue as the current density increased. Such a color change is most annoying in television pictures, since the brighter portions of the image show one color and the less brilliant portions another color.

Figure 2 shows performance results for natural vs. synthetic willemite, and indicates the low efficiency of the mineral product compared with a good present-day zinc silicate phosphor.

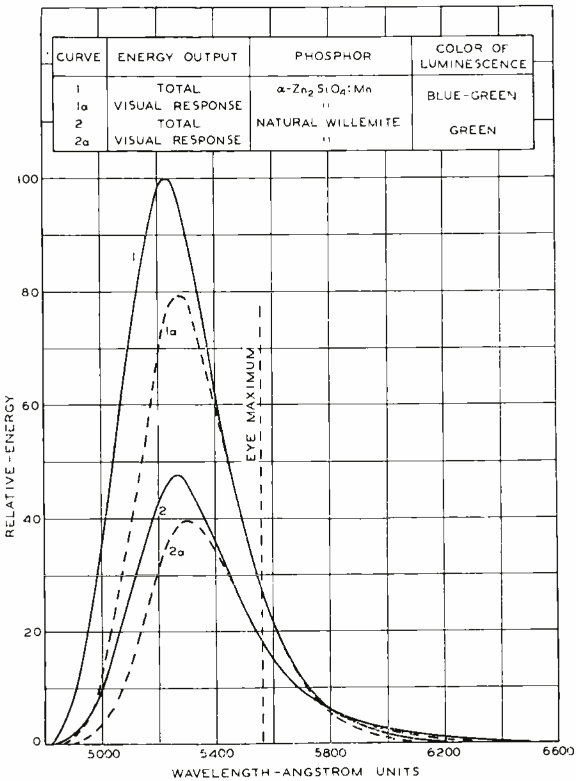


Fig. 2—Emission spectra of natural and synthetic willemite (manganese-activated α-zinc orthosilicate).

There were many more objectionable features about the phosphors known at the outset of electronic television; serious disadvantages such as poor secondary emission, further reduction of initially low efficiencies when the materials were ground and processed for Kinescope application, and a restricted choice of green colors. However, since better phosphors were vital to electronic television, which has the inherent advantage of practically inertialess picture scanning, extensive research was instigated in the RCA Manufacturing Co., Inc., under the sponsorship of Dr. V. K. Zworykin, Director of Electronics

Research, to determine the possibilities of improving phosphors for use in Kinescopes.

An intensive search of the literature on the subject of luminescence disclosed a plethora of phosphor recipes, which, as Dr. Saul Dushman said,[6] "read just like a cook book", but unfortunately most of the recipes were like matches: they worked but once. It appeared certain that the strong green luminescence of zinc silicate phosphor (i.e., willemite) required about one per cent of manganese activator "impurity" and that of zinc sulphide required about one one-thousandth per cent of copper, but reproduction of results was an apparent impossibility unless extraordinary precautions were taken to purify all ingredients to a degree better than "spectroscopic purity".

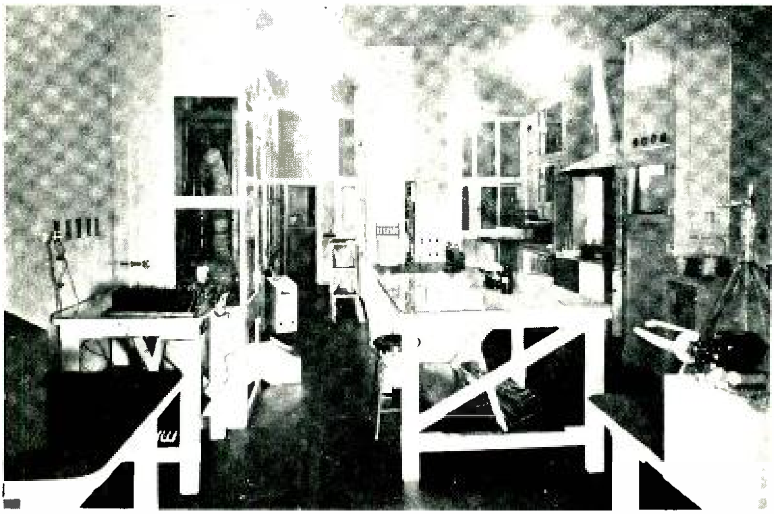


Fig. 3—View of part of the RCA chemico-physics laboratories in which luminescence research is conducted.

Accordingly, air-conditioned laboratories were designed and constructed in the engineering buildings of the RCA Manufacturing Company's Camden, New Jersey plant. Improved laboratories were constructed in the Harrison, New Jersey plant when the chemico-physics research group was transferred to Harrison in 1939. Figure 3 gives a view of one of the Camden laboratories, showing the "hospital operating room" type of simple construction which facilitated thorough cleaning of linoleum walls and floors, glass bench tops, etc.

Thus equipped, it was practical to synthesize the two principal phosphor systems, oxides (especially silicates) and sulphides in many permutations of exceptionally pure substances.

[6] Meeting of the Optical Society of America, New York, October 1936.

BASE MATERIALS

ATOMIC NUMBER	ION	ATOMIC RADIUS A°	IONIC RADIUS A°	GROUND STATE OF ATOM	GROUND STATE OF ION	SULFIDE BEST ACT-IVATOR	SULFIDE CRYSTAL STRUC-TURE	SULFIDE MP $\frac{\Delta H_f}{S}$	OXIDE BEST ACT-IVATOR	OXIDE CRYSTAL STRUC-TURE	OXIDE MP $\frac{\Delta H_f}{O}$	TUNGSTATE (MOLYBDATE) BEST ACT-IVATOR	TUNGSTATE (MOLYBDATE) CRYSTAL STRUC-TURE	TUNGSTATE (MOLYBDATE) MP $\frac{\Delta H_f}{O}$	SILICATE (GERMANATE) BEST ACT-IVATOR	SILICATE (GERMANATE) CRYSTAL STRUC-TURE	SILICATE (GERMANATE) MP $\frac{\Delta H_f}{O}$	ALUMINATE (BORATE) BEST ACT-IVATOR	ALUMINATE (BORATE) CRYSTAL STRUC-TURE	ALUMINATE (BORATE) MP $\frac{\Delta H_f}{O}$	POLAR-IZING POWER OF ION	POLAR-ISABIL-ITY OF ION	ELEMENT
4	Be^{++}	1.13	0.34	$^1S_0$	$^1S_0$		B3			B4	2570° 135					H13		Cr	H12		16.0	0.04	Be
12	Mg^{++}	1.60	0.78	$^1S_0$	$^1S_0$		B1	82.2		B1	2800° 146	—	H4?	345.2	Mn	H12	d.1557 347.5	Cr	H11	2135°	3.3	0.094	Mg
20	Ca^{++}	1.97	1.06	$^1S_0$	$^1S_0$	Bi	B1	113.4		B1	2572° 152	—	H4	387	Mn		1540° 377			1600°	1.8	0.47	Ca
30	Zn^{++}	1.34	0.83	$^1S_0$	$^1S_0$	Cu Ag	B3 B4	1185° 44	—	B4	1975° 85	—	?		Mn	H13	1512° 349	Mn	H11		>3.3		Zn
48	Cd^{++}	1.49	1.03	$^1S_0$	$^1S_0$		B4	780° 34.6		B1	900–1426° 65.2	—	?		Mn		1242°				<3.3		Cd
5	B^{+++}	0.97	0.20	$^2P_{\frac{1}{2}}$	$^1S_0$						93										0.02		B
13	Al^{+++}	1.43	0.57	$^2P_{\frac{1}{2}}$	$^1S_0$				Cr	D51	2050° 127				Cr		d.1810° 139				9.2	0.052	Al
14	Si^{++++}	1.17	0.39	$^3P_0$	$^1S_0$					C8-10	1470–1713° 101.5										26.0	0.0165	Si
32	Ge^{++++}	1.22	0.44	$^3P_0$	$^1S_0$					C8	1713° 64												Ge
74	W^{++++++}	1.37	0.68	^5D																			W
8	O^{--}		1.32	$^3P_2$																		3.88	O
16	S^{--}		1.74	$^3P_2$																		10.2	S

ACTIVATORS

ATOMIC NUMBER	ION	ATOMIC RADIUS A°	IONIC RADIUS A°	GROUND STATE OF ATOM	GROUND STATE OF ION	SULFIDE CRYSTAL STRUC-TURE	SULFIDE MP $\frac{\Delta H_f}{S}$	OXIDE CRYSTAL STRUC-TURE	OXIDE MP $\frac{\Delta H_f}{O}$	POLAR-IZING POWER OF ION	POLAR-ISABIL-ITY OF ION	ELEMENT
29	$Cu^{+,++}$	1.28	0.96^{+}	$^2S_{\frac{1}{2}}$	$^1S_0$?	? C1	$1100^{\circ ++}$ 11.6^{++}	? C3	$1235^{\circ ++}$ 38.5^{++}	>Ag		Cu
47	$Ag^{+,++}$	1.44	1.13^{+}	$^2S_{\frac{1}{2}}$	$^1S_0$?	?	825° 5.5^{+}	? C3	d.300° 6.95^{+}	<Cd		Ag
83	$Bi^{+++,+++++}$	1.82	0.74^{+++++}	$^4S_{\frac{3}{2}}$	$^1S_0$?	?	685° 26^{+++}	?	$860^{\circ +++}$ 49.5^{+++}			Bi
24	$Cr^{++,+++}$	1.24	0.64^{+++}	$^7S_3$	$^5D_0$ $^4F_{\frac{3}{2}}$	B8 ?	? ?	? D51	$1990^{\circ +++}$ 91^{+++}			Cr
25	$Mn^{++,++++}$	1.29	0.70^{++}	$^6S_{\frac{5}{2}}$	$^6S_{\frac{5}{2}}$ $^5D_0$	B1 C2	d. 47^{++}	B1 C4	1705° 96.5^{++}			Mn

STRUCTURE TYPES

B1 = CUBIC, FACE-CENTERED (NaCl)
B3 = CUBIC (SPHALERITE)
B4 = HEXAGONAL (WURTZITE)
B8 = HEXAGONAL (NICKEL ARSENIDE)
C1 = ORTHOCUBIC (FLUORITE)
C2 = CUBIC (PYRITE)
C3 = CUBIC (CUPRITE)
C4 = TETRAGONAL (RUTILE)
C8 = HEXAGONAL (α-AND β-QUARTZ)
C9 = CUBIC (β-CRISTOBALITE)
C10 = HEXAGONAL (β-TRIDYMITE)
D51 = HEXAGONAL - RHOMBOHEDRAL (CORUNDUM)
H11 = CUBIC (SPINEL)
H12 = RHOMBIC (OLIVINE)
H13 = RHOMBOHEDRAL (PHENACITE)

Table 1

Since sulphides had been originally discovered and developed abroad, European luminescence researchers concentrated mainly on sulphide phosphors. American researchers, because of their use of the naturally-occurring mineral phosphor, willemite, followed by use of the improved synthetic willemite, favored the oxygen-containing phosphors, especially since oxide phosphors are inherently more rugged than sulphides or selenides.

An important product of RCA's television-luminescence research is the zinc beryllium silicate phosphor system which is a major component of the light-emitting coatings used in the new highly-efficient, tubular, luminescent lamps (usually called "fluorescent lamps" despite the need for a considerable phosphorescence in order to minimize flicker). The luminescent art is commencing to expand into the ultraviolet and infrared regions of the spectrum and should employ the unique advantages of the high efficiencies of phosphors, and their easily controllable emission spectra in those invisible radiation ranges.

II. Constitutions and Syntheses of Phosphors

It should be mentioned that, despite unusually favorable conditions for synthesizing luminescent materials, exact reproduction of phosphors is still difficult. Each phosphor sample tends to be individualistic, differing noticeably from identically constituted and similarly synthesized samples in one or more of its properties such as spectral emission characteristic, phosphorescent-time constant, secondary-emission qualities, etc. Attainment of practically identical phosphors is possible, but requires extreme care and extraordinary skill.

Phosphors are both impurity- and structure-sensitive materials. The addition or subtraction of as little as 0.0001 per cent (one part in a million) of a foreign substance can alter some phosphors' properties by 50-100 per cent. Maintaining identical chemical composition, but changing crystal structure by polymorphic transitions also produces equally pronounced changes in some phosphors' characteristics.

The best phosphors are well-crystallized, inorganic materials (termed "base materials" or simply "bases"), usually containing a small trace of one certain metallic salt which is called the "activator". Whereas minute concentrations of some foreign salts greatly enhance the basic crystals' luminescence, similar concentrations of other metal salts, notably those of iron, cobalt, and nickel, "poison" luminescence.

In general, the best phosphors have a relatively colorless bulk crystal which is of the excess-cation type of a high-temperature semi-conductor and contains a very small concentration of a salt of some easily polarizable multivalent element. Table 1 lists the more important phosphor constituents and a few of their pertinent properties.

The dashes in the columns titled "best activator" indicate that an efficient phosphor may be prepared from the indicated base materials without adding an activator. The vertical connecting lines, shown in the same columns, link cations which may be intersubstituted in the particular base material and yet produce a good phosphor. Galliate and selenide phosphors are not included in the table.

Synthesis of phosphors is chiefly chemical work. Obviously, the best available analytical reagent chemicals are much too impure for use in phosphors. Therefore, the chemist must further purify the substances used, add the tiny quantities of activator and perhaps a flux to assist in crystallizing the phosphor. The intimate mixture must then be skillfully heated to produce the crystal size and modification having greatest efficiency and ease of application in cathode-ray tubes.

It is especially remarkable and, from the chemist's standpoint, aggravating, that the limits of chemical purification processes coincide with the order of magnitude of activator impurity usually necessary in efficient phosphors or the magnitude of the "poisoning element" detrimental to phosphors. The coincidence occurs in the range of 10^{-3} to 10^{-8} part of activator, or impurity, to one part of bulk crystal.

The syntheses of phosphors are typified by the following two examples:

A. Synthesis of blue-emitting zinc sulphide phosphor.

Purify zinc sulphate ($ZnSO_4$) by conventional chemical methods until no spectrographically detectable impurities remain. Electrolyze the aqueous zinc sulphate solution to remove any copper, manganese, and lead which the spectrograph may not have indicated. Precipitate pure zinc sulphide with well-washed hydrogen sulphide, and wash the precipitate. $ZnSO_4 + H_2S = ZnS \downarrow + H_2SO_4$. Add sufficient solution of a silver salt to equal a silver concentration of 0.01 per cent of the weight of the zinc sulphide and further add sufficient sodium or potassium chlorides (in aqueous solution) to equal 2 per cent of the weight of the zinc sulphide. Stir well, while evaporating to dryness and heat in a quartz crucible at 800-1500° C. The time and temperature of heating may be adjusted to determine the phosphor's particle size and form. The resultant phosphor is symbolized by $ZnS{:}Ag$; since the alkali halide reacts with the zinc sulphide to form volatile zinc chloride and soluble alkali sulphide which are removed during heating and subsequent elutriation. Cadmium sulphide may be substituted in part for the zinc sulphide to alter the phosphor's spectral emission over the entire visible spectrum and into the infra-red.[7]

[7] H. W. Leverenz and F. Seitz, "Luminescent Materials", J. Applied Physics, *10*, 7, pp. 479-493, 1939.

B. Synthesis of yellow-green emitting zinc beryllium silicate phosphor.

Zinc and beryllium nitrates ($Zn(NO_3)_2$ and $Be(NO_3)_2$) are purified and mixed in aqueous solution such that the ratio of zinc to beryllium is approximately nine to one on a gram-molecular-weight (mole, or molar) basis. Approximately 0.006 mole of pure manganese nitrate ($Mn(NO_3)_2$) is added per mole of zinc plus beryllium. Very pure, finely divided silica (SiO_2), such as colloidal silica or a substance such as an organic silicate is added to the nitrate solution and the carbonates of zinc, beryllium, and manganese precipitated around the silica by adding ammonium carbonate.

The amount of silica added may be exactly ortho-proportion or up to several hundred per cent over ortho-proportion. "Ortho-proportion" is two moles of (zinc + beryllium) to one mole of silica.

Stir and evaporate to dryness and heat in a clean platinum crucible at 900-1600° C depending on the degree of chemical combination, crystal type and size required. A shorthand notation for the finished phosphor is $ZnO_u{:}BeO_v{:}SiO_{2w}{:}Mn$. In this example, $u/v \cong 9$ and $w \geqq (u+v)/2$, but u, v, w, and the Mn concentration may be varied to produce a wide variety of emission colors and other phosphor characteristics.[8]

Mechanical mixtures of certain blue-emitting and yellow-emitting phosphors, prepared as described in A and B, will give a resultant white light under cathode-ray excitation.[9]

III. Theories of Phosphor Luminescence

There is no theory of luminescence adequate to explain quantitatively all the properties of known phosphors or to predict the properties of new phosphors.

All efficient phosphors are definitely crystalline. See Figure 4 for examples of the regular arrays required in order to have efficient phosphors. The attainment of an ordered state is evidently necessary to provide a minimum of traffic obstruction to electrons liberated in the crystals. In view of the high efficiencies obtained, especially with corpuscular excitation, it appears that the bulk crystal, as well as the "centers" associated with the small concentration of the activators, absorbs the radiant or corpuscular exciting energy and redistributes it in smaller, more digestible packets of low-velocity free electrons or excitons (electron-hole pairs). The liberated electrons or excitons

---

[8] See Figs. 24, 25, and 26 of reference 7.

[9] H. W. Leverenz, "Optimum Efficiency Conditions for White Luminescent Screens in Kinescopes", J.O.S.A., *30*, 7, 309-315, 1940.

travel through the crystal for considerable distances from their origins and eventually return to their own or, more probably, to other centers which are capable of transforming the energy into luminescent emission. The actual emission act is probably performed by a neutral (non-ionized) atom or by a negatively charged ion associated with a multivalent activator center. The centers may be visualized as loosely bound units regularly distributed throughout the crystal lattice, substituted in place of lattice units or located in places where lattice units are missing, or else associated with crystal faults.

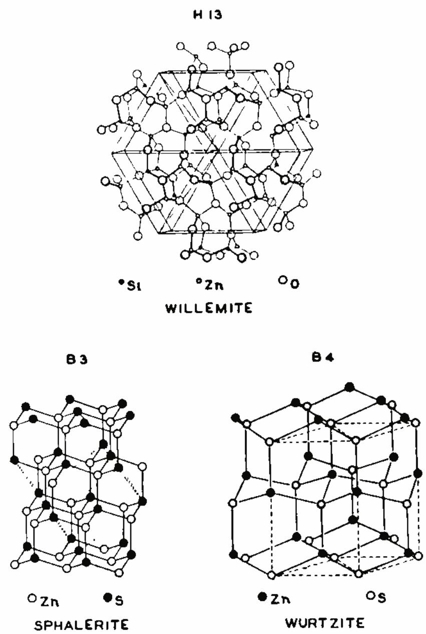


Fig. 4—Crystal structures of α-zinc silicate (willemite), α-zinc sulphide (wurtzite), and β-zinc sulphide (sphalerite).

A crude description of the *modus operandi* of a phosphor is the following: Imagine a three-dimensional lattice-work of elastic bands joining together a regular array of identical bells. This corresponds to the basic crystal of a phosphor. Dispersed at regular intervals throughout the ordered structure, in a concentration of one to a thousand, there are much smaller bells substituted for, or suspended between, bells of the main bell network. These smaller bells are the activator centers. It is apparent that considerably less force will be required to cause the smaller bells to sound than the larger bells. It is also apparent that the smaller bells may be rung by either direct application of energy or by receiving energy which has been transshipped through the elastic bands after being absorbed by some larger unit. The reason for the greater luminescent efficiency of the crystalline state versus the amorphous state is deducible from the model. Energy transfer through elastic bands having widely varied degrees of tension would be short-lived because the different bands would not pass the same frequencies.

In order to give a more tenable picture of the action corresponding to free electron liberation in the phosphor crystal, it would be necessary to imagine that the clappers of the bells could become detached and slide along the joining bands until they encountered a small bell which would ring, whereas the momentum of the clapper was insufficient to ring a larger bell. The absorption of energy and re-emission of sound entirely by a single small bell represents fluorescence. Absorption of energy by any unit of the lattice-work with subsequent transmittal to a small bell, possibly far removed, which emits the eventual sound represents phosphorescence. The distinction between the two luminescence acts is seen to be primarily one of localization versus decentralization and of time required for energy transport. Vigorous jangling of the entire structure would cause the main lattice bells to sound and disturb the more sensitive efficiencies of the smaller bells. The model thus portrays the effect of incandescence in phosphors.

Unfortunately, the bell model fails in several respects in simulating the operation of an actual phosphor. For example, it allows high-amplitude, low-frequency vibration to ring the small bells which emit higher frequencies than were possessed by the exciting energy. This is opposite to phosphor action as expressed in Stoke's law, "The emitted light is of a longer wavelength than the exciting radiation". Since frequency, ν, wavelength, λ, and speed of propagation, c, are related by

$$c = \lambda\nu$$

it is seen that the bell model violates Stoke's law by absorbing low-frequency long wavelength energy and emitting high-frequency, short-wavelength sound. Exceptions to Stoke's law are unimportantly rare.

The model correctly portrays the independence of a phosphor's emission spectrum with respect to means of excitation and time during or after excitation. This is true only when the phosphor's emission is a single band, since more than one band would indicate different centers (different small bells in the model) which usually vary greatly with respect to excitation-saturation and decay rate. In the latter case, the total emission color of a phosphor changes markedly during phosphorescence, while the former case has been demonstrated in Figure 1.

Refinements and ramifications of the foregoing mechanical simile provide stimulation for experiment, yet fail to depict the complex atomic dynamics of real phosphors because the actions within phosphor crystals involve the vaguely comprehended transition zone between corpuscular and undulatory energy.

Foreign elements, such as iron, nickel, etc., previously classified as phosphor poisons are deleterious by virtue of: (1) occupying positions which might otherwise be advantageously occupied by the luminescence activator units, (2) absorbing energy and then emitting radiation in an invisible (viz. ultraviolet or infra red) region of the spectrum, and (3) decreasing phosphorescence by absorbing transshipped energy more readily than the luminescence centers and thus

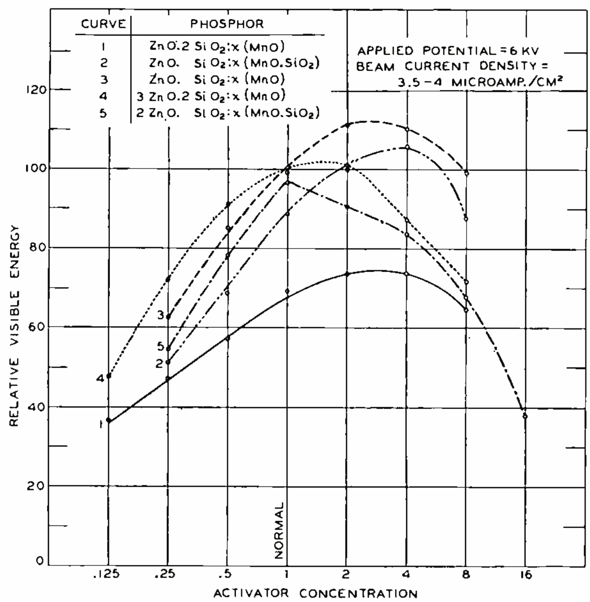


Fig. 5—Luminescence efficiency of α-zinc orthosilicate as a function of composition and manganese activator concentration.

destructively diminishing the amount of potential energy stored in the phosphor crystal. Use of more than one activator in a phosphor fails to increase efficiency since each activator is occupying positions which might be used by the other.

There is an optimum concentration of activator, but it is not critical. There is no lower limit or threshold value, except possibly as expressed in terms of the human eye's sensitivity. The completely dark-adapted eye requires a minimum of approximately 17 x 10^{-10} ergs/sec. visible radiation through the pupil for recognizable stimulation.[10] Increasing the activator concentration above the optimum reduces efficiency by: (1) exceeding the number of suitable faults or interstitial

[10] LeGrand Hardy, "Eye as Affected by Illumination", Am. Illum. Soc., Trans., *29*, pp. 364-384, 1934.

positions available for activator units in the basic crystal, or (2) in the case of isomorphic substitution, allowing the activator units to approach each other so closely that they produce mutual interference. The distance from center to center of manganese ions (assumed homogeneously distributed) in an α-zinc silicate phosphor with optimal activation is 9.08 Å. Using the ionic radius value $Mn^{++} = 0.91$ Å, it is found that the distance of closest approach is $9.08 - 1.82 = 7.16$ Å.

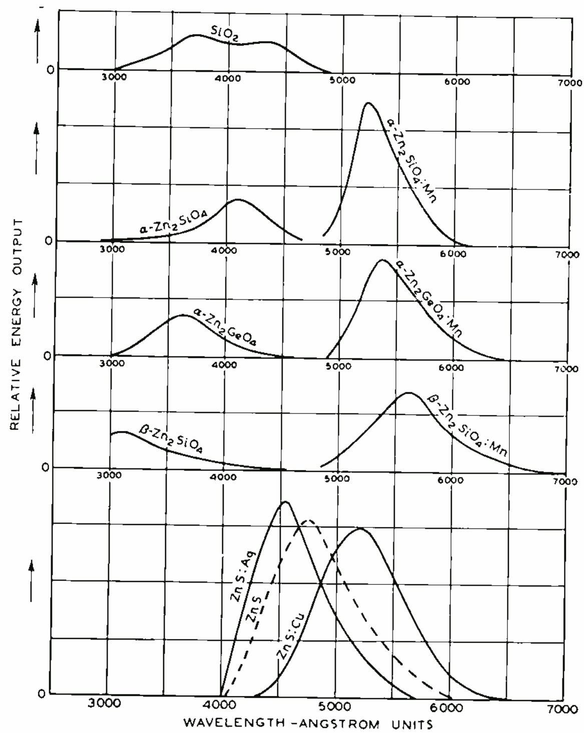


Fig. 6—Luminescence emission spectra of some activated and unactivated phosphors. The unactivated substance's emission disappears when the activator is present, and a new band having greater energy efficiency appears in the visible spectrum.

Figure 5 shows how variation of activator content affects efficiency in the α-zinc silicate system. It is noticeable that the optimum concentration of manganese increases with increasing silica concentration in the initial composition of the phosphor.

Pure, presumably unactivated, crystals also luminesce. In fact, under cathode-ray bombardment all materials luminesce. The luminescence emission spectra of several pure crystallized sulphides and silicates are shown in Figure 6 contrasted with activated phosphors of the same substances. The ordinate scales of the various curves, including those of the same phosphor (activated and unactivated), are not drawn to relative scale.

Hitherto unpublished results, obtained in the RCA Laboratories, show the emission spectrum of pure SiO_2 (crystallized at 1300° C) to be located in approximately the same ultraviolet region as that of the pure silicates (crystallized at 1100-1300° C) of zinc, magnesium, calcium, cadmium, strontium, and barium. In the foregoing pure-substance phosphors the emission mechanism is thus determined by the *Si-O* bond in the crystal lattice while the metal cation (*Zn, Mg, Ca, Cd, Sr,* or *Ba*) has very little effect. This result is interesting, in that it is quite the opposite of the case for the same silicates activated with manganese. The manganese activator is therefore associated with the cation lattice positions whereas the pure crystal's emission centers are located in the $Si\text{-}O_x$ radicals or chains.

The emission spectrum of pure unactivated beryllium silicate phosphor appears at considerably shorter wavelengths than that characteristic of the silica and silicates listed above. The highly polarizing beryllium ion has a much smaller radius ($Be^{++} = 0.34$ Å) than any of the previously listed cations and is smaller than the silicon ion ($Si^{++++} = 0.39$ Å). In the case of beryllium silicate, it seems as logical to call the substance silicon berylliate since the beryllium-oxygen linkage is stronger than the silicon-oxygen binding.

Pure, unactivated β-zinc silicate, which is formed by quenching molten zinc silicate,[11] has its emission spectrum located at shorter wavelengths than that of the normal unactivated α-zinc silicate. Similarly, the emission spectrum of unactivated α-zinc germanate is at shorter wavelengths than that of pure α-zinc silicate. However, the emission spectra of manganese-activated β-zinc silicate and α-zinc germanate are located at longer wavelengths than that of α-zinc silicate. The spectrum shifts of the activated compared with the unactivated materials are seen to be in opposite directions. Evidently the binding forces of the luminescent-active optical electrons associated with the $Si\text{-}O_x$ groups are increased by expanding the lattice from α- to β-zinc silicate and similarly by expanding the lattice through substitution of germanium for silicon. The increased binding force may be due to the diminishing of the cation's polarizing influence by increasing the distance between the cations and the $Si\text{-}O_x$ groups or chains. Since the activator units are located at cation positions, the lattice expansion weakens the binding forces of the activator's valence electrons. It appears from these results that valuable information regarding strengths of crystal lattice bonds

[11] See page 489 of reference 7.

and their directivities may be gained by further studies of the emission spectra of cathode-ray excited substances.

There are many more interesting, though apparently anomalous, phosphor emission-spectrum shifts which have been observed in the course of our research work, but these must be withheld for future publication, since their discussion is not suited for this review.

Figure 7 shows the efficiency of manganese-activated α-zinc silicates superimposed over the phase diagram of the zinc-silica system. There is only one true compound formed, as shown by the single melting point maximum at 1512° C at the ortho-proportion of 2 $ZnO - 1\ SiO_2$. The

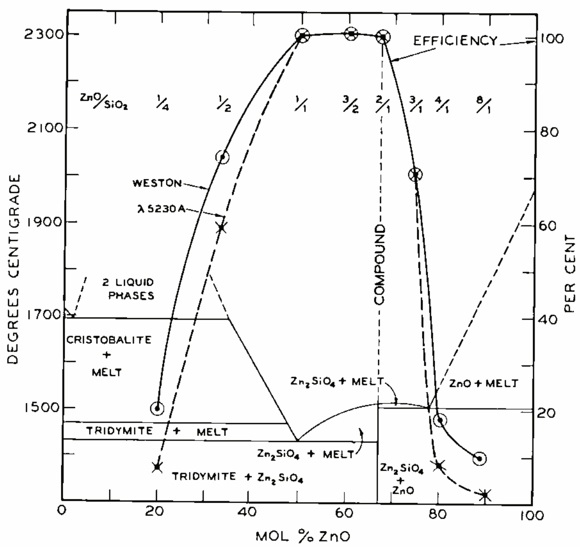


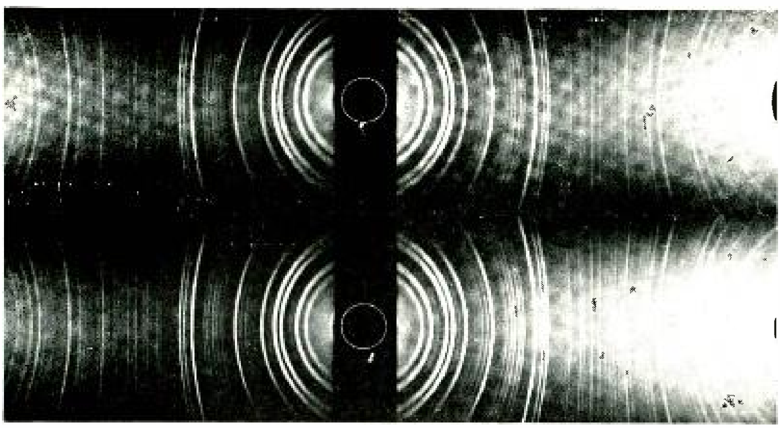
Fig. 7—Luminescence efficiency of manganese-activated α-zinc silicate as a function of composition, with constant ratio of manganese to zinc.

sharp drop of efficiency on the excess zinc oxide side, as contrasted with the much slower decrease on the excess silica side, coupled with the pertinent information given in connection with Figure 5, has led the writer to propose a "deficiency structure"[12] explanation concerning α-zinc silicate's ability to use higher activator concentrations efficiently when the silica concentration is increased.

Figure 8 shows some X-ray powder photographs which were made by Professor B. E. Warren of the Massachusetts Institute of Technology. The photographs show no structure change in increasing the ZnO/SiO_2 ratio from that of the compound (2/1) to 100 per cent

[12] H. W. Leverenz, "Relative Emission Spectra of Zinc Silicates and Other Cathodoluminescent Materials", Paper #30, American Physical Society meeting, Washington, D. C., April 28, 1938.

excess SiO_2 (1/1). The orthosilicate (see Figure 4) phenacite-type structure persists despite the inclusion of a large excess of silica. It seems logical to propose that the silica excess continues to build a normal orthosilicate structure with the exception that some of the zinc and corresponding oxygen units are lacking. The resultant lattice, then, instead of being named an α-zinc orthosilicate with excess silica, should be called an α-zinc orthosilicate with a deficiency of zinc oxide. The distinction is important, for the absent zinc oxide positions are partially filled by manganese oxide activator units and the structure can therefore efficiently use a higher concentration of manganese activator. All the compositions, even those with several hundred per cent excess silica, are still orthosilicates despite the deviation from stoichiometric ortho-proportions. Luminescence research has valuable potentialities in



(a) α-$2ZnO{\cdot}SiO_2$. 1200 °C — 2 Hours
(b) α-$ZnO{\cdot}SiO_2$. 1200 °C — 1 Hour

Fig. 8—X-ray powder diffraction photographs of (a) ortho-proportion α-zinc orthosilicate, and (b) ortho-proportion α-zinc orthosilicate containing 100 per cent excess silica.

disclosing some of the newer facts of crystal chemistry, a study essentially different from that of the conventional chemistry of solutions and gases.

IV. Phosphor Properties and Application of Phosphors in TCR Tubes

Good phosphors must meet numerous requirements for use in TCR tubes.[13] The necessary qualifications may be divided into two groups according to whether they are: (1) objective (independent of the seeing act), or (2) subjective (directly related to the processes of seeing).

A. *Objective Qualities:*

(1) *Ease of applying phosphors to form TCR tube screens.*

---

[13] See reference 7.

This subject has been given considerable discussion elsewhere.[14] The chief objective from the phosphor standpoint is to produce a non-aggregated, smooth-flowing phosphor powder having a narrowly-limited crystal size whose average magnitude is best suited for the method of screen application, provides good screen adherence and gives sufficient screen contrast. The crystal size is best controlled by the crystallization process, since grinding of phosphors seriously reduces

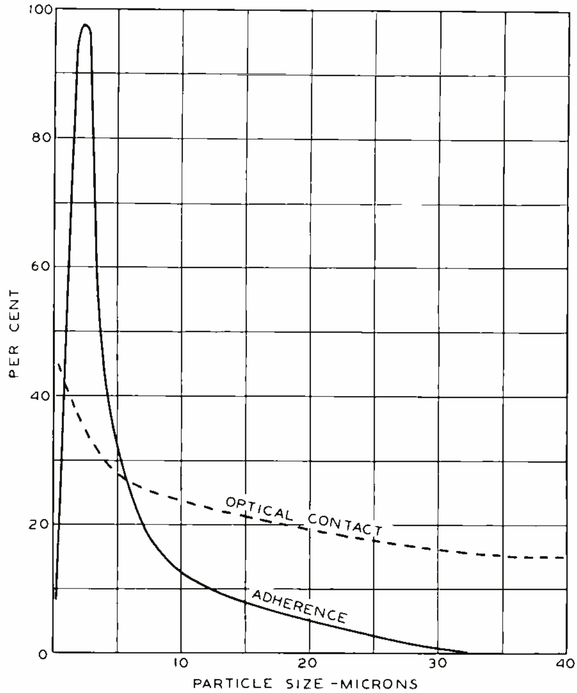


Fig. 9—The effect of phosphor particle size on optical contact and adhering power.

their efficiencies and stabilities. Table 2 shows the effect of grinding an α-zinc silicate phosphor.

Table 2

Effect of grinding upon performance of $\alpha\text{-}Zn_2SiO_4:Mn$

Hours grinding	Original efficiency
0	100%
16	80
24	66
64	36

[14] H. W. Leverenz, "Problems Concerning the Production of Cathode-Ray Tube Screens", J.O.S.A., *27*, 1, 25-35, 1937.

3) *Secondary emission of phosphors.*

Phosphors are good insulators, usually having resistivities greater than most glasses ($10^{12} - 10^{16}$ ohm cm. at room temperature). The negative charge imparted to the screen by the exciting electron beam cannot be dissipated effectively by conduction through the crystals to the accelerating-anode coating, A_2, in the Kinescope (Figure 11), but must be maintained at a low value by emission of secondary electrons from the phosphor, S. The secondary electrons are attracted to the anode coating as long as the collector voltage is positive with respect to the phosphor crystals, or only slightly negative, within the voltage range corresponding to the emission velocities of the secondary electrons (0 to approximately $-10\ V$).

It is very important that a phosphor's secondary-emission ratio (ratio of emitted secondary electrons to incident primary electrons), be unity or greater for the particular voltage applied to the Kinescope. Should the ratio be less than unity, the potential on the screen will decrease with respect to the applied voltage until unity ratio is established or, failing to attain a ratio of unity or greater, the screen potential will fall to cathode potential so that no further current can reach the luminescent screen.

Figure 12 shows a secondary-emission vs. voltage curve representative of phosphors and insulators in general.

Practically, the important secondary-emission characteristics are the "limiting potential"[20], V_L, and "deviation angle", θ. The limiting potential is the applied voltage corresponding to the second unity-crossover (V_L) of Figure 12, and is the point beyond which further increase of applied voltage produces less than linear increase in the actual potential of the phosphor-coating. This potential determines the velocity of the impinging electrons. Figure 13 shows typical plots of some phosphor-coating potentials, relative to applied voltage, indicating the limiting potentials.

The "deviation angle", θ, represents the degree of the screen potential's non-linear conformity beyond the limiting potential. Use of applied voltages greater than the limiting potential gives a greater gain of accelerating potential the smaller the value of θ. The higher applied voltages have advantage even in the case of large values of θ,

[20] W. B. Nottingham, "Electrical and Luminescent Properties of Willemite under Electron Bombardment", J. Appl. Phys. *8*, 762-778, 1937.
H. Nelson, "Method of Measuring Luminescent Screen Potential", J. Appl. Phys. *9*, 592-599, 1938.
W. B. Nottingham, "Electrical and Luminescent Properties of Phosphors under Electron Bombardment", J. Appl. Phys. *10*, 73-83, 1939.
S. T. Martin and L. B. Headrick, "Light Output and Secondary Emission Characteristics of Luminescent Materials", J. Appl. Phys. *10*, 116-127, 1939.

in that the beam current density may still be increased somewhat independently of the screen potential.

Both V_L and θ are greatly affected by extraneous influences such as Kinescope screen composition and thickness, residual gas, evaporated material from heated tube parts or getters, and deleterious effects of continued electron bombardment during tube life.

Efficient phosphors belong to the group of colorless insulators having good secondary emission. Bruining and deBoer[21] have outlined the

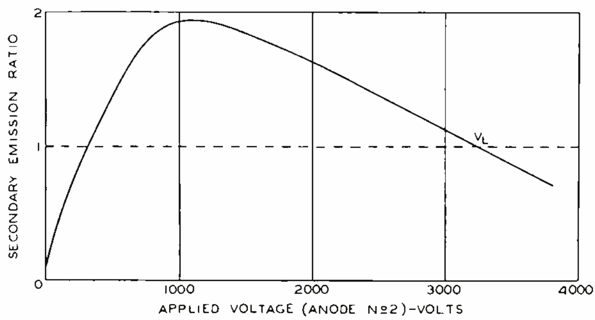


Fig. 12—Typical secondary-emission characteristic of insulators.

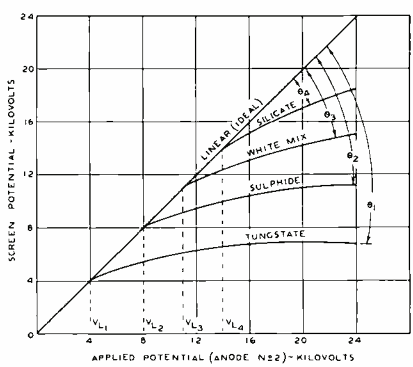


Fig. 13—Screen potential of phosphor screens as a function of applied potential.

conditions favorable for high secondary emission. They specify that the red limit of the external photoelectric effect of a material should correspond to the first absorption band on the red side of the absorption spectrum.

The quantitative influences of structure variations and changes in compositions of phosphors, as affecting V_L and θ, have yet to be

[21] H. Bruining and J. H. deBoer, "Secondary Emission", Parts IV and V, Physica, VI, *8*, 823-840, 1939.

investigated. With the use of higher voltages, over 20 kilovolts for projection Kinescopes, it becomes important to know more about possible means of increasing V_L and decreasing θ without resorting to extraneous devices such as supplying a conducting coating under the phosphor screen.

4) *Stability of phosphors.*

Phosphor centers are intrinsically delicate, as indicated by their sensitivity to exciting radiation of the order of 2.5 electron volts energy or more (5000 Å or less). Luminescence may be excited in phosphors with six-volt electrons, whereas television cathode-ray tubes are operated at several kilovolts. The possible destructive thermal agitation occasioned by a 10-kilovolt electron striking a phosphor center, assumed as an isolated atom, is indicated by the "temperature" which the atom would attain were the entire energy of the 10-kilovolt electron absorbed by the atom. This "temperature" would be

$$T = 1/2\, m\, v^2 / 3/2\, k = \frac{1.57 \times 10^{-7}}{2.06 \times 10^{-16}} = 7.7 \times 10^8 \ {}^\circ\mathrm{K}$$

where

m = mass of electron
v = velocity of electron
k = Boltzman's constant $= 1.371 \times 10^{-16}$ erg deg^{-1}

The absorption of energy by an atom's immediate neighbors in a crystal considerably reduces the value of T. Energy is usually subtracted in small (approximately 30 electron volts) quantities from the swiftly moving primary electron; the probability of an absorption act being inversely proportional to the electron's volt velocity. These considerations, nevertheless, do not detract from the 10-kilovolt electron's potential destructive effect.

The underlying differences between the two most important phosphor species, sulphides (including selenides) and oxides (including silicates, tungstates, borates, etc.), are the differences between sulphur and oxygen with respect to their combining affinities. Oxygen and sulphur are both members of group 6B of the periodic system[22] and are, therefore, chemically equivalent. The principal difference between oxygen and sulphur (or selenium) is in the physical size of their atoms or ions. Table 3 lists the atomic and ionic radii as well as other physical data of oxygen and sulphur.

---

[22] H. W. Leverenz, "A Convenient Periodic Chart of the Elements", Foote Prints, *12*, 1, 22-24, 1939.

Table 3

	Atomic Radius A	Ionic Radius A	Melting Point °K	Boiling Point °K	Deformation (Polarization) of the ion	First Ionization Potential el. v.	Electron[23] Affinity
Oxygen	0.60	1.32 (O−)	54	90	3.88	13.56 (O+)	+3.8
Sulphur	1.04	1.74 (S−)	402	717	10.2	10.3 (S+)	+2.1

It is to be expected that sulphides and selenides will have less resistance than oxides have to decomposition under the so-called "burn-

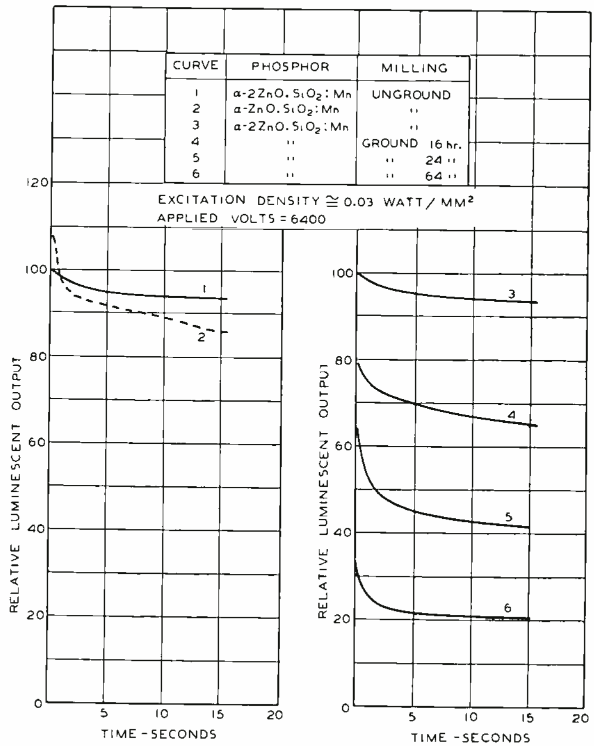


Fig. 14—"Burning" of manganese-activated α-zinc orthosilicates as affected by composition and degree of comminution.

ing" action of an electron beam. Chemically speaking, an electron is a reducing agent except when its velocity is such that it ejects two or more electrons from an atom or ion, in which case it is an oxidizing agent. The same relative stabilities of phosphors obtain with respect to other injurious actions such as comminution, exposure to moisture, light, air, tube processing (exhaust and baking), and high operating temperatures (viz. as encountered in projection Kinescopes) as well as contamination or "poisoning".

[23] G. Glockler, "Estimated Electron Affinities of the Light Elements", Phys. Rev. *46*, 111-114, 1934.

The effect of utilizing a bulk crystal having greatest stability is illustrated by the "burning" tests shown in Figure 14. It is seen that the manganese-activated α-zinc silicate composed of ortho-proportions ($2\ ZnO{\cdot}SiO_2$), corresponding to the compound's melting point (1512° C) in the phase diagram of the ZnO-SiO_2 system (see Figure 7), has less initial decrease of efficiency under intense electron bombardment and less rapid decay of efficiency on continued bombardment than is true of the hypothetical meta-proportion ($ZnO{\cdot}SiO_2$), corresponding to the lowest melting point (1437° C) in the phase diagram. The remainder of Figure 14 shows that decreasing the bulk crystal's stability by grinding, which increases the concentration of strains and faults as well as increasing the ratio of surface tension to lattice energy, greatly reduces a phosphor's ability to withstand "burning".

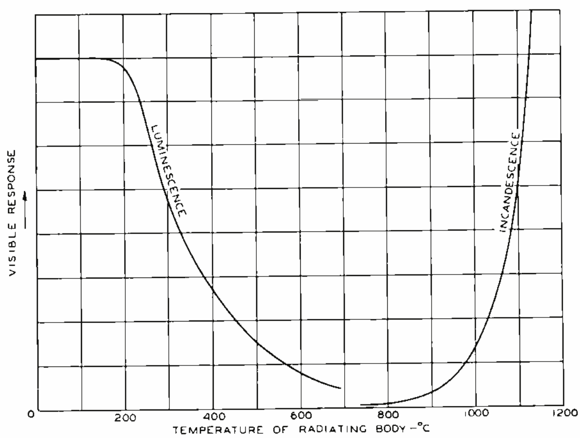


Fig. 15—Typical temperature dependence of luminescence and incandescence efficiencies of a phosphor.

5) *Heat and infra-red effects on phosphors.*

The general effect of temperature on a phosphor's luminescence efficiency is shown in Figure 15. When the bulk lattice elements of a phosphor become agitated to the extent that they unduly "jostle" the sensitive activator centers, luminescence ability is lost. Further increase in temperature eventually produces incandescence, which is light emission occasioned by the excitation energy mutually imparted by violent oscillation of neighboring lattice elements. The ordinate scales for the two curves of Figure 15 are not drawn to a common scale.

At very low temperatures, viz., liquid air temperature, phosphorescence may be frozen-in, or stored, for later release by applied heat or infra-red radiation. The storing process comprises trapping free electrons (detached bell clappers in our previous description of phosphor action) near activator centers and in crystal faults. A definite quantity of thermal agitation is required to re-liberate the electron that it may wander to either another trapping location or to a suitable luminescence activator center where it may excite light emission. Electric or magnetic fields may also be used to effect dislodgement of trapped electrons.[24]

The maintenance of phosphor efficiencies at elevated temperatures is approximately proportional to their stabilities as discussed in the preceding section. That is, oxide and silicate phosphors will generally operate efficiently at higher temperatures than sulphide and selenide phosphors. Each phosphor has an optimum operating temperature which depends not only on the phosphor, but also on the intensity of excitation.

As an example of the vital role played by the bulk lattice in phosphor stability with respect to temperature, the following description is given of a 1935 experiment with a projection tube having a screen of yellow-luminescing β-zinc silicate phosphor. Several spots on the phosphor coating were heated with a 500 microampere, 10,000 volt electron beam held stationary until some of the β-zinc silicate had been converted to the green-luminescent α-form. Starting with a 2.2 cm^2 pattern, these spots were scanned over a decreasing area so as to increase the power input per unit area. When the pattern area was decreased to about 1.5 cm^2, the green spots "burned" and appeared black. Further reduction of the pattern area to less than 0.4 cm^2 showed that the β-zinc silicate increased in brilliancy without discernible "burning". It was only when the beam was concentrated into such a very small area that the generated heat raised the temperatures of the tiny crystals above 900° C that the β-zinc silicate reverted to the α-form and "burned". During this experiment, the scanned area became noticeably more efficient than the unbombarded area outside the scanned pattern. This fact was determined by occasionally expanding the scanned area for momentary observations of the relative brilliancies of the scanned and the previously unscanned phosphor.

B. *Subjective Qualities:*

6) *Emission spectra of phosphors.*

The absorption spectra of phosphors are not discussed in this paper, since cathode rays are capable of providing energy in effectively any spectral region down to the wavelength of complete conversion as given by the equation

$$\lambda_{\text{min.}} = \frac{1.234 \times 10^4}{\text{electron volts}} \text{ Å}$$

---

[24] See pages 263-279 of reference 2.

Ultraviolet sources are less versatile than electron excitation and must be chosen to produce energy within the individual absorption band of any specific phosphor. Thus, silicate, borate, and tungstate phosphors respond efficiently to low-pressure mercury discharges (predominantly 2537 Å) while sulphide phosphors respond efficiently to high-pressure mercury discharges (largely 3650 Å).

In commencing a discussion of emission spectra of phosphors, it is now necessary to introduce the subjective aspect by consideration of the process of seeing. The eye is a very selective receiver of radiation, sensitive only in the narrow band of 3800-7200 Å (approximately one octave) with a maximum of 5560 Å, as shown in Figure 16. Phosphors intended for Kinescope use must have their emissions located well

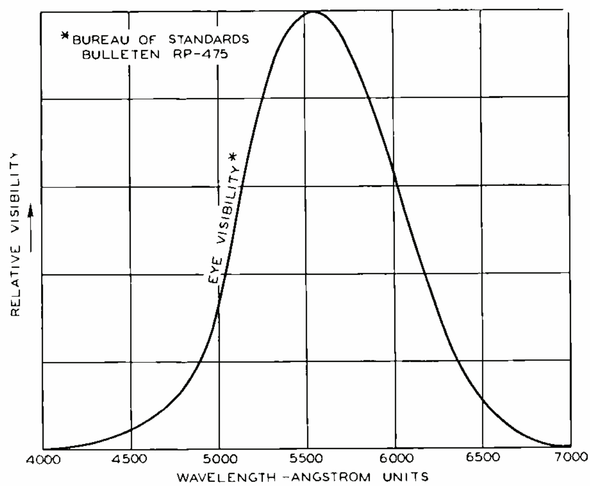


Fig. 16—Sensitivity curve of the human eye.

within the visible region and preferably near the wavelength of maximum visual response (5560 Å) if optimum efficiency is desired. Thus, many phosphors having very high energy efficiencies in the visually ineffective spectral regions are unsuitable for direct use in Kinescopes.

Seeing is a voluntary process as distinguished from hearing and breathing which are practically involuntary. The ability to interpret visual impressions is developed in each individual just as are the arts of walking and speaking. Seeing is influenced not only by objective factors such as the spectral quality, intensity and duration of light, but also by physiological factors such as fatigue, degree of abnormality of an individual's seeing mechanism, and by psychological factors such as immediate environment and the emotional state of the individual. Individuals differ greatly with respect to their visual impressions of identical objects, since each person's sight is dependent upon his own experience in evaluating color, contrast, brightness, distance, size, aesthetic appeal, etc.[25] It is necessary to have well-weighted averages of a large number of persons in order to formulate a general rule concerning seeing. Similarly, the desirability or undesirability of an object or action as estimated visually must be statistically determined to have value as a representative opinion.

CURVE	PHOSPHOR	TEMP. FOR 2 hr.	RELATIVE VISUAL RESPONSE	NATURAL COLOR	COLOR OF LUMINESCENCE
1	ZnS:0.008% Ag	940°	18.5(1a)	WHITE	LIGHT BLUE
2	ZnS(80) CdS(20):0.01% Ag	"	27.0	LIGHT GREEN WHITE	VERY LIGHT BLUE GREEN
3	ZnS(60) CdS(40): "	"	66.3	VERY LIGHT GREEN	VERY LIGHT CREAM GREEN
4	ZnS(50) CdS(50): "	"	100.0(4a)	LIGHT YELLOW	LIGHT GREEN YELLOW
5	ZnS(40) CdS(60): "	"	63.7	LIGHT CREAM YELLOW	LIGHT YELLOW ORANGE
6	ZnS(20) CdS(80): "	"	9.4(6a)	TAN ORANGE	LIGHT RED
7	CdS:0.02% Ag	"	—	LIGHT BROWN ORANGE	RED

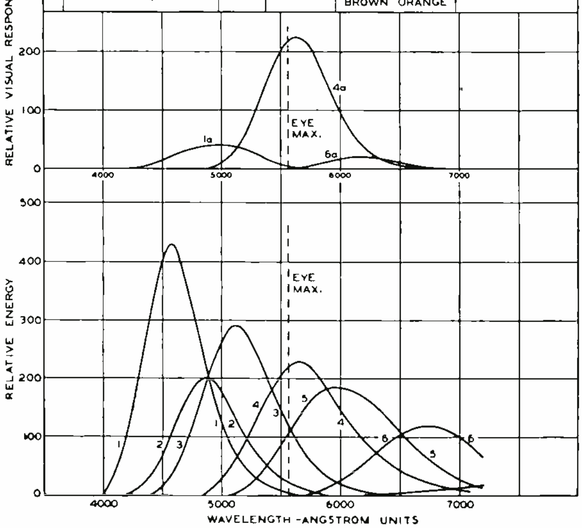


Fig. 17—Relative emission spectra and relative visible efficiencies of silver-activated zinc-cadmium sulphide phosphors.

Decades of acquaintance with printed matter, photographic reproductions and the early motion pictures have instilled a taste for black and white, or black versus some very pale color, rather than black versus a strong hue such as yellow or green. While the demand for white Kinescope screens may be largely traditional, there are some features favoring the choice of white. The contrast of white to black is greater than that of saturated colors to black and the simultaneous

[25] Much of the information in this paper regarding the seeing process is obtained from (a) M. Luckiesh and F. K. Moss, "The Science of Seeing", D. van Nostrand, 1937, and (b) J. P. C. Southall, "Introduction to Physiological Optics", Oxford University Press, 1937.

stimulation of all the color sensations comprising white may be more desirable physiologically than the continued use of but one part of the eye's presumably tri-stimulus mechanism of color vision. Most authorities agree that light is physiologically better the nearer it approaches the spectral quality of diffuse daylight.[26] However, since the eye is a simple lens it cannot focus blue and red in the same plane due to chromatic aberration. Purple, therefore, can never appear distinctly in focus and white should not deviate toward lavender shades if sharp detail is to be observed. Green or yellow shades of white are usually less detectable as "off-white" than are blue or red shades of

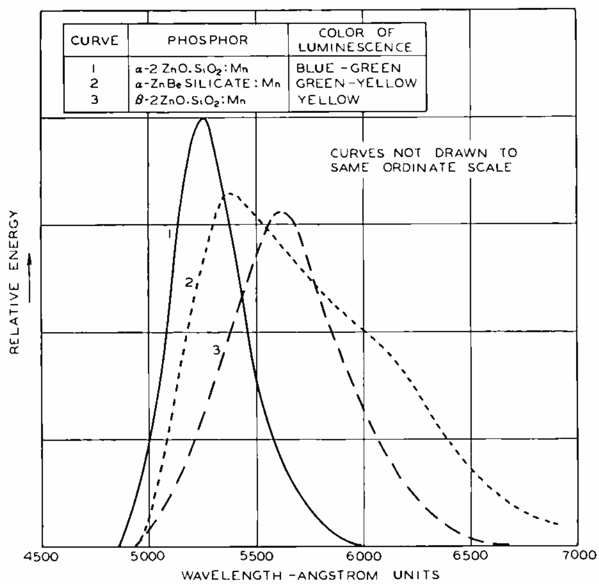


Fig. 18—Relative emission spectra of some manganese-activated silicate phosphors.

white, since the eye's sensitivity to purity (saturation discrimination) is greatest in the blue and red, and least in the green and yellow.

Through centuries of experience the human eye has associated blue-white, viz. daylight, with high illumination levels (200-5000 foot lamberts) and yellow-white, viz. candlelight or incandescent lamps, with considerably lower brightnesses (1-200 foot lamberts).[27] Table 4 shows some data taken from Luckiesh and Moss[28] with additions pertinent to this review. An ideal white light would probably comprise equal energy continuously spread over the entire visible spectrum, but the sensation of white may be produced by but two monochromatic emission

[26] See reference 10.

[27] P. J. Bouma, "Colour Reproduction in the Use of Different Sources of 'White' Light", Philips Tech. Rev., *2*, 1, 1-8, 1937.

[28] See page 325 of reference 25(a).

lines paired in wavelength and relative energies as shown by the complementary white-stimulating pairs $E_{1_1} + E_{2_1}$, $E_{1_2} + E_{2_2}$, etc. in Figure 19.[29]

The emission spectra of phosphors are almost all narrow bands, such as shown in Figures 17 and 18. Only one white-emitting single phosphor has been described and its efficiency is too low to be of commercial importance at present.

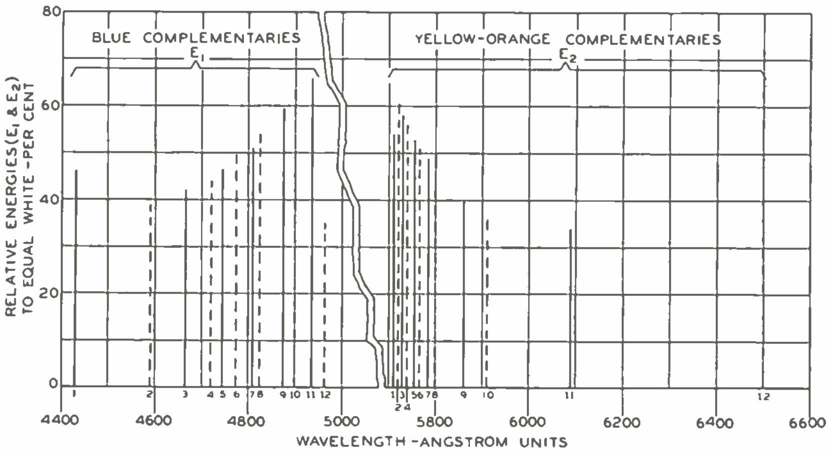


Fig. 19—Relative locations and energy ratios of binary monochromatic white-stimulating spectral complementaries.

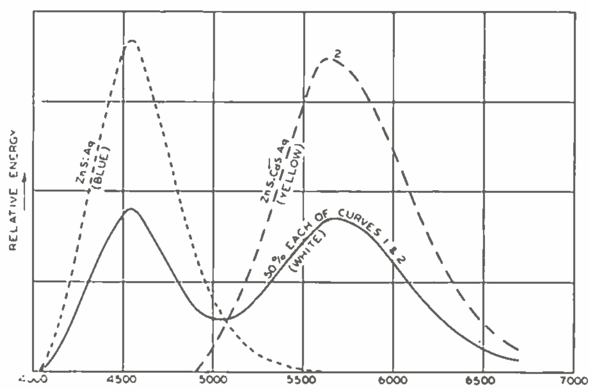


Fig. 20—White-emitting binary phosphor mixture.

Phosphor-emissions are thus usually quite saturated hues and different phosphors must be mechanically mixed to provide paler, i.e., whiter, colors.[29]

Kinescope screens emitting white light are usually composed of mixtures of complementary blue-emitting and yellow-emitting phos-

[29] See reference 9.

Table 4

	Candles/in²	Foot-Lamberts
Outdoors, daylight (December):		
Fresh snow	11.0	5000
Bare ground	0.45	200
White cloth	8.8	4000
Black cloth	0.55	250
Outdoors, night:		
Concrete highway, artificial lighting	0.002	1
Indoors, artificial lighting:		
Ceiling above office lighting unit	0.3	140
Buff wall of same room	0.018	8
Floor of same room	0.005	2
Light sources:		
Sun	1,000,000	450,000,000
Full moon	3.3	1,500
600-watt capillary mercury lamp	285,000	129,000,000
250-watt type H-2 mercury lamp	650	294,000
200-watt tungsten frosted lamp	144	65,000
100-watt tungsten frosted lamp	110	50,000
40-watt tungsten frosted lamp	33	15,000
Candle flame	9.5	4,300
Miscellaneous sources:[30]		
Well-lighted printed page	0.022	10
High-light brilliance on theater screen	0.006–0.012	2.7–5.2
High-light brilliance on 16 mm movie screen	0.006	2.7
High-light brilliance of 12" Kinescope television picture (α-willemite) (6 kv)	0.04	18.2
High-light brilliance of 2¼" x 3" projection Kinescope screen (α-willemite) (15 kv)	1.95	880
High-light brilliance of projection Kinescope projected on a screen 1.5' x 2' (α-willemite) (15 kv)	0.0042	1.9
High-light brilliance of a projection Kinescope using β-willemite (10 kv) and high beam current	13	5,900
Front-surface, zinc sulphide screen, at 70 kv and 0.4 ma[31] in a 0.5 x 0.5 cm² scanned area	3,100	1,400,000

phors as shown in Figure 20. The blue-emitting zinc sulphide phosphor is very susceptible to contamination and easily acquires a green-emission band at the expense of, and in addition to its normal blue-emission if subjected to careless handling or abuse in tube processing. The resultant screens then luminesce very green instead of white. Thus, Kinescope manufacturers sporadically rediscover and are plagued by green zinc sulphide, a phosphor first synthesized in 1886.

In order to produce and maintain a pure white emission color from a composite screen, it is necessary to have the component phosphors accurately matched and, if not entirely stable, at least unstable to the same relative degree. The complementary phosphors should have invari-

[30] V. K. Zworykin and W. H. Painter, "Development of the Projection Kinescope", *Proc. I.R.E. 25*, 938-954, 1937.

[31] K. Scherer and R. Rübsaat, "Helligkeitsmessungen an Zinksulfidschirmen bei Anregung durch Kathodenstrahlen", Archiv für Elektrotechnik, XXXI, 12, 821-826, 1937.

ant individual and relative spectral distributions with respect to the operating range of the completed Kinescope. They must be substantially similar with respect to their secondary-emission characteristics, variations of light output with varied temperature, current density and accelerating voltage, phosphorescence characteristics and, as previously mentioned, their effective particle sizes.

The occurrence of more than one band of emission from a material is indicative of the presence of more than one type of activating center or else of more than one crystal form or chemical combining proportion.

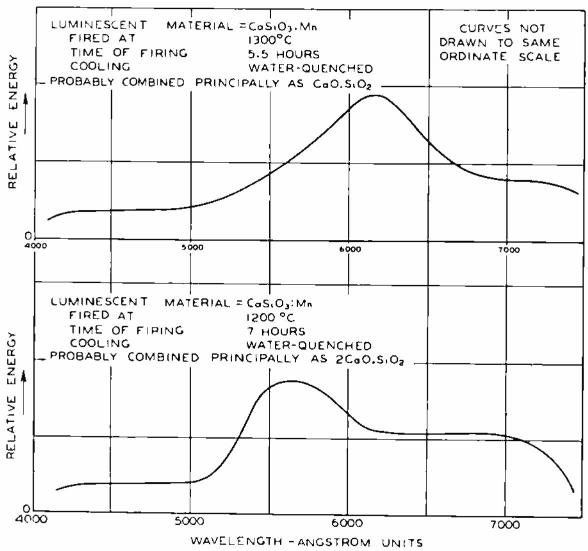


Fig. 21—Luminescence emission spectra of manganese-activated calcium silicate phosphors.

For example, the double band of *ZnS:Cu* (Figure 1) represents the emission of (1) *ZnS:Zn*[32] and (2) *ZnS:Cu* (see Figure 6), while the multiple bands of manganese-activated calcium silicate, shown in Figure 21, are indicative of the various chemical combining proportions of calcium oxide and silica. Luminescence studies are thus uniquely useful as analytical means in determining compositions and constitutions of materials.

Emission spectra of phosphors allow exceptional control and degree of variation and may, therefore, be practically "made to order". Characteristics of invariant optical media (viz. absorption filters) and spectral responses of photoelectric devices may be matched by phosphor

[32] A. Schleede, "Ueber die Ursachen der Luminescenz von reinem ZnS und ZnO", Angew. Chemie. *50*, 908, 1937.
F. Seitz, "Interpretation of the Properties of Zinc Sulphide Phosphors", J. Chem. Physics, *6*, 454-461, 1938.

spectral emission distributions for interesting applications such as in color television.

7) *Brilliancy and efficiency of phosphors.*

Theoretically, enormous brilliancies could be produced with present type, fine-crystal phosphor screens. The following general equation represents the maximum magnitude of luminescence energy output (E_{max}) disregarding phosphor material, but assuming unlimited rate of energy input

$$E_{max} = \frac{ZPE}{t} \frac{\text{ergs}}{\text{cm}^2\ \text{sec}}$$

Z = optimum concentration of activator centers/cm^3 ($\leqq 10^{21}$)
P = penetration distance of the exciting radiation into the phosphor (cm)
$E = h\nu$ = the energy value of the quanta of emitted radiation (ergs)
t = the length of time required by the phosphor to convert the exciting energy into the emitted energy (seconds)

If it is assumed that every luminescent center is "loading and firing" without interruption and that the time of the fluorescent act is 10^{-8} second, and if any loss of emitted light by absorption or scattering is disregarded, the substitution of values for 10,000 volt electrons ($P \cong 0.00025$ cm)[33] exciting light at 5560 Å ($E = h\nu = 3.54 \times 10^{-12}$ erg) yields

$$E_{max} = \frac{10^{21}\ (2.5 \times 10^{-5})\ 3.54 \times 10^{-12}}{10^{-8}} = 9 \times 10^{12} \frac{\text{ergs}}{\text{cm}^2\ \text{sec}}$$
$$= 9 \times 10^{5}\ \text{watts/cm}^2$$
$$= 6 \times 10^{8}\ \text{lumens/cm}^2$$
$$= 6 \times 10^{11}\ \text{foot lamberts}$$

The foregoing analysis is mainly of academic interest, but serves to give an upper limit for light output of conventional phosphor screens. Attainment of such great brilliancy requires practically 100 per cent conversion of excitation energy into light in order to avoid overheating of the phosphor and requires tiny crystals in which the phosphorescence action is negligible compared with fluorescence.

Efficiencies of phosphors are higher than other conventional light sources. Tungsten lamps for home and office lighting purposes have conversion efficiencies of but 2-4 per cent while the best phosphors are approximately 5-10 per cent efficient under cathode-ray excitation and 50-80 per cent efficient under suitable ultraviolet excitation.[34]

---

[33] See page 27 of reference 14.

[34] R. N. Thayer and B. T. Barnes, "Basis for High Efficiency in Fluorescent Lamps", J. Opt. Soc. Am. *29*, 131-135, 1939.

A. Rüttenauer, "Uber die Lumineszenzausbente des Zinksilikat-Leuchtstoffes in der Gasentladung", Zeits. f. techn. Physik, *19*, 148-151, 1938.

Overall efficiency of phosphors depends primarily upon the amount of energy which is usefully absorbed and secondarily upon the quantum deficit relationship. This latter relationship expresses the loss due to energy difference (ΔE) between the exciting ($h\nu_1$) and the emitted ($h\nu_2$) radiation.

$$\Delta E = h(\nu_1 - \nu_2)$$

E = energy
h = Planck's constant, 6.56×10^{-27} erg sec.
ν = frequency of the light (sec^{-1}) $= c/\lambda$
λ = wavelength
c = speed of light in vacuo $= 3 \times 10^{10}$ cm/sec

Since, by Stoke's law, λ excit. $<$ λ emitted, then

$$E \text{ emitted} < E \text{ exciting}$$

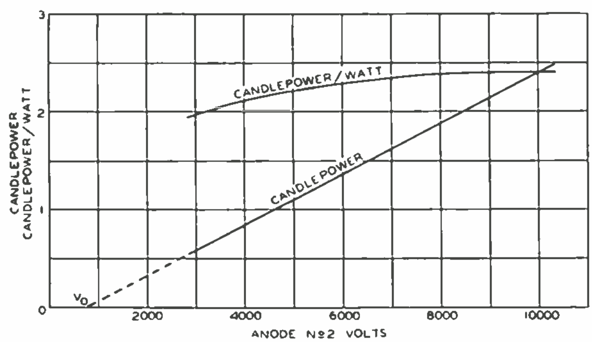


Fig. 22—Typical candlepower and efficiency curves of a phosphor as a function of the Kinescope's applied potential.

As mentioned in section IV-6, cathode rays are capable of providing energy in quantities equal to or less than that given by the equation for complete conversion

$$\lambda_{\text{excit.}} \geqq \lambda_{\text{min.}} = \frac{1.234 \times 10^4}{\text{electron volts}} \text{ Å}$$

For a 10 kilovolt electron, $\lambda_{\text{min.}} = 1.2$ Å. If the emitted wavelength of the excited phosphor be 5230 Å (maximum of α-zinc silicate:Mn) then the quantum deficit allows only $1.2/5230 = 0.02$ per cent efficiency, if each electron produces but one quantum of light.

Efficiencies greater than 5 per cent are actually obtained and the light outputs of phosphors increase at a power of the electron voltage approximately between one and two, as shown in Figure 22. From a set of experimental measurements, calculations for the particular case of a zinc cadmium sulphide phosphor excited by an electron beam carrying 5 microamperes at 10,000 volts show that the phosphor was being struck by $3 \cdot 10^{13}$ electrons/second and emitting 10^{16} light quanta/second.

Each 10,000-volt primary electron was producing 330 light quanta, besides ejecting at least one secondary electron and having 90 per cent or more of its energy converted into heat. An average of 30 electron volts per quantum was expended. The particular phosphor had an activator concentration of one part of silver per million parts of sulphide, hence it was calculated that there was one silver activator center for each 2,400,000 atoms of the bulk crystal, or one activator center in each crystal segment measuring 460 Å on its cube edges and having 210 bulk crystal atoms along an edge. The probability of a primary electron scoring a direct hit on an activator center as compared with a bulk crystal unit is very small, being only: $1/2.4{\cdot}10^6 = 4{\cdot}10^{-7}$. Thus, out of the original $3{\cdot}10^{13}$ electrons/second there would be an "effective" $3{\cdot}10^{13}$

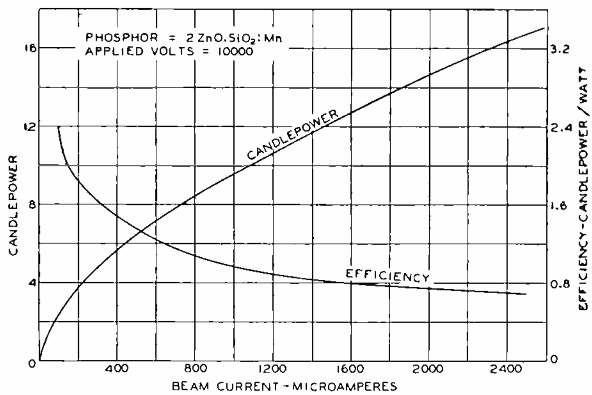


Fig. 23—Typical candlepower and efficiency curves of a phosphor as a function of the Kinescope's electron beam current.

$\times 4{\cdot}10^{-7} = 1.2{\cdot}10^7$ electrons acting solely on the silver activator centers. Assuming the entire energy of each "effective" primary electron to be converted into light, there would be $10^4 \times 1.2{\cdot}10^7/10^{16} = 1.2{\cdot}10^{-5}$ electron volt/quantum.

Since, even with 100 per cent efficiency, at least 2.2 electron volts are necessary to produce a quantum of light at 5560 Å, a correction factor of over 10^5 must be applied to the calculated energy conversion.

One must conclude, from the foregoing data, that a primary electron either acts over a distance of 100-200 Å in the phosphor lattice or else, as is more probable, the energy of the beam is absorbed by the bulk lattice and trans-shipped to the activator centers. The trans-shipped energy is probably in small packets corresponding in energy to radiation in the region of approximately 2000-4000 Å, thus reducing the previously mentioned quantum deficit from 99.98 per cent to 20-60 per cent.

The variation of light output with varied current density is shown in Figure 23. A mathematical formulation[35] of light output in terms of independent variables in Kinescope operation is

$$L = K_1 f_1 (I,a) (V_a - V_o) - K_2 f_2 (V)$$

L = luminous intensity
K_1, K_2 = constants characteristic of the phosphor
I = beam current
a = beam radius
V_a = applied voltage
V_o = extrapolated "dead" voltage (see Figure 22)
$K_2 f_2 (V)$ = secondary emission function ($\propto \Theta$)

Reverting again to the human eye's role in utilizing luminescence, it is found that the eye sensitivity (S) to a change ($\triangle J$) of brightness at a certain brightness (J) may be roughly expressed by

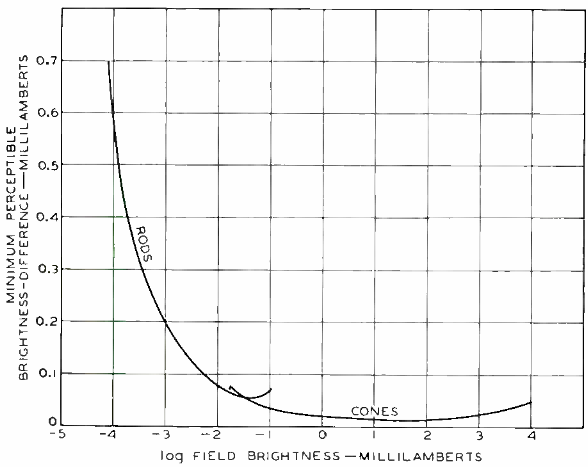


Fig. 24—Minimum perceptible brightness difference as a function of field brightness.

$$dS = K\, dJ/J, \text{ or } S = k \cdot \log J$$

This form of expression is generally true of all the senses, in that the detectable difference of stimulus must be varied as the total stimulus varies (Weber's Law). In general, therefore,

$$\triangle J/J = 1/k = K \text{ (constant)}$$

The ratio $\triangle J/J$ is practically constant over the comparatively narrow range of intensities from 18.6 to 1860 foot candles which is the normal range of daylight illumination.[36]

Figure 24 shows the relationship between field brightness and minimum perceptible brightness-difference for scotopic vision (dark-

[35] See page 29 of reference 14.
[36] See pages 38-42 of reference 25(b).

adapted or rod-vision, *R*) and photopic vision (daylight or cone-vision, *C*) as determined by Hecht.

The sense of brightness interval is demonstrated in Table 5, taken from Luckiesh and Moss.[37] This table shows averaged estimates of ten equal brightness intervals ranging from black to white. The white illumination was 22.8 foot candles and the reflectance of the surrounding field was 19.1 per cent.

Table 5

Estimated Brightness Value	10	20	30	40	50	60	70	80	90	100%
True Reflectance	1.12	2.90	5.95	11.05	18.0	27.3	38.9	53.6	72.8	100%

The scale given in Table 5 provides comparison factors to be applied to the thousands of subjectively determined relative brilliancies of luminescent materials reported in the literature, but the user should remember to consider color differences as well as the brightness differences.

8) *Phosphorescence.*

Until the 19th century and the advent of invisible forms of exciting energy, such as ultraviolet radiation, cathode-ray energy and radioactive emanation, phosphorescence was the principal demonstrable feature of luminescent materials. Quantitative measurements of phosphorescence have been made on thousands of materials since the first phosphoroscope was constructed by Becquerel in his 30 years of research on luminescence prior to 1867.[38] Lenard and his co-workers, starting in the 1880's, constructed many improved phosphoroscopes and experimented with all the important phosphor types under widely varied conditions including the following: temperature, means and degree of excitation, thickness of phosphor layer, size of phosphor crystals, phosphor composition and preparation, etc.[39]

Modern investigators[40] have increased the exactness of phosphorescence measurements, but have not discovered any phosphor exhibiting

---

[37] See page 75 of reference 25(a).

[38] E. Becquerel, "La Lumiere, ses causes et ses effets", I, 247, Didot Freres, Fils et Cie., Paris, 1867.

[39] See pages 103-194 of reference 2.

[40] (a) R. B. Nelson, R. P. Johnson and W. B. Nottingham, "Luminescence during intermittent electron bombardment", J. Appl. Phys. *10*, 335-342, 1939.

(b) G. R. Fonda, "Phosphorescence of zinc silicate phosphors", J. Appl. Phys. *10*, 408-420, 1939.

(c) R. P. Johnson and W. L. Davis, "Luminescence during intermittent optical excitation", J.O.S.A. *29*, 283-290, 1939.

(d) W. deGroot, "Luminescence decay and related phenomena", Physica, VI, 275-289, 1939.

(e) A. Schleede and B. Bartels, "Untersuchungen ueber das An- und Abklingen des Leuchtvorganges bei Phosphoren", Zeits. f. techn. Physik, *11*, 364-396, 1938.

a decay curve contrary to the normal initially rapid decrease in light output, followed by a "tapering-off" of decay rate. According to Johnson and Davis, "It appears that the older phosphors differed chiefly in efficiency, not in any essentials of behavior, from the materials recently developed for television and fluorescent lighting." Figure 25

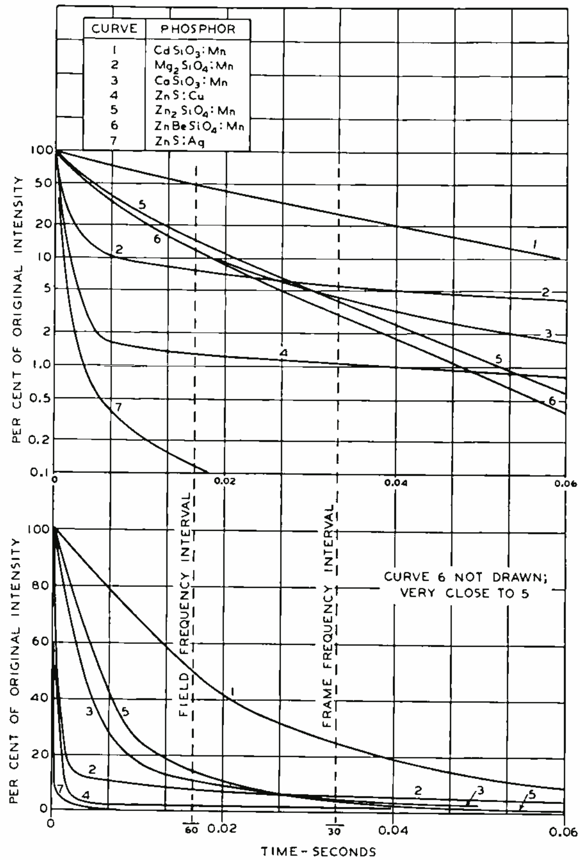


Fig. 25—Phosphor decay curves (persistences). Peak intensities arbitrarily set equal to 100.

shows typical phosphorescence curves for silicate phosphors (exponential-monomolecular) and for sulphide phosphors (hyperbolic or bimolecular).[41] The persistence curves are shown plotted in normal, linear fashion as well as on a semi-logarithmic scale. The latter scale emphasizes the slower rate at the end of a phosphor's decay when the light output is very low. Pure tungstate phosphors have very short

[41] The curve for Cd SiO_3:Mn is from reference 40(c) (loc. cit.) while the remainder are measurements made by T. B. Perkins, Research & Engineering Dept., RCA Mfg. Co., Inc., Harrison, N. J.

decays, lasting about 10^{-5} second. The decay curve of β-zinc silicate:Mn is practically identical with the α-form.

Equations representing the elementary types of phosphor decay curves are as follows:

1) Exponential. Characteristic of a monomolecular process
$L = L_0 e^{-kt}$ typical of silicates and possibly tungstates.

2) Hyperbolic, bi- or poly-molecular type
$L = a/(b+t)^{\alpha}$ typical of sulphides

where L = light output at time t
L_0 = light output at time $t = 0$
$a = L_0 b^{\alpha}$
and k, b, and α = constants characteristic of the phosphor. α has values between 0.8 and 3.

No simple equation will fit any one decay curve over its entire length. The rate of initial decay of sulphide phosphors increases rapidly with the degree of excitation, while phosphorescences of silicate phosphors are less affected by degree of excitation. The long-persistence "tail" of silicate phosphors is more concave upward than is the first nearly exponential part of the decay curve and is strongly temperature-dependent. The "tail" disappears at high temperatures ($> 100°C$) and at very low temperatures $< -100°C$), while phosphor decay curves with or without the "tail" are invariably concave upward.[42] Obviously, no combination of phosphors can yield a decay characteristic deviating from that obtained by superposition of their individual persistences and, hence, the combination's decay will follow the same trend of "upward concavity".

Phosphorescence, as previously mentioned, is light emitted as a result of freeing electrons trapped in or near activator centers and in crystal faults throughout the phosphor. It is obvious that the rate of emission of light during phosphorescence decreases with decrease of the number of electrons remaining trapped. Thus, the intensity of phosphorescence naturally decreases rapidly with time, especially during the initial decay when the concentration of trapped electrons is high. Furthermore, large concentrations of trapped electrons strain the crystal lattice, accelerating the system's rate of return to the unexcited state. The initial persistence decay rate increases with increasing degree of excitation.[43] In general, larger crystals phosphoresce more slowly than smaller crystals, but the decay curves of the various sizes follow the same general equations, differing only in their constants. The phosphorescence of larger crystals is slower than that

---

[42] See reference 40.
[43] N. Riehl. "Aufbau und Wirkungsweise leuchtfähiger Zinksulfide und anderer Luminophore". Ann. d. Physik, *29*, 7, 636-664, 1937.

of smaller crystals because free electrons have opportunity to wander farther afield in the more extended lattices. Their return to light-emitting centers is akin to the game of "musical chairs", only on a more statistically-scrambled three-dimensional scale.

During almost a century of extensive study, the non-occurrence of a single exception to the preceding statement that phosphor decay curve characteristics are concave upward is noteworthy as a statistical weight of probability precluding the attainment of speculative phosphorescent decays having concave downward characteristics, such as

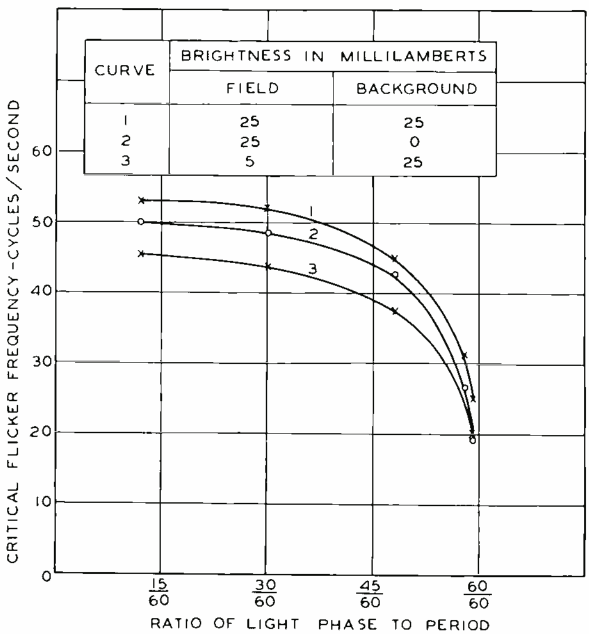


Fig. 26—Critical flicker frequency as a function of the illuminated portion of the field frequency interval.

have been suggested to permit the use of relatively few frames and fields per second for television images.

Figure 26 shows results published by P. W. Cobb[44] demonstrating the reduction of critical flicker frequency by increasing the amount of time the light was on in relation to the total time between light-dark intervals. It is seen that filling 59/60 (98.3 per cent) of the field time with light reduced the critical frequency to only 20-26 cycles/second, and that filling 45/60 (75 per cent) reduces the critical frequency to 39-47 fields per second, the exact values being dependent on illumination

[44] P. W. Cobb, "The Dependence of Flicker on the Dark-Light Ratio of the Stimulus Cycle", J.O.S.A., *24*, 109, 1934.

levels. It follows, from Cobb's data, that the use of phosphors such as long-persistence orange-emitting calcium- or cadmium silicate would decrease the critical frequency to only about 40 fields per second at low brightness levels. At the brightness levels ordinarily demanded for television pictures, the critical frequency for these materials would be about 44 to 48 fields per second. Because it is desirable to use the most efficient light-producing phosphors, shorter persistence characteristics with still higher critical frequencies must be accepted. Thus, a mixture of blue-emitting zinc sulphide and yellow-emitting zinc beryllium silicate—which provides a brilliant white screen—fills less than 45 per cent of the field interval and has a critical frequency of 44 to 52 fields per second for a brilliancy range of 5 to 25 millilamberts.

Minor reduction in field frequency by use of long-persistence phosphors is dearly bought, in that

1) The efficiencies of the long time-lag screens, having at least approximately an exponential decay, are less than 50 per cent of the present white screens. A phosphor's excitation rate is directly proportional to its decay rate. The imposition of a longer persistence further possibly reduces the inherently low efficiencies of such phosphors by decreasing the amount of excitation obtained from the scanning electron beam which bombards each phosphor crystal for but 1.5×10^{-7} seconds in the case of a 12″ Kinescope operating with 507 lines and 30 frames, interlaced.

2) The brilliancy of a long-persistence screen is naturally lower than that of a short-persistence screen because the storage process of long-lag phosphors places a two-fold limitation upon attainment of large concentrations of free electrons in the phosphor lattice:

a. As the crystal becomes excited, there are fewer sources of further electrons which may be ejected to wander into the lattice.

b. As the trapping positions in the crystal become filled, the later ejected electrons are summarily prevented from finding suitable positions and are thus rendered more probable prey for conversion into undesirable heat energy rather than into useful light.

3) Suitable white screens are not obtainable with the orange-emitting long-persistence phosphors by virtue of (a) the inefficiency of orange as a complementary color[45] and (b) the lack of a blue-green complementary phosphor which matches the orange component's decay curve. A mixture of cadmium silicate with copper-activated zinc sulphide produces an approximate lavender-white which, in addition to its undesirable color, is unsatisfactory because rapidly moving

[45] See reference 9.

objects are shown blue on their leading edges and orange on their trailing edges. This phenomenon is the result of incompatible phosphor excitation and decay characteristics. See decay curves shown in Figure 25.

4) Use of lower field frequencies would necessitate restricting phosphors for Kinescope use to but one or two inefficient materials out of the thousands of efficient materials which are known and in development. Such restriction is definitely poor engineering practice, since suitable persistence has been shown to be but one of the many important qualities which must be possessed by a phosphor in order that it be serviceable in Kinescopes.

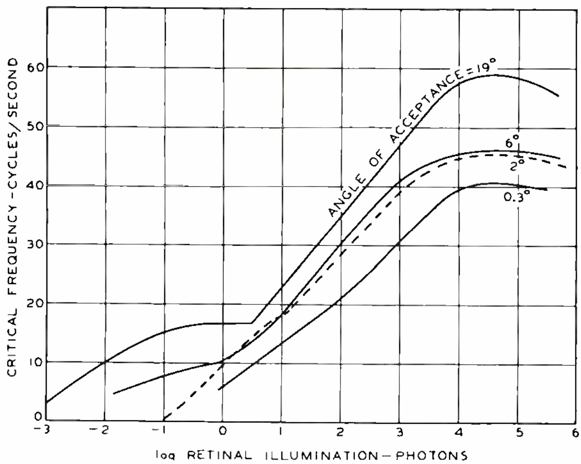


Fig. 27—Critical flicker frequency as a function of retinal illumination. (One photon = unit of visual stimulation = 0.2914 foot-lambert.)

It has been proposed to "straighten" phosphor decay curves by applying infra red or heat to accelerate their phosphorescent outputs inversely as their decay characteristics. Such proposals usually fail to take into account the refractory natures of phosphors and their non-conduction of heat. If temperatures high enough to dissipate heat effectively by radiation are used, the luminescent efficiencies of phosphors will be decreased to negligible values.

Demonstrations of reduced flicker are best carried out at low brightness levels, since the eye's sensitivity to flicker decreases rapidly with decreasing level of light intensity as shown in Figure 27.[46] However, even at low brightnesses, the tendency to approach a dim image more

[46] S. Hecht and E. L. Smith, "Intermittent Stimulation by Light", VI, "Area and the Relation Between Critical Frequency and Intensity", J. Gen. Physiol. *19*, 979, 1936.

closely introduces aberrate vision and greatly increases flicker. Aberrate vision intensifies the retina's peripheral stimulation (rods) with respect to excitation in the fovea centralis (cones). This is shown in Figure 27, in that increasing the angle of acceptance from 0.3° to 19° increases the critical frequency by approximately 30 per cent. As television images increase in size and the dimensions of the home remain fixed, it will become increasingly important to take every precaution to preclude flicker, especially with respect to children who naturally tend to place themselves too close to the television screen.

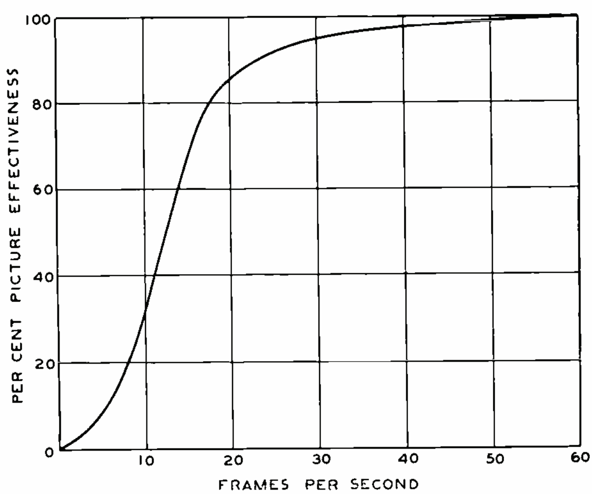


Fig. 28—Relative degree of a television picture's satisfactoriness as a function of frame frequency.

The flicker sensitivity of the human eye as a function of color is directly proportional to the eye's photopic sensitivity to that color. Thus, the critical flicker frequency for colored lights of equal intensity decreases in the order: yellow, 5800 Å; orange, 6100 Å; green, 5100 Å; red, 6500 Å; blue, 4800 Å.[47]

The general average illumination level has been increasing rapidly in the past few decades. In 1896, 0.4 footcandle was deemed sufficient; in 1914, the average was 1-3 footcandles, while in 1931 the standard was 11 footcandles.[48] The brilliancies of television images will undoubtedly tend to increase with the trend of increasing illumination levels.

There is a definite increase in visual acuity up to 15-20 footcandles and increased illumination directly improves the ability of the eye to

[47] See page 374 of reference 25(b).

[48] LeGrand Hardy, "The Measurement of Light", Trans. Illum. Soc. XXXV, 7, 605-624, 1940.

see quickly. There is an increase in "speed of vision" between 0.4 footcandle and 12.0 footcandles amounting to approximately 600 per cent for visual detail of size 1.15′ and 331 per cent for a detail of double this size (Ferrie & Rand as quoted in Hardy).[49] Accordingly, utilization of finer detail in television images (such as by increasing the number of lines) would require increasing picture brilliancy and correspondingly increasing the eye's susceptibility to flicker.

It is well understood that a frame repetition rate of less than 30 per second is inadequate to avoid loss of continuity in rapidly moving objects since even 30 frames per second is insufficient for some television scenes of rapid motion. A thorough study by R. Thun has resulted in the graph of Figure 28 showing degree of satisfactoriness of average pictures (either motion pictures or television) as a function of frame frequency while keeping all other variables at optimum values.[50] An extensive investigation by E. W. Engstrom[51] gives the essentials involved in choosing frame and field frequencies. The study showed it to be illogical to expect that the flicker problem could be solved by using phosphor screens of longer persistence.

Another important fact, which weighs against the use of longer-persistence phosphors, is the blurring or "trailing" effect which is a comet-like tail appended to moving objects reproduced on a long-persistence screen. The flattening shape of the persistence curve for very long phosphorescences intensifies the ratio of light emitted after the picture time-interval as contrasted with the maximum light intensity. Furthermore, the "carry-over" of long-persistence phosphors leaves a higher general level of illumination on the Kinescope screen, thereby seriously reducing contrast.

Summarizing: it appears most sensible to operate electronic television at 30 frames and 60 fields (or more) per second in order to assure: (1) being able to provide Kinescope screens emitting a white light comparable with the ingrained standard of daylight; (2) allowing improvement in Kinescope performance, as new phosphors are developed, without automatically discarding the more efficient materials because their decay characteristics are not slow enough; (3) affording smoothness and continuity in the ever-increasing percentage of scenes which contain rapid motions; (4) avoiding loss of contrast and blurring due to the unavoidable "carry-over" of long persistence screens; and

---

[49] See reference 48.

[50] R. Thun, Die Kinotechnik, April 5, 1935, p. 117; ibid Nov. 5, 1935, pp. 358-362; ibid Nov. 20, 1935, pp. 374-376. (Abstracted by Dr. D. W. Epstein and furnished through the courtesy of Mr. E. W. Engstrom, Director of Research, RCA Mfg. Co. Inc., Camden and Harrison, N. J.)

[51] E. W. Engstrom, "A Study of Television Image Characteristics", Proc. I.R.E., *21*, 1631-1651, 1933; ibid, *23*, 295-311, 1935.

(5) providing high enough levels of flicker-free picture brilliancy to be adequate, not only for the "normal eye", but also for the refractively defective eye. The latter receives greater benefit from increase in quality and quantity of illumination than does the "normal eye". Approximately two-fifths of our population, comprising millions of people, have defective visual functions[52] which must be considered in establishing a public service such as television.

Appreciation is expressed for the stimulating interest accorded these investigations by Mr. E. W. Engstrom, Dr. G. R. Shaw and Dr. E. A. Lederer as well as other members of the Research Department.

Reading List

(Supplemental to the cited references)

(1) *General References:*
H. Rupp, "Die Leuchtmassen und ihre Verwendung," Borntraeger, Berlin, 1937.
"Symposium on Luminescence," Transactions of the Faraday Society, *35*, 1, 1939.
V. K. Zworykin and G. A. Morton, "Television. Electronics of Image Transmission," J. Wiley, 1940.
F. Fritz, "Leuchtfarben," Gustav Bodenbender, Berlin, 1940. (Very extensive bibliography of over 830 papers and patents.)

(2) *Recent Specific References:*
N. C. Beese, "The Response of Several Fluorescent Materials to Short Wave-Length Ultraviolet Radiations," J.O.S.A. *29*, 278-282, 1939.
H. Bey, "Aufladepotentiale elektronen-bestrahlter Leuchtmassen," Physik. Zeits., *39*, 605-611, 1938.
G. G. Blake, "Infra-Red Radiations with Special Reference to their Quenching Effects upon Zinc Sulphide Phosphors," Roy. Soc. N. S. Wales, Journ. & Proc. *73*, 112-124, 1939.
J. Ewles, "Resolution and Interpretation of the Luminescent Spectra of Some Solids at Low Temperatures," Proc. Roy. Soc., Series A, No. 928, Vol. 167, 34-52, 1938.
G. R. Fonda, "The Fundamental Principles of Fluorescence," A.I.E.E. Technical Paper 38-57, May 1938.
G. R. Fonda, "Characteristics of Silicate Phosphors," J. Physical Chem., *43*, 561-577, 1939.
G. R. Fonda, "The Preparation of Fluorescent Calcite," J. Physical Chem., *44*, 435-439, 1940.
J. Frenkel, "On the Absorption of Light and the Trapping of Electrons and Positive Holes (also Excitons) in Crystalline Dielectrics," Sowj. Physik, *9*, 2/3, 158-187, 1936.
J. H. Gisolf, "The Absorption Spectrum of Luminescent Zinc-Sulfide and Zinc-Cadmium Sulfide in connection with some Optical, Electrical and Chemical Properties," Physica, VI, 84-96, 1939.
G. Graue and N. Riehl, "Untersuchungen über die Porenweite und das innere Volumen amorpher und kristalliner Stoffe," Zeits. anorg. allgem. Chem., *233*, 365-375, 1938.
W. Grotheer, "Ermüdungserscheinungen Elektronenbestrahlter Leuchtmassen," Zeits. f. Physik, 112, 9/10, 541-559, 1939.
C. Hagen, "Aufladepotentiale, Sekundär-emission und Ermüdungserscheinungen elektronenbestrahlter Metalle und Leuchtsubstanzen," Physik. Zeits., *20*, 621-640, 1939.
S. T. Henderson, "Band Spectra of Cathodo-Luminescence," Proc. Roy. Soc., Series A, No. 954, Vol. 173, 323-338, 1939.

[52] See reference 10.

R. P. Johnson, "Luminescence of Sulphide and Silicate Phosphors," J.O.S.A., *29*, 387, 391, 1939.
K. Kamm, "Über die Zinksulfid-Cadmiumsulfid-Phosphore," Ann. d. Physik, *30*, 333-353, 1937.
F. A. Kröger, "Some Optical Properties of Zisc Silicate Phosphors," Physica, VI, 764-778, 1939.
F. A. Kröger, "Fundamental Absorption of ZnS-MnS and ZnS-CdS-MnS Mixed Crystals," Physica, VI, 779-784, 1939.
J. W. Marden and G. Meister, "Effects of Impurities on Fluorescent Compounds," Trans. Illum. Eng. Soc. XXXIV, 503-513, 1939.
F. Möglich, "The Mechanism of Luminescence in Phosphorescent Crystals," Angew. Chemie, *53*, 54-56, 1940.
J. T. Randall, "The Fluorescence of Compounds containing Manganese," Proc. Roy. Soc. Series A, No. 940, Vol. 170, pp. 272-293, 1939.
N. Riehl, "Über einen neuen Effekt an lumineszierendem Zinksulfid," Zeits. f. techn. Physik, 5, 152-155, 1939.
N. Riehl and M. Schön, "Der Leuchtmechanismus von Kristallphosphoren," Zeits. f. Physik, *114*, 682-705, 1939.
E. Streck, "Über die Zerstörung des Zinksulfids durch Licht und α-Strahlen," Ann. d. Physik, *34*, 96-112, 1939.

SS "AMERICA" RADIO INSTALLATION

By

I. F. Byrnes

Chief Engineer, Radiomarine Corporation of America, New York

THE largest and fastest passenger liner built in the United States up to the present time is the new S.S. *America*. This vessel (Figure 1), which is owned by the United States Lines Company, was placed in commercial service in August of this year. The *America* is a 35,000-ton ship and has a length of 723 feet, a beam of 93 feet and is designed to carry approximately 1200 passengers. The radio installation is specially designed to provide the utmost in safety and communication efficiency and in all respects is in keeping with the modern design features of the vessel itself. The radio room is located near the forward section of the sports deck, and is in close proximity to the bridge and chart room.

Fig. 1

Antenna System

A fairly elaborate antenna system is provided on the *America* in order to accommodate the various radio transmitters and receivers which must operate over wide frequency bands for both telegraph and telephone service. The main flat-top transmitting antenna is divided into two sections, and consists of a single wire supported between the foremast and the mainmast. The aft flat-top section is approximately

240 feet long, while the forward section is 60 feet long. The flat-top is divided with insulators directly above the radio room, and two 70-foot downleads are brought into the trunk, as shown in Figure 2. By means of switches in the radio room, the larger flat-top section may be used

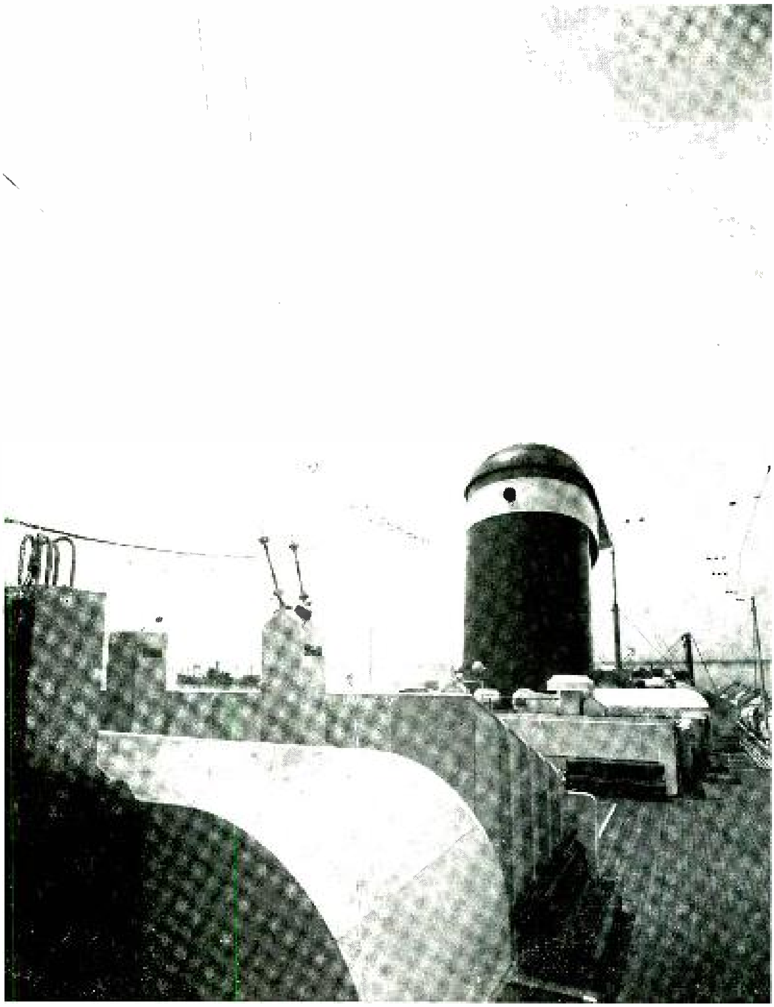


Fig. 2

for the intermediate-frequency transmitter and the forward section for high-frequency telephone or telegraph operation. Both flat-top sections may be connected in parallel for low-frequency or long-wave transmission. The main transmitting antenna is approximately 107 feet above the water line. The effective capacity of the aft section is about 1000

$\mu\mu$f, while the forward section has a capacity of approximately 500 $\mu\mu$f.

There are five doublet antennas, for high-frequency reception, located between the stacks and supported in the thwartship plane of the vessel. These doublets are cut for various frequencies between 4 and 22 megacycles. In addition, there is also provided a horizontal "V" antenna which may be used with the emergency transmitter or the high-frequency telegraph transmitter.

Fig. 3

Radio Room Transmitters

The four main transmitters in the radio room are shown in Figure 3. These transmitters, while differing considerably in electrical design, are built in frames of equal size in order to permit a symmetrical layout and uniform appearance. The transmitter on the right is for low-frequency radiotelegraph service, and is designed to deliver one kilowatt of power to the antenna on any one of ten crystal-controlled frequencies between 110 and 160 kilocycles. The next transmitter is for intermediate-frequency service, and provides ten crystal-controlled frequencies and one kilowatt of antenna power in the band from 350 to 500 kilocycles. The third unit, for high-frequency radiotelegraph transmission, covers five frequency bands between 4140 and

22,200 kilocycles, with provision for a maximum of thirty crystal-controlled output frequencies and one kilowatt of antenna power. The fourth transmitter is the radio-frequency unit for the 600-watt radiotelephone transmitter. This is a five-band transmitter, crystal-controlled, and is normally set up for the ship-to-shore telephone bands between 4000 and 18,000 kilocycles. Additional details on the construction of these transmitters is considered further in this paper.

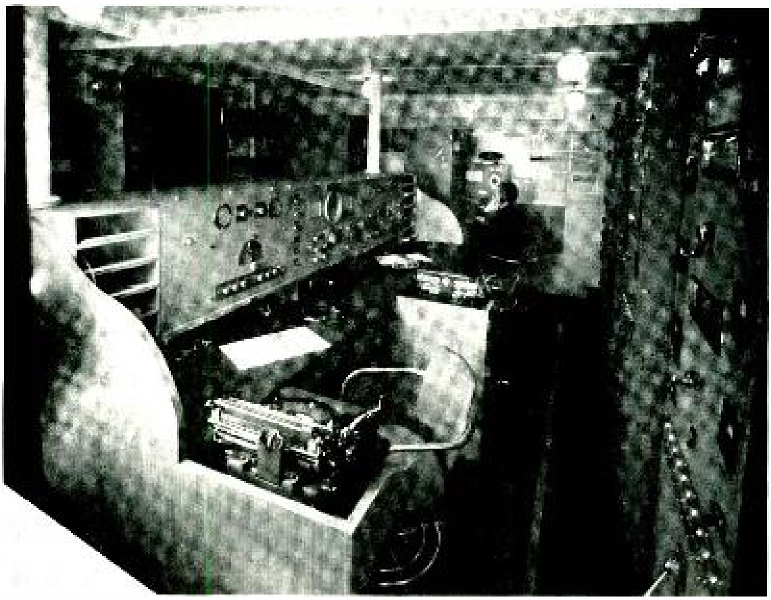


Fig. 4

Radio Room Receivers

For radiotelegraph reception, a special receiver console was designed as an integral part of the radio operating position in order to provide the most efficient and convenient facilities for the radio operators. This arrangement is shown in Figure 4. The radio receiver to the left is a superheterodyne unit covering a band from 75 to 1500 kilocycles. The receiver unit in the center of the console, directly beneath the radio room clock, is employed for general stand-by or auxiliary service and is a tuned radio-frequency unit covering the band from 15 to 600 kilocycles. The third receiver at the right end of the console is used for high-frequency reception, and is a superheterodyne covering all frequencies between 540 and 30,000 kilocycles. All receivers are arranged for both loud-speaker and headphone reception, and with switching facilities to enable time signals to be passed on to a loudspeaker located in the chart room. The receiver console is also provided with the necessary control switches for the radio transmitters. In the far end of the radio room, there is located the radiotelephone operator's position, which includes the operator's control panel, radiotelephone receiver and the radiotelephone rectifier-modulator unit. The control panel provides circuits for connection to the radiotelephone booth located just outside the radio room, and in addition is arranged for connection to stateroom telephones.

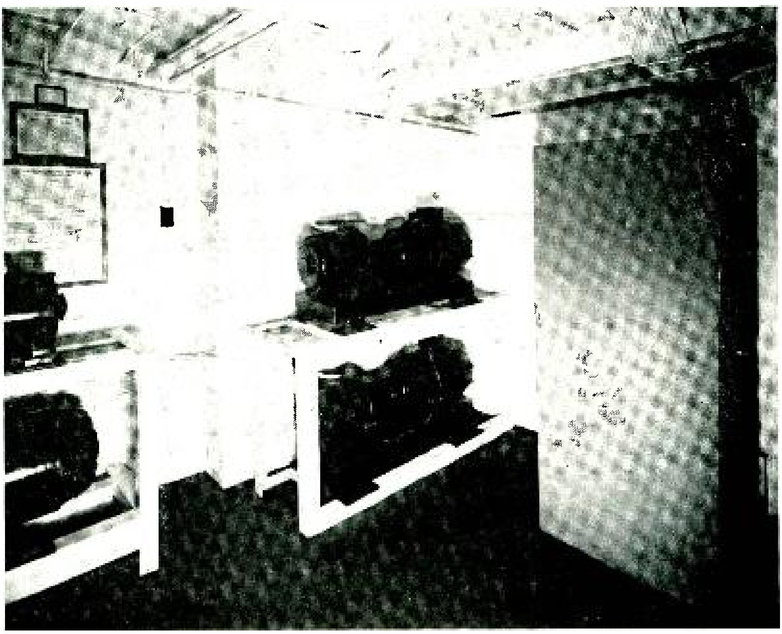


Fig. 5

Power Supply

While a small amount of 60-cycle power supply is available on the *America*, the principal power source is 230 volts d.c. For this reason it was necessary to use rotating equipment in order to provide the necessary voltages for the four main radio transmitters. In shipboard radiotelegraph service, an important consideration is to provide facilities to enable the radio operator to go "on the air" in the shortest possible time. This factor makes it undesirable to employ mercury vapor rectifier tubes or other types of tubes which may require thirty seconds or more time delay before reaching their operating temperature. For these reasons, it was decided to employ high-voltage d-c motor generator sets for plate supply, and to use filament type tubes in the radio transmitters. With such an arrangement, any of the radiotelegraph transmitters may be placed "on the air" within

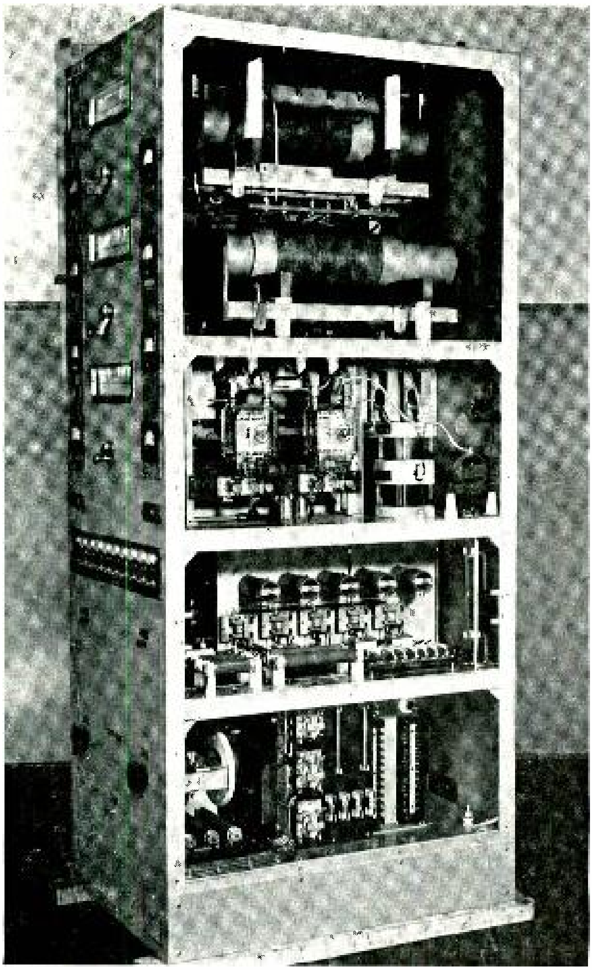


Fig. 6

approximately three seconds. The high-voltage motor generators and their control panel are located in the fire control room just aft of the chart room. These machines and the "starter-filter panel" are shown in Figure 5. There are two 2500-volt motor generator units, one of which is used to supply power to either the intermediate- or low-frequency transmitter, while the second high-voltage machine supplies the one-kilowatt high-frequency telegraph transmitter. Provision is made in the "starter-filter panel" to switch the high-frequency motor generator so that it may be used to furnish power to the intermediate or low-frequency transmitters in the event of machine failure. The

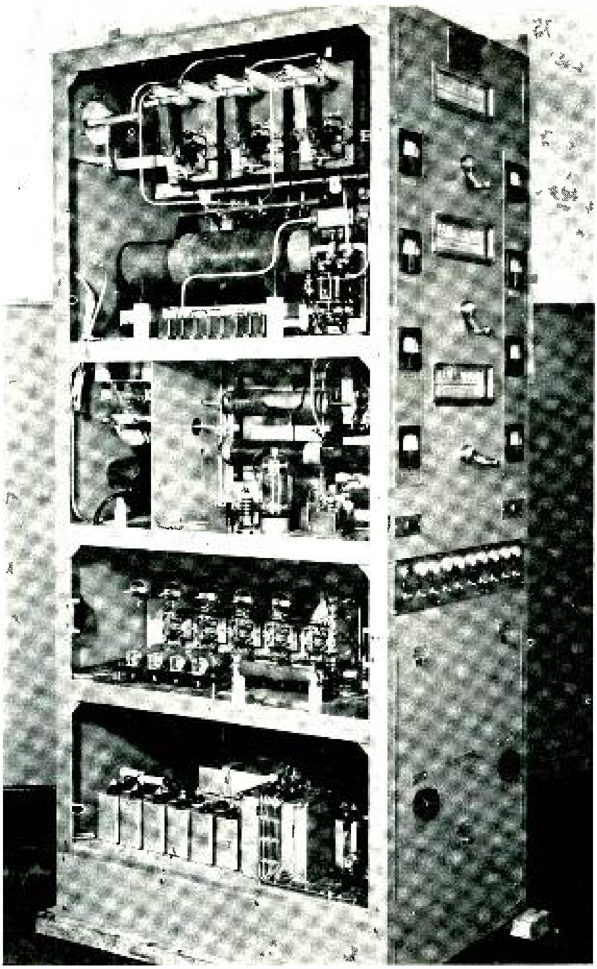


Fig. 7

third motor generator in the fire-control room furnishes 625-cycle power for modulating the intermediate-frequency radiotelegraph transmitter. The fourth machine is a standard motor alternator unit, and supplies 60-cycle power for the 600-watt radiotelephone equipment.

TRANSMITTER DESIGN

A view of the interior construction of the one kilowatt intermediate-frequency transmitter is shown in Figures 6 and 7. The transmitter employs only four tubes. The crystal oscillator stage consists of a 1624 pentode, and is arranged with ten pre-tuned oscillator circuits, the tuning adjustment being obtained by means of movable iron cores. Small relays, controlled by toggle switches on the front panel, enable the appropriate crystal and oscillator circuit to be quickly selected. The buffer stage of the transmitter consists of an 813 tube, and the power amplifier stage uses two 833 tubes which operate in parallel. For the so-called "working" waves, that is, all frequencies in the intermediate band except 500 kilocycles, there are provided ganged buffer and power amplifier tuned tank circuits which are controlled from the front panel. Movable iron cores are used here for inductance variation. In like manner, the antenna circuit may be resonated to any one of the working frequencies by means of a panel control which moves the iron cores in or out of the antenna loading inductance.

The international distress and calling frequency of 500 kilocycles is pre-tuned throughout in the intermediate-frequency transmitter, and relays are provided in the various stages to instantly connect the 500-kc channel to the antenna. By means of this arrangement, the radio operator need only throw a small switch, from his operating position, in order to change from 500 kilocycles to any one working wave that he has previously selected. This facility enables expeditious handling of traffic and eliminates the need on the part of the operator to manipulate controls on the radio transmitter when he wishes to change from the calling to a working frequency or vice versa.

The intermediate-frequency radiotelegraph transmitter must be designed for both A-1 (continuous wave) and A-2 (modulated wave) emission. In the *America's* transmitter, modulation is accomplished by means of a 625-cycle alternator which is coupled to the plate circuit of the radio power amplifier tubes through a suitable step-up transformer. About 750 watts of audio power are required to modulate the carrier at 80 per cent modulation. Suitable precautions have to be taken to insulate the antenna loading inductances as well as the relays in the antenna circuit of the transmitter, since peak voltages to the order of 17,000 volts are developed in the output circuit. The antenna current on A-2 transmission is approximately 22 amperes.

The design of the low-frequency (110-160 kilocycles) transmitter is in general similar to that employed for the intermediate-frequency set. The low-frequency transmitter is arranged for A-1 emission only, as modulated wave transmission is not permitted in this band. A power transfer switch is used on the low-frequency set so that the high-voltage motor generator may be readily connected to this set or to the intermediate-frequency transmitter. Both transmitters are never used simultaneously, and for this reason it is possible to operate with the same motor generator power supply and the same antenna system. Since a large value of CL is required for operation around 110 kilocycles, it is desirable to have an antenna of the highest possible

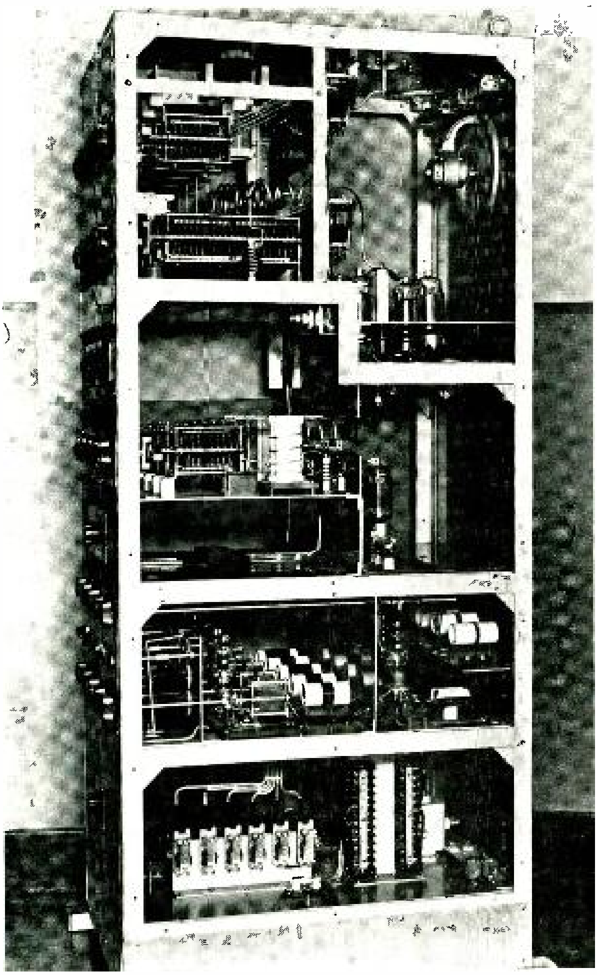


Fig. 8

effective capacitance in order that the necessary antenna loading inductance may be kept down to as low a value as possible. This is accomplished on the *America* by paralleling the forward and aft sections of the main flat-top. Voltages to the order of 25,000 volts exist at the antenna output terminal of the low-frequency transmitter.

For high-frequency telephone or telegraph transmission, the design of the transmitters follows conventional practice with the exception that provision is made for pre-tuned circuits throughout in each of the high-frequency bands. An interior view of the radio-frequency unit of the 600-watt telephone transmitter is shown in Figure 8. Separate oscillator and buffer amplifier tubes are used for each of the five bands between 4000 and 18,000 kilocycles. Type 807 tubes are used in each of these stages. The third stage, which functions as a fundamental amplifier or as a frequency multiplier, is a Type 813 tube, while the final power amplifier stage consists of four Type 813 tubes operating in parallel. Relays, operated either from the front panel or from the operator's position, enable the various pre-tuned stages to be instantly selected. If the operator requires a different frequency in any one band, he selects the appropriate crystal from the transmitter panel. This arrangement therefore provides a total of five "quick" frequencies within the entire range. As each band has facilities for six crystals, it is possible to use the transmitter on a total of thirty output frequencies.

The power amplifier stage of the 600-watt telephone set is plate modulated with power derived from a separate "rectifier-modulator" unit. This unit comprises four Type 805 modulator tubes which function as Push-Pull Class "B" amplifiers. These tubes are preceded by conventional stages of Class "A" audio amplifiers, finally terminating in the output of the microphone. A separate audio amplifier-rectifier unit, operating from the microphone output, is employed to permit voice control of the transmitter carrier wave. This arrangement, in effect, keys the transmitter unit each time speech is impressed on the microphone and at the same time renders the radio receiver inoperative. When the party aboard ship stops talking, the transmitter carrier is cut off and the receiver input and output circuits are energized. Operation with voice-controlled carrier is of considerable value when receiving signals under unfavorable conditions, as it eliminates the necessity of the ship's radio receiver being required to work through the strong outgoing signal from the transmitter. Under more favorable receiving conditions, it is preferable to operate with a constant carrier, and when connections are made to staterooms, because this eliminates "clipping" of initial speech syllables and more nearly simulates land-line telephone operation. The control unit for the telephone transmitter is fitted with a privacy device to invert outgoing speech, as well as to restore the incoming inverted received signal.

The high-frequency radiotelegraph transmitter unit is designed similarly to that employed for telephone, except that A-1 emission only, with an antenna power of one kilowatt, is provided. In addition, the high-frequency telegraph transmitter provides operation in the 21,000-kilocycle ship telegraph band.

EMERGENCY EQUIPMENT

Passenger vessels are required by law to be equipped with a separate emergency transmitter and receiver, energized from a separate power supply so that operation for a period of at least six hours may be obtained independently of the ship's normal power source. On the *America*, a completely separate emergency station is installed in one

Fig. 9

corner of the main radio room. This layout is shown in Figure 9. A small 50-watt transmitter, which derives its power from a motor generator set, and a large 12-volt storage battery, is employed. A standard shipboard 15-600 kilocycle intermediate-frequency receiver and an auxiliary crystal receiver are also installed. An auto alarm, mounted on the bulkhead, for standing a distress watch on 500 kilocycles is provided as a further aid to safety to permit reception of the international alarm signal in the event that the radio operators are engaged in handling traffic on other frequencies.

AUXILIARY RADIOTELEPHONE

A noteworthy feature of the radio installation of the *America* is the coastal harbor type of radiotelephone transmitter and receiver that is installed in the radio room. This is a ten-frequency 75-watt crystal-controlled transmitter and a ten-frequency receiver, which operates in the coastal harbor band between 2000 and 3000 kilocycles. This equip-

Fig. 10

ment is shown in Figure 10. The auxiliary radiotelephone provides a convenient means of communication with tugboats or coastal stations when approaching port.

RADIO DIRECTION FINDER

A modern direction finder is installed on the bridge of the *America*, as shown in Figure 11. This instrument employs a highly sensitive and selective superheterodyne receiver, and is employed in the conventional manner for taking bearings on marine beacon stations or on other vessels. The main compass card of the direction finder is driven by the vessel's gyro repeater system, thereby enabling great circle radio bearings to be obtained with respect to true north.

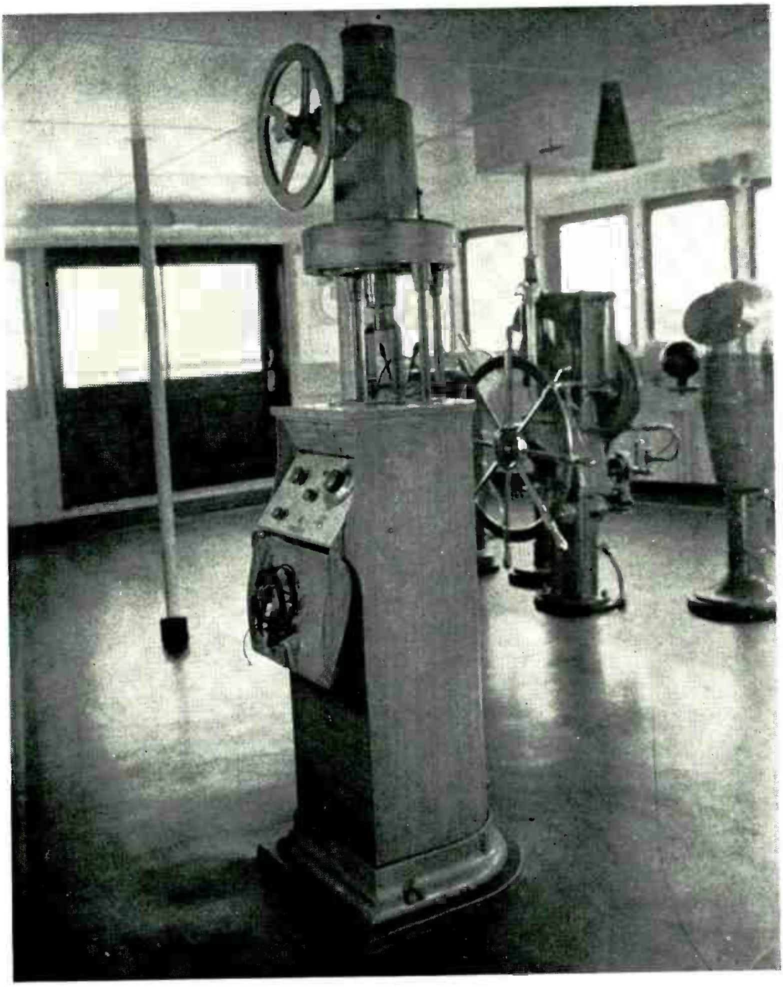


Fig. 11

Lifeboat Radio

The port and starboard motor lifeboats of the *America* are each completely equipped with a radiotelegraph installation to comply with Government Regulations. The radio transmitter in each lifeboat is pre-tuned to 500 kilocycles, while the radio receiver covers the conventional intermediate-frequency band of 350-500 kilocycles. Each motor lifeboat is fitted with a large 12-volt storage battery which operates the radio transmitter and receiver for a period of at least six hours. The antenna available on the lifeboats is necessarily small, and consists of a single wire approximately twenty feet long and twenty feet above the waterline. The masts for the antenna are hinged so that they may be readily stowed when the lifeboats are on the davits. The storage batteries in the lifeboats are maintained on continuous charge by means of suitable charging panels located in the radio room.

Miscellaneous

A total of 122 vacuum tubes are used in all of the radio equipment on the *America*. The maximum current required from the ship's main 230-volt d-c line is approximately 85 amperes. About 95 leaded and armored conductors, with a total weight of 4½ tons, are required for the various incoming and outgoing circuits of the radio room. The total weight of the radio transmitter, receivers, motor generators, etc., is 6½ tons.

For the convenience of passengers, there is available a small reception room adjacent to the main radio room and near one of the elevators on the vessel. The reception room includes a sound-proofed telephone booth for the ship-to-shore telephone service. Under normal conditions, simultaneous operation of two telegraph transmitters, the main telephone set and the auxiliary radiotelephone may be accomplished.

NBC FREQUENCY-MODULATION FIELD TEST

BY

RAYMOND F. GUY AND ROBERT M. MORRIS

National Broadcasting Company, Inc., New York

Summary—*The full theoretical advantages of frequency modulation may be obtained in practice if the transmitting and receiving apparatus are properly designed.*

For primary service, amplitude modulation on the ultra high frequencies offers some advantages over standard broadcasting. Frequency modulation offers advantages over amplitude modulation on the ultra high frequencies. The advantages to the listener of frequency modulation on the ultra high frequencies consist of freedom from the 10-kc beat note and side-band interference which result from the frequency allocation of standard broadcasting, and also the reduction of locally generated noise, atmospherics, and interference from distant stations operating on the same channel. Standard broadcasting has the advantage of providing clear channel night-time service to vast areas which would not be served by frequency modulation on the ultra high frequencies.

TWENTY years ago all frequencies above 1500 kc were generally considered to be of such little value that even the amateurs had objections to being confined to them. There is no need to state here what has since occurred on these "useless" frequencies nor to dwell on the fact that the surface has but been scratched. One service after another has wholly or partially transferred the bulk of its activities to them, and a multitude of new and invaluable services have been made possible by their use. So-called "Standard Broadcasting" had a most humble beginning on 830 kc, which was then in the middle of the marine band of 500 kc to 1000 kc. Broadcasting quickly crowded the original occupants out of most of this band. It is not one of the services which have since moved into the high-frequency spectrum. It remains on the former marine frequencies where it started. But there is a possibility that a shift may be approaching.

The use of the ultra-high frequencies for sound broadcasting offers some technical advantages. These advantages consist of escaping the present 10-kc channel limitation, getting away from static and eliminating all except spasmodic long-distance interference. We have known this for many years, have for a decade experimentally operated low power u-h-f stations and at times have had the experience of receiving good service from our low-powered u-h-f transmitters when static ruined reception from our high-powered standard broadcasting plants. Five years ago the FCC had applications for, or had licensed, over 100 u-h-f transmitting stations and it seemed that a trend was developing toward u-h-f broadcasting, but this trend was not sustained. Interest has been revived in recent months through wide-spread discussion of the advantages of frequency modulation on the ultra high frequencies.

Amplitude-modulated u-h-f stations can provide greater coverage than standard broadcasting shared-channel stations limited by night-time interference from distant stations operating on the same channel. This interference usually causes such a station's useful service area to shrink to a small fraction of its daytime area and many of these stations are, in addition, required to reduce power at night to minimize similar interference to other stations. Few such stations are free from night-time interference within their 2000 microvolt contour and many are limited to 8000 microvolts. The ultra high frequencies offer an escape from such limitations by virtue of practical freedom from static and shared-channel interference. Frequency modulation can provide a much greater degree of improvement than can amplitude modulation on these ultra high frequencies.

Purpose of F-M Field Test

The National Broadcasting Company has been one of the groups which viewed realistically the possibilities of the ultra high frequencies and has pioneered in conducting experimental operations in that field for many years. It has been NBC's belief that at some time in the future these ultra high frequencies would come into much wider use. It is the policy of NBC to investigate every technical development affecting its field and apply it where possible to provide better public service. Toward that end 18 months ago NBC formulated plans for an exacting field test of frequency modulation similar to the field test of television which preceded the dedication of that service on April 30, 1939. The purposes of the field tests were to quantitatively evaluate the advantages of frequency modulation over amplitude modulation on the ultra high frequencies, not by confined laboratory measurements which had already been made by others, nor merely by operating a frequency-modulated station, but *by painstakingly measuring and comparing the two systems under all kinds of actual service conditions in the field* and determining how much of the theoretical advantage of FM could be obtained in practice.

The impression has been gained by some, that only by the use of FM may the public now enjoy high fidelity sound reproduction in the home. With ultra high frequencies, the same fidelity can be provided with either amplitude or frequency modulation. Improved fidelity is made possible by increasing the transmitting channel width beyond the 10-kc channel allocations of the standard broadcasting band and not by using a particular type of modulation. Irrespective of the type of modulation, receivers for "high fidelity" reproduction require equally costly high power, low distortion audio amplifiers and expensive loudspeakers and acoustical systems.

Present-day transmitters of NBC and others in the standard broadcasting band transmit much higher fidelity than is reproduced in moderate-priced receivers. To reproduce sound in the degree of fidelity which is now available, requires more costly receivers. Popular interest has been much more pronounced in low-priced receivers than in high fidelity at more cost. The problems of high fidelity are problems of cost and of widespread public appreciation of improved fidelity and not of a type or method of modulation. High fidelity eventually will receive the recognition it merits. Provision has been made for it in the ultra-high-frequency channel allocations.

In the NBC field test of frequency modulation attention was directed toward the evaluation of the frequency modulation system in the suppression of undesired noise and interference, using the amplitude modulated system as the reference, or standard of comparison.

Scope of the Tests

To carry out this project properly, and for the first time completely, it was decided that:—

1. The same transmitter, transmitting antenna, receiving antennas, receivers and measuring equipment should be utilized for each system of modulation.
2. The transmitter and receiver should be equipped for instantaneous switching to either amplitude modulation or frequency modulation and the transmitter should be of 1000 watts power.
3. The transmitter power should be continuously variable over a range of 10,000 to 1 on frequency modulation, and a means of accurately measuring it should be installed.
4. The most important comparisons should be between amplitude modulation, frequency modulation with a deviation of 15 kc (total swing of 30 kc), and frequency modulation with a deviation of 75 kc (total swing of 150 kc). A minimum of 15 kc deviation was chosen because it represents a deviation ratio (deviation divided by maximum audio frequency transmitted) of 1 and because a 30-kc i-f system would still be required for a smaller deviation, to accommodate the side-bands.
5. Order wires should be used between the transmitting and main

receiving points to expedite the work and insure accurate results.

6. The observations and measurements should be conducted at a number of scattered and representative receiving locations throughout the service area of the station.
7. The observations would include shared channel and adjacent channel operation of F-M stations.

The transmitter was equipped with relays for instantly selecting at will the condition of modulation desired. Since the degree of frequency deviation is directly proportional to the audio input voltage, pads, selected by a relay, served to produce various frequency deviations. Herein, for frequency modulation, the term "per cent modulation" refers to the passband of the system and is the ratio of the total swing being used divided by the total pass-band.

Tone modulation was used for most measurements. For measurements of distortion or signal-to-noise ratios, with modulation present, the tone output of the receivers was cleaned up by passing it through filters and then impressed upon RCA noise and distortion meters.

For brevity the following designations will be used herein:

AM ....Amplitude Modulation.
FM-15 ....Frequency Modulation with a deviation of 15 kc, or total swing of 30 kc.
FM-30 ....Frequency Modulation with a deviation of 30 kc, or total swing of 60 kc.
FM-75 ....Frequency Modulation with a deviation of 75 kc, or total swing of 150 kc.
S/N ....Signal-to-noise ratio.

Preliminary work on this field test was started in April, 1939, and the project was completed on May 4, 1940. The results of this work were submitted to the FCC during the F-M hearings of March, 1940 and presumably were of assistance to that body in formulating rules and standards of good engineering practice covering u-h-f broadcasting.

THE TRANSMITTER

The 1000-watt transmitter was an RCA unit, originally designed for amplitude modulation, which was modified and equipped to meet the requirements for both FM and AM. The receivers, of which there were four, were specially designed and built throughout so as also to meet the requirements.

Special temporary authorization was requested and obtained from the FCC to use either FM or AM on 42.6 Mc, ordinarily an exclusive F-M frequency.

The station was licensed as W2XWG and was constructed on the Empire State Building in New York City. It is shown in Figure 1. This transmitter utilized an FM modulator of the type developed by Murray G. Crosby of the RCA. Some manufacturers of F-M transmitters have adopted this system as standard equipment for commercial F-M transmitters. The transmitter was equipped with a reactance control in the primary circuits of the main rectifier which

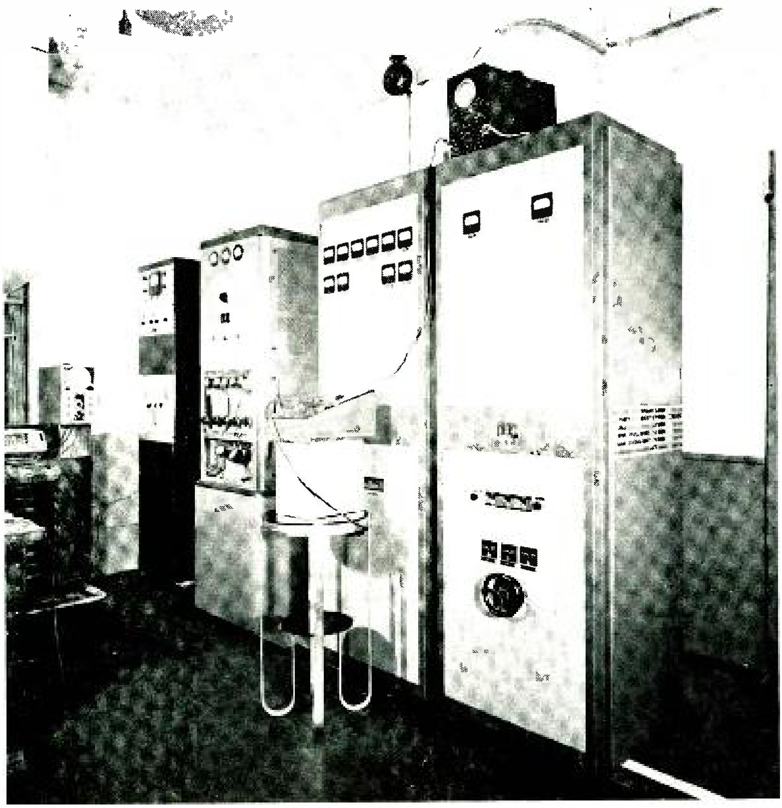


Fig. 1

permitted continuously variable control of the carrier power from 1000 watts down to less than 1/10 of one watt during transmission with frequency modulation. A General Radio multi-range vacuum tube voltmeter was provided to measure the transmission line voltage and power accurately at any point over this wide range. Both of these instruments are illustrated by Figure 2.

The frequency modulator consisted of several tubes, including an oscillator, a reactance tube, a crystal beating oscillator, a discriminator and a filter. The operation is as follows:—

The reactance tube produces a change of reactance in its output circuit when the grid voltage is actuated by d-c bias or a superimposed audio frequency such as tone or program. This reactance controls the frequency of a connected oscillator which drives the transmitter. The other items in the unit are provided to maintain accurate average carrier frequency stability. This is accomplished by a series of simple operations. A fixed crystal oscillator beats against the fre-

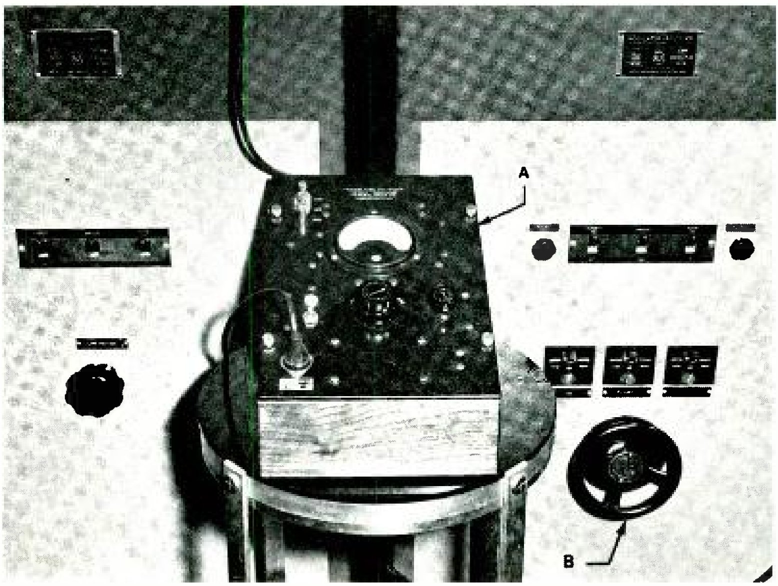


Fig. 2

quency of the modulated oscillator to produce an average difference frequency of 1500 kc. This 1500 kc is brought up in level through an amplifier and impressed upon a discriminator. When the carrier frequency is such that exactly 1500 kc is produced no voltage appears in the output of this discriminator. However, when the reactance tube average frequency changes, 1500 kc is no longer produced and a voltage is developed in the output circuit of the discriminator. A discriminator output filter removes all modulation frequencies to produce a direct current which controls the reactance tube d-c bias and thus the average frequency of the oscillator. The purpose of the filter is to prevent the reactance tube bias from changing with modulation.

The transmitter modulation characteristics are shown on Figure 3.

On this curve there are plotted a-c volts as the abscissa and carrier frequency as the ordinate. The audio input to the transmitter is superimposed on the reactance tube d-c bias.

Figure 4 shows the overall frequency response of W2XWG and the Bellmore receiver before pre-emphasis and de-emphasis were inserted.

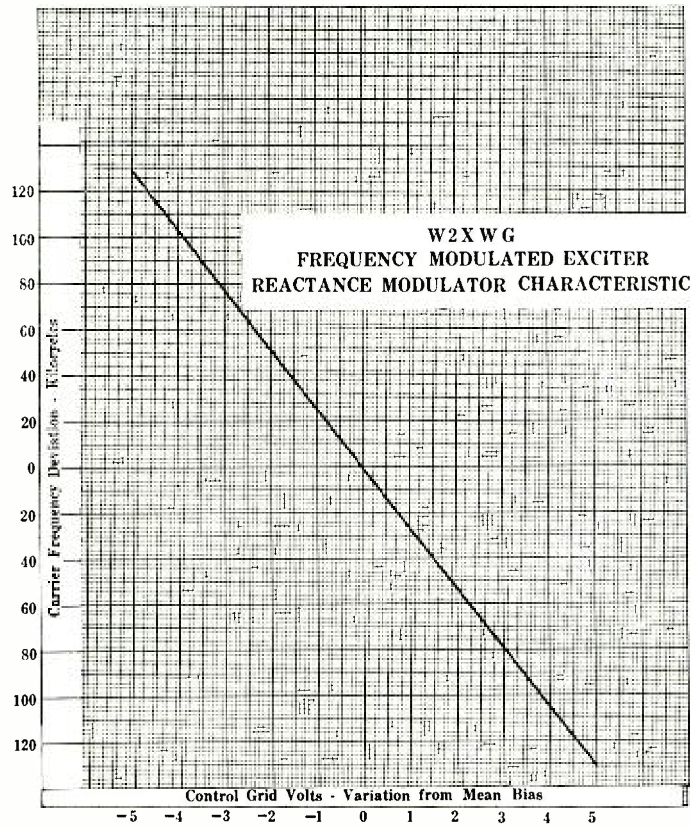


Fig. 3

The antenna normally used for transmitting for the field test was the television video unit on top of the Empire State Building.[1] However, an auxiliary antenna was built consisting of a folded dipole. This was directed eastward towards Bellmore and used when the video antenna was occupied for television schedules.

The photograph on page 5 shows the video antenna, which is 1300

[1] Television Transmitting Antenna for Empire State Building, Nils E. Lindenblad, RCA Review, pp. 387-408, April, 1939.

feet above the sidewalk. Figure 6 shows the folded dipole in operating position.

The circuit arrangement of the special 1000-watt transmitter is illustrated in the form of a block diagram in Figure 7. The frequency modulator comprises the five blocks at the upper right side, containing RCA 6V6, 6K8 and 6H6 tubes.

The degree of frequency deviation was measured by a method recently described in the RCA REVIEW[2]. The method consists of applying a constant frequency tone to the transmitter input terminals and gradually increasing the voltage until the carrier amplitude drops to zero. When this occurs, the frequency deviation is 2.405 times the audio

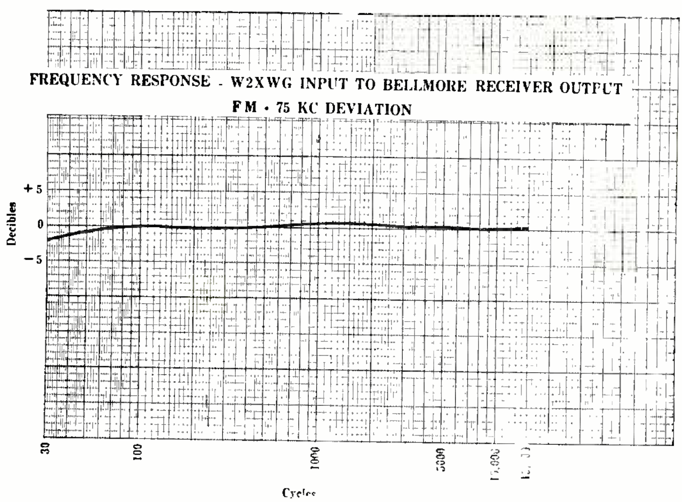


Fig. 4

modulating frequency. This method is simple and very satisfactory if reasonable precautions are taken.

THE RECEIVERS

In order that an exacting comparison of amplitude and frequency modulation can be made, it is very desirable that the receivers for the two systems be as nearly ideal and alike as possible. This practice minimizes otherwise possible errors due to differences in receiver performance. The ideal way to build such receiving systems would be the

[2] A Method of Measuring Frequency Deviation, M. G. Crosby, RCA REVIEW, pp. 473-477, April, 1940.

common use of as many parts as possible. This was done. They were built on the same chassis. The r-f amplifier, first oscillator and the converter were used to drive two different i-f systems in parallel. Switching and chassis space was provided for a third i-f system but it was not used. The i-f systems were used separately and each was designed particularly for the reception of the modulating system to be tested. Each contained an F-M and an A-M detector and each detector had a separate audio output amplifier. The outputs of these audio

Fig. 5

amplifier tubes were all in parallel, but the system desired was selected by screen grid blocking of the other amplifiers. For AM and FM-15, parts of the same i-f system were used but the detectors and limiters were separate. For FM-75 a separate i-f system was used. This was also provided with a separate A-M detector, but it was little used because the 150-kc i-f passband was not representative of amplitude receivers. One meter was provided to show the diode currents and another was provided to show the discriminator currents. The latter is a zero-center meter which indicates zero discriminator current when the receiver is exactly tuned for FM. Separate gain controls were provided for the i-f systems and also for the individual audio amplifiers. The receiver output circuit was provided with a very sharp 8-kc low-pass filter with a disconnecting key and the high-frequency de-emphasis circuit at this point was also provided with a disconnecting key. A single switch served to select the type of modulation which it was desired to receive.

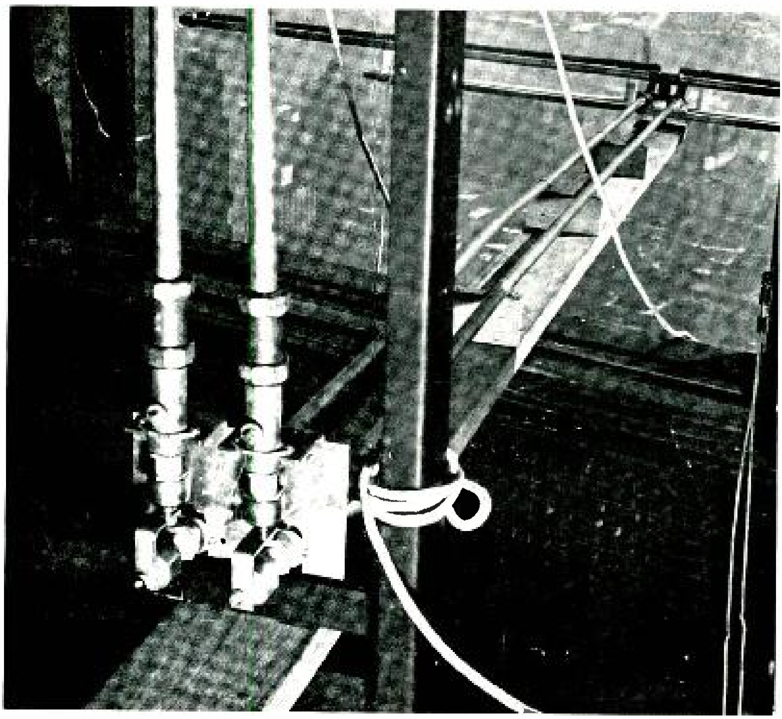


Fig. 6

Four receivers were specially built by the RCA Manufacturing Company for this test and would theoretically give full output with an r-f voltage of only 1/10 microvolt across the input terminals. The hiss level of 1 microvolt peak in these receivers was representative of the best modern receivers and was produced by thermal agitation in the antenna circuits and by tube hiss. They were made as good as receivers can be built in order that the final conclusions concerning frequency modulation as a system would not be erroneous due to apparatus shortcomings. The sacrifice of receiver design to price will not permit

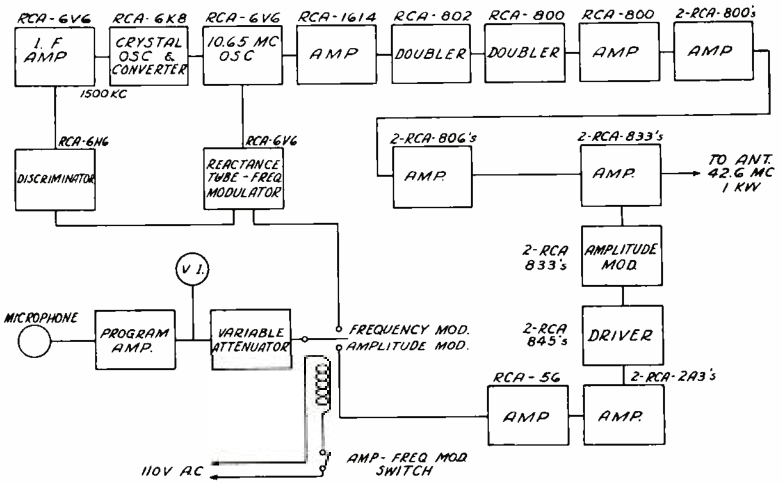


Fig. 7

the full gain of frequency modulation, as reported herein, to be realized.

A block diagram of the duplicate receivers is shown in Figure 8. A photograph of one of the receivers is shown in Figure 9. At Bellmore and also in the NBC laboratory in Radio City there were provided and used a special F-M receiver, dipole antenna and trans-

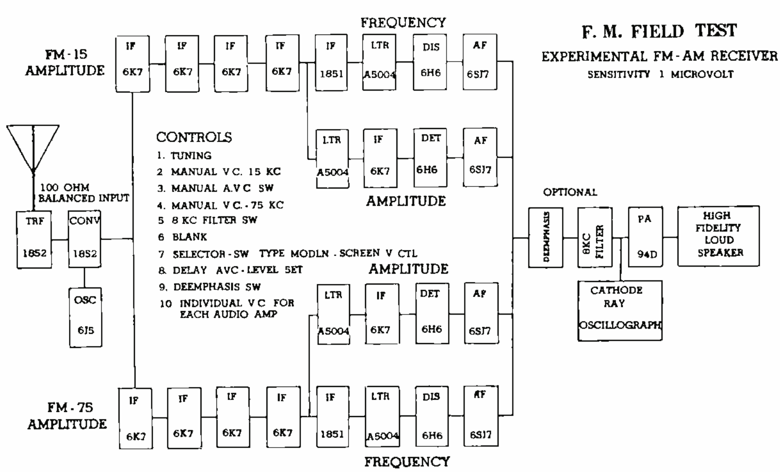


Fig. 8

mission line, commercial receivers of various makes, cathode-ray oscillographs, a special high-fidelity high-power audio amplifier, an RCA audio oscillator, an RCA noise and distortion meter, a harmonic analyzer, disc-recording equipment, a high quality loudspeaker, unweighted volume indicators, volume indicators weighted for ear response, a u-h-f signal generator, a u-h-f field intensity measuring (RCA type 301A) set, r-f transmission line calibrated attenuators, noise-producing devices including automobiles, diathermy machines,

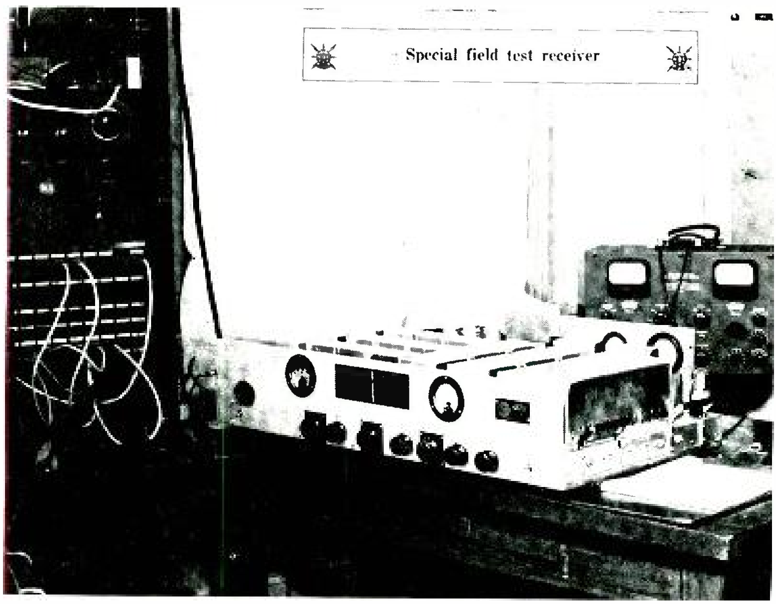


Fig. 9

etc., audio voltage amplifiers, balanced variable pads, and a microphone for recording.

The i-f systems in these receivers were painstakingly designed and adjusted for FM-15, FM-75 and AM, and at intervals during the field tests they were checked with respect to filter characteristics, i-f characteristics, limiter action, discriminator performance, audio frequency response, distortion, etc.

A shielded mutual inductance type antenna attenuator was built and calibrated for controlling the amount of signal or noise voltage reaching the Bellmore receiver input terminals from the antenna. By controlling the power of W2XWG, the antenna attenuator and the noise sources themselves, any desired carrier or noise voltages could be produced. Separate means were provided for controlling the noise amplitudes without changing the character of the noise.

Figure 10 shows measurements made with the special RCA field test receiver and two commercial receivers. Receiver input microvolts are plotted against audio output level. The drop in the commercial receivers is due to r-f gains insufficient to operate the limiters at low input

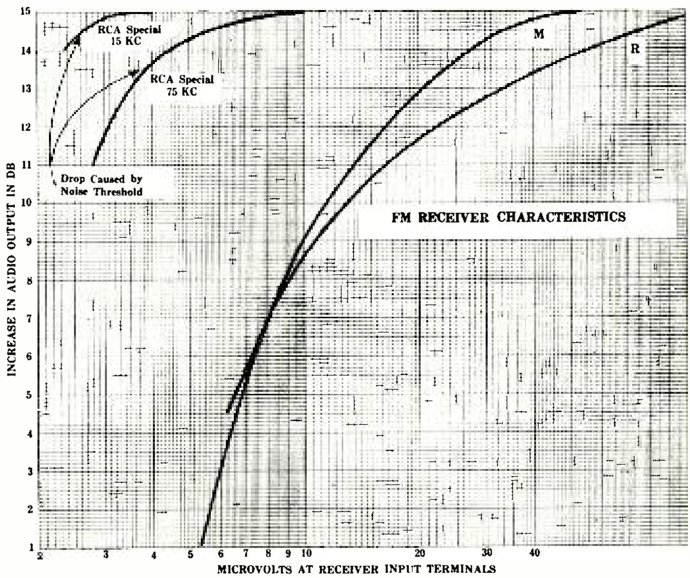


Fig. 10

voltages. The drop in the field test receiver is due to the noise threshold limitation and not lack of r-f gain.

THE FIELD INTENSITY SURVEY OF W2XWG

At the Bellmore receiving location all observations and measurements were correlated with field intensity in microvolts per meter, and also with the number of microvolts across the receiver input terminals. Similar measurements were also made of the noise voltages. In order that all of these measurements could be directly related to miles service radius, a field intensity survey was made of W2XWG. The measurements were carried out to 3.5 microvolts per meter, corresponding to a distance of approximately 85 miles, as shown on the survey map of Figure 11. The radiation index for W2XWG is 910. This is defined as the product of the antenna height, the antenna gain and the square root of the power in kilowatts.

One of the Radio Facilities engineering cars used by NBC for such measurements is shown in Figure 12. Each of these cars is a passenger sedan in which the body and interior has been modified to house a 75B RCA 500-kc to 20-Mc field intensity measuring set, a 301A RCA 18-Mc to 155-Mc field intensity measuring set, a modified RCA 302A

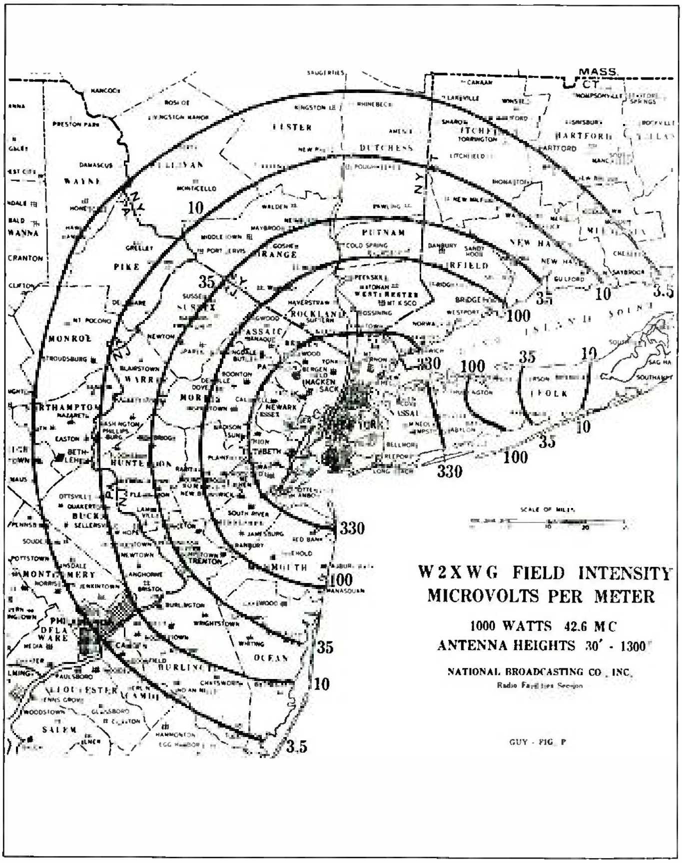


Fig. 11

noise meter, 2 Esterline Angus recorders, a driveshaft-to-recorder gear and clutch system, a special vernier speedometer and distance recorder, an automobile receiver, an aviation type broadcast loop, a u-h-f universally mounted antenna, a swivel-mounted searchlight, a compass and a steel mounting frame with aviation shock absorbers for instruments. Quick-demountable mechanisms permit easy removal of measuring gear.

The Measuring Locations

As a part of this project, field tests and electrical transcriptions were made under a variety of conditions at the following locations:—

Collingswood, N. J.	—85 miles	(temporary station)
Hollis, L. I.	—12 miles	(temporary station)

Fig. 12

Floral Park, L. I.	—15 miles	(temporary station)
Port Jefferson, L. I.	—50 miles	(temporary station)
Commack, L. I.	—36 miles	(temporary station)
Riverhead, L. I.	—70 miles	(temporary station)
Hampton Bays, L. I.	—78 miles	(temporary station)
Bridgehampton, L. I.	—89 miles	(temporary station)
Eastport, L. I.	—65 miles	(temporary station)
Bellmore, L. I.	—23 miles	(permanent station)
NBC Laboratory	— 1 mile	(permanent station)

In addition to making thousands of measurements at the permanent stations, field intensity measurements, listening tests and orthacoustic recordings were made at each of the temporary stations. The recordings compared on one disc, in each series, AM, FM-15 and FM-75. Several discs were recorded at each location under various noise conditions, including the random neighborhood noise encountered. For the observations at the temporary stations a two-car "F-M caravan" was assembled, equipped and moved from station to station. Figure 13 shows the two cars used to transport the equipment and also shows one of the typical locations selected. Figure 14 shows the Bellmore

Fig. 13

receiving station, at which most of the measurements reported herein were made. This receiving station is in a typical suburban neighborhood 23 miles from the transmitter atop the Empire State Building in New York City.

At the Bellmore station a simple horizontal dipole antenna was mounted 25 feet above the ground on a wooden pole. At the temporary receiving stations the receiving antenna consisted of a tripod-mounted dipole. The temporary station locations were selected over a range of distances which covered all grades of service ranging from zero to excellent and a sufficient number of stations were used to insure authoritative conclusions. The observations, orthacoustic recordings and comparisons made at these temporary stations provided a broad overall picture of u-h-f broadcasting without which this thorough field test would have been incomplete. The recordings remain as permanent exhibits of the results.

F-M THEORY

The theory of FM has been presented in the literature[3,4,5]. However, since the following pages compare the actual performance in the field with the theoretical, there is presented for the convenience of the reader a very brief review of the reasons why FM is superior, and the nature of the superiority.

The advantages of FM over AM in noise suppression are contributed by three factors:

1. The triangular noise spectrum of FM.
2. Large deviation ratios.
3. The greater effect of de-emphasis in FM compared to AM.

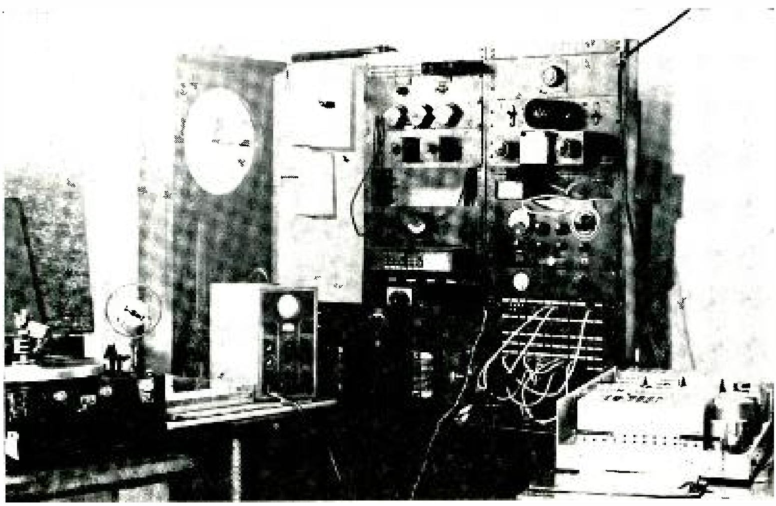


Fig. 14

THE TRIANGULAR NOISE SPECTRUM

An F-M system with a deviation ratio of one has an advantage in signal-to-noise ratio of 1.73 or 4.75 db for tube hiss or other types of fluctuating noise.

Tube hiss consists of a great many closely overlapping impulses. When combined with a steady carrier of fixed frequency, the noise

[3] A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation, Edwin H. Armstrong, *Proc. I.R.E.*, pp. 689-740, Vol. 24, May, 1936.

[4] Frequency Modulation Characters, Murray G. Crosby, *Proc. I.R.E.*, pp. 472-514, Vol. 25, April, 1937.

[5] The Service Range of Frequency Modulation, Murray G. Crosby, RCA REVIEW, pp. 349-371, January, 1940.

peaks beat with it. In the following it is convenient to consider an individual noise frequency as a separate carrier, of which there are many present at any time.

Since a combination of two station carriers differing in frequency is equivalent to a station carrier and a single noise voltage, both cases may be considered at the same time. The effect is most easily shown and understood by means of a simple vector diagram.

The desired carrier vector continuously rotates through 360 degrees and is indicated in Figure 15. A weaker carrier, or a noise voltage, rotates around the carrier vector at a frequency which is equal to the difference between the two. Amplitude modulation is produced as shown. As the undesired vector rotates around the desired vector,

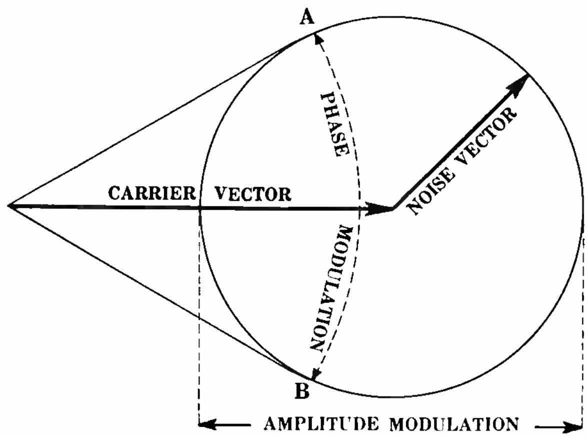


Fig. 15

phase modulation also is produced between the limits A and B. The faster the undesired vector rotates, or the faster the rate of phase change becomes, the greater becomes the momentary change in frequency and, therefore, the greater the frequency modulation becomes, because frequency modulation is a function of the first differential of phase modulation. Therefore, the amplitude of the F-M noise or beat note varies directly with beat frequency. This results in a triangular noise spectrum.

In amplitude modulation there is no such effect as this. All noise components combine with the carrier equally, resulting in a rectangular noise spectrum. The ratio of r-m-s fluctuation noise voltages in FM and AM is, therefore, the ratio between the square root of the squared ordinate areas of these spectrums. This ratio is 1.73 or 4.75 db.

The Deviation Ratio

The deviation ratio, or modulation index, is obtained by dividing the maximum carrier deviation by the highest audio frequency transmitted. For an F-M system, the suppression of fluctuation noise is directly proportional to the deviation ratio. In Figure 16 the A-M noise spectrum corresponds to the total hatched area below 15 kc because the i-f and a-f system would cut off there. The FM-75 receiver i-f system actually accepts noise out to 75 Mc and it has the usual F-M triangular noise characteristic. However, the audio amplifier and the ear respond only to noise frequencies within the range of audibility, around 15 kc, and reject everything else. Therefore, the FM-75 noise we actually hear corresponds only to the small cross-hatched triangle.

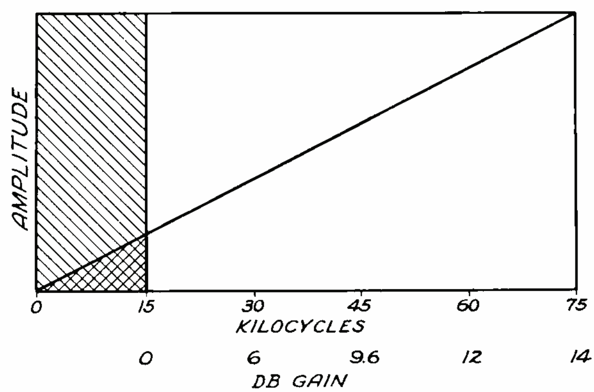


Fig. 16

The maximum height of this F-M triangle, corresponding to voltage, is only one-fifth of the height of an FM-15 triangle or the A-M rectangle. Such being the case, the FM-75 advantage over FM-15 is 5 to 1, or 14 db, and over AM it is 1.73 x 5 or 18.75 db.

De-emphasis

The use of a 100-microsecond filter to accomplish this high-frequency pre-emphasis and de-emphasis has been adopted as standard practice in television and u-h-f sound broadcasting by the Radio Manufacturers Association and recently by the FCC.

It was shown that in FM the noise amplitude decreases as its frequency decreases whereas in AM it does not. Therefore, de-emphasis is more effective in FM. This is shown in Figure 17.

The full rectangle at the left is the A-M noise spectrum. The full triangle at the right is the F-M spectrum. The application of de-emphasis reduces these areas to those combining the hatched and black sections. Squaring those ordinates gives the black areas, corresponding to power. Extracting the square root of the ratios of these black areas gives the r-m-s S/N advantage of FM over AM. It is slightly over 4, corresponding to about 12.1 db. This includes the gains contributed by both the triangular noise spectrum and de-emphasis. The spectrum advantage is 4.75 db. Hence the de-emphasis advantage is 12.1 db minus 4.75 db or 7.35 db. To sum up, for hiss noise the F-M noise spectrum advantage is 4.75 db, the de-emphasis advantage is 7.35 db and the deviation ratio of FM-75 is 14 db, giving a total of about 26 db.

It has been reported[5] that the use of 100 microsecond pre-emphasis produces overswing or overmodulation of from 2.5 to 4.5 db with

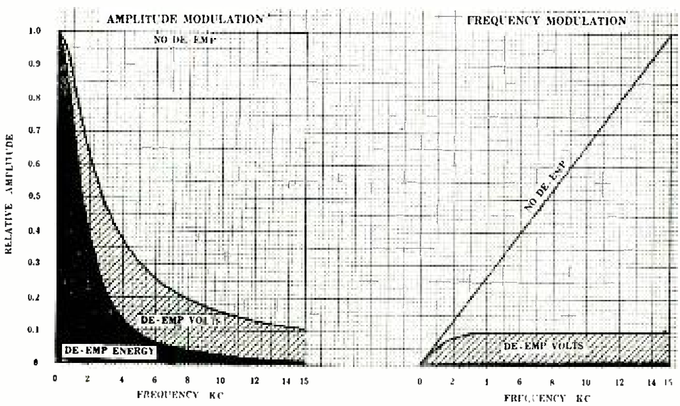


Fig. 17

program modulation, depending upon the character of the sound source. The NBC field test confirmed these conclusions. Therefore, in presetting transmitter gain controls, using low-frequency tone modulation, a 2.5 db correction should be made for program modulation.

Results of the F-M Field-Test Measurements

One of the first facts sought and determined was the lowest field intensity which could provide good service if no external r-f noise were present, and receiver hiss only were the ultimate limiting factor. Figure 18 shows the ultimate limit in service range due to the receiver noise alone, in the complete absence of any noise received on the antenna. The signal-to-noise ratio is shown for the three systems over a significant range of receiver input microvolts, field intensity in microvolts-per-meter and miles-service range. The signal-to-noise ratios were rated by subjective listening tests and the ratings are shown on the right side of the figure. The noise threshold values are indicated in such a manner that they show where the increase of noise with modulation becomes severe but do not indicate the absolute value where an insignificant increase of noise results. The curves are shown dotted below the threshold value because the increase of noise with

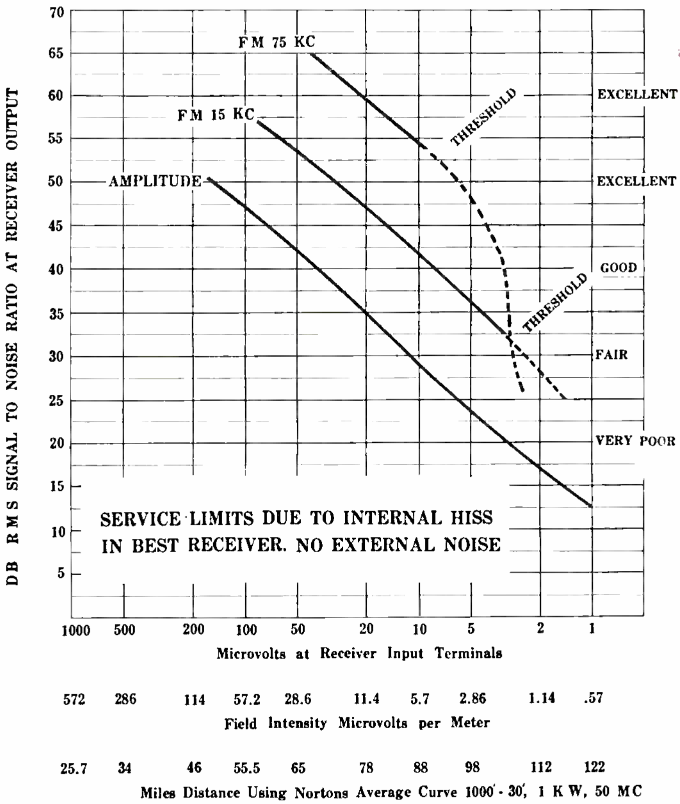


Fig. 18

modulation is very severe. A study of these curves will show the ratio of field intensity or power necessary to produce equivalent performance on the three types of modulation plotted. It is possible to determine directly the distance at which equivalent grades of service may be obtained with the three types of modulation plotted. The three curves may be projected on a straight line toward the upper left corner of the figure if it is desired to compare the results at very high signal-to-noise ratios.

The noise level in the receiver may be read directly from this figure. For instance, where there are two microvolts at the receiver input wise retain their characteristics, if the noise were predominantly steady in character. Here again it may be observed that the full theoretical gain of FM-15 and FM-75 over AM is obtained.

In order that this figure may be more easily interpreted, the same data has been plotted on a bar chart comprising Figure 20. This shows the number of miles service range for the types of modulation shown and also includes calculated values for a frequency-modulation system with a deviation of 30 kc. This also shows the service range with powers of 5 kw and 50 kw, assuming that the same neighborhood noise level existed at all receivers.

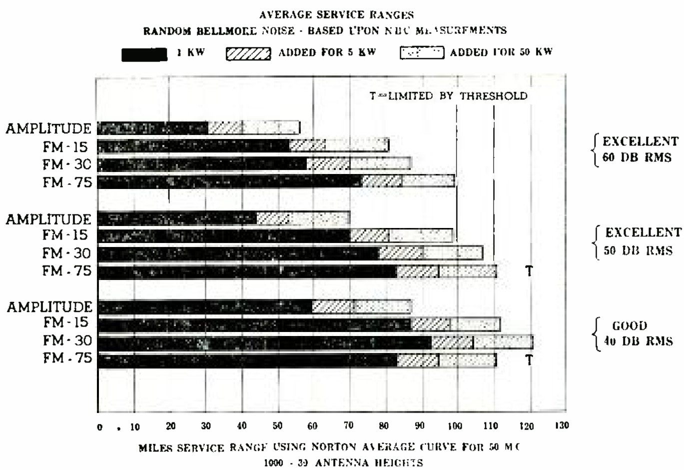


Fig. 20

It will be seen that with a signal-to-noise ratio of 40 db, FM-30 produces a greater service range than any other modulating condition, assuming that in each case the receiver is designed specially for the system being received. This would indicate that if 40 db r-m-s signal-to-noise ratio is considered satisfactory as a minimum, FM-30 would be about the optimum deviation to use. The service range for a swing of 150 kc does not extend beyond the limit imposed by the noise threshold and therefore FM-75 is limited to the distance at which a S/N ratio of 53 db is obtained in all cases.

The threshold effect actually starts on FM-75 at about 60 db but does not become severe until the unmodulated S/N ratio is about 53 db. Therefore, the threshold is indicated on the curves as 53 db.

If neighborhood noise were to be greater than that measured at Bellmore, these bars would all shorten, but would retain their relative terminals the signal-to-noise ratio on AM is 17 db. Therefore, the noise is 17 db below two microvolts RMS. This curve was made using a signal generator as a transmitter, feeding directly to the input terminals to block out all external noise. It may be seen that the full theoretical gain of FM-15 and FM-75 over AM was obtained.

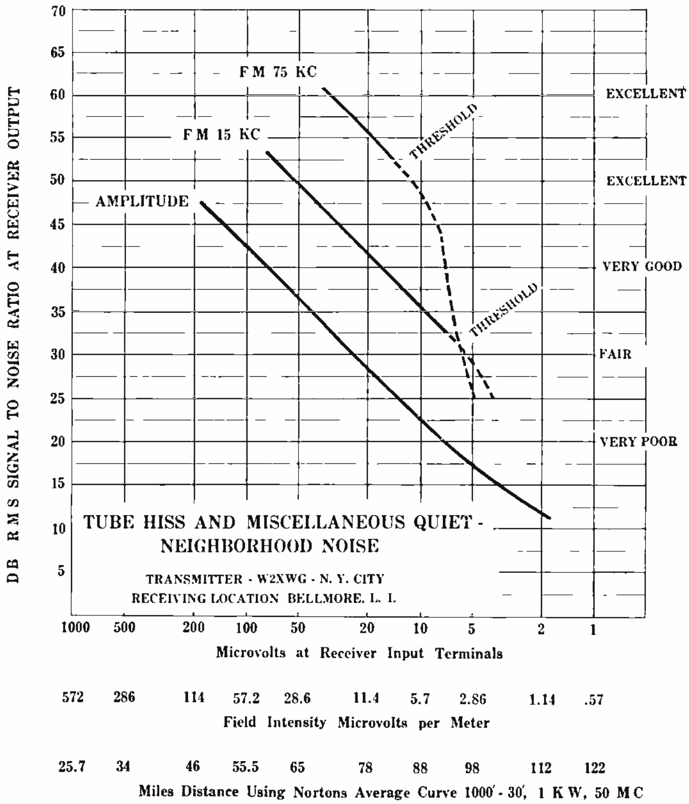


Fig. 19

Figure 19 shows the results of a series of measurements made under conditions which would be representative of a receiver operated in a home. The dipole was used and the noise received was a combination of random Bellmore neighborhood noise plus receiver hiss and thermal agitation. This receiver was located about 400 feet from the nearest public street and there was little automobile traffic in the neighborhood. The measurements were not made during a time when the noise was particularly low but the noise was taken as it existed, at random times. In a location with a higher noise level the three curves plotted on this figure would slide to the left but would otherlengths. It will be noted that FM-75 is superior, even with the threshold limitation, at signal-to-noise ratios of 50 or more decibels.

Regardless of the carrier-to-noise or the signal-to-noise ratio coming out of the discriminator, the full benefits of FM cannot be obtained unless the audio amplifier hum level is sufficiently low. The advantages of FM do not extend into the audio amplifier.

Figure 21 shows measurements of tube hiss. The figure shows the ratio of r-m-s to peak values using an 8-kc audio band width and a 15-kc audio band width with the triangular F-M noise spectrum and de-emphasis. The hiss levels shown are representative of what can be expected from the best modern high-gain receiver.

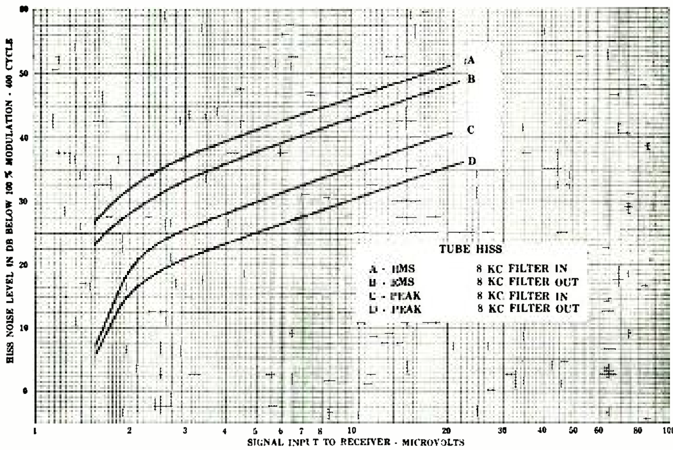


Fig. 21

From the curve, it can be observed that the ratio of peak to r-m-s values with 8 kc is slightly over 10 db. With 15 kc it is slightly over 13 db. It can also be seen that the r-m-s noise is reduced 3 db when the pass-band is reduced from 15 kc to 8 kc. The peak noise is reduced 5 db under the same conditions.

Operation of Two F-M Stations on the Same Channel

By referring to the section covering noise interference it can be seen that the worst condition of shared channel operation occurs when both stations are unmodulated, and a fixed beat note, therefore, results. It will also be seen that the higher this beat note the greater will be its amplitude up to about 5,000 cycles. Figure 22 was made on the basis of the worst conditions, which occur when the difference in carrier frequency reaches approximately 5,000 cycles. Were it not for the effect of de-emphasis in the receiver the beat note amplitude would continuously rise with frequency. However, de-emphasis of the high frequencies prevents that from happening. The effect may be further understood by referring to the section on pre-emphasis and de-emphasis. It will be noted that the noise on the desired station caused by the undesired station varies inversely with the deviation ratio. Here again the theoretical advantage of FM was obtained in our field test.

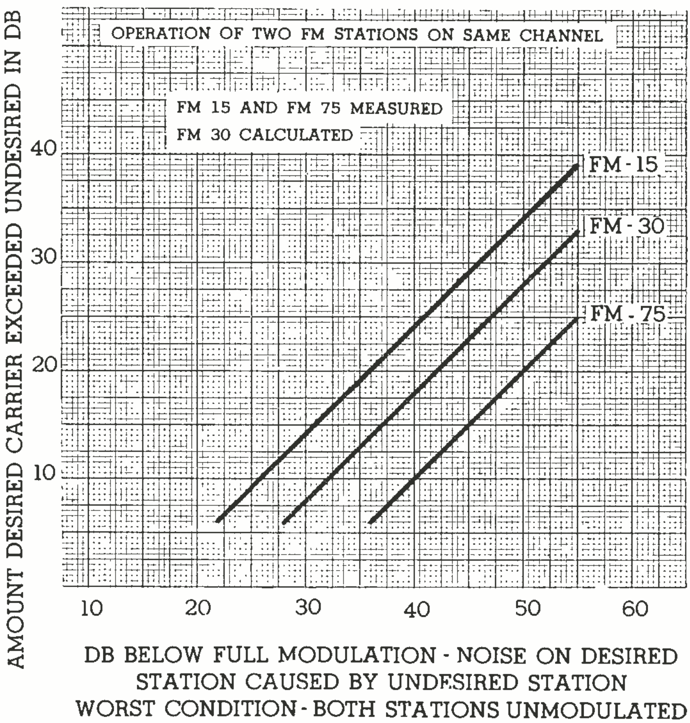


Fig. 22

When either of the stations producing the beat note becomes modulated, the beat note disappears, because one carrier sweeps across the other one. When the desired station is approximately 20 db stronger than the undesired station, all interference and cross talk effects become unnoticeable. At 12 db difference they are noticeable, but it is the opinion of some engineers that the 12 db ratio would be tolerable. Frequency modulation offers a great advantage over amplitude modulation in the allocation of stations on the same frequency. In AM the carrier amplitude of the desired station must be 100 times, or 40 db greater than the undesired carrier amplitude for a 40 db signal-to-beat note ratio. For FM-75 it need be only 10 db, or 3 times greater; for FM-30, 17.5 db, or 8 times greater; for FM-15 24 db, or 10.5 times greater. The result is that F-M stations may be located much closer geographically than A-M stations, and therefore many more station assignments can be made per channel.

Operation of Frequency-Modulation Stations on Adjacent Channels

Figure 23 shows the results of listening tests on adjacent F-M channels using in each case FM-75. On the basis of this information it is

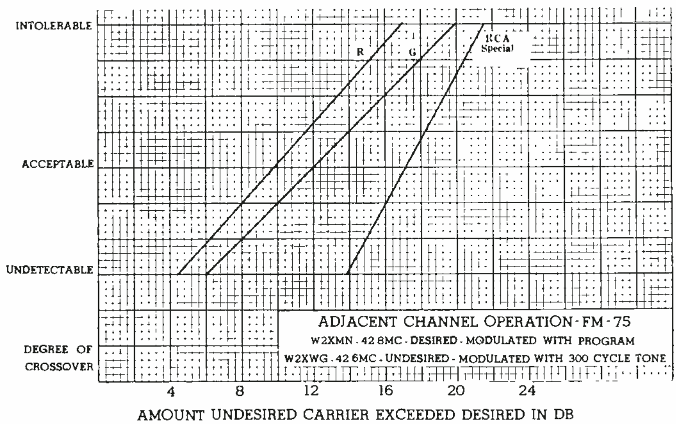


Fig. 23

seen that adjacent channel stations must not be located in the same geographical area. The channels were 200 kc wide and were adjacent, one being W2XMN on 42.8 Mc and the other W2XWG on 42.6 Mc. W2XMN was used as the desired channel, since its field intensity was constant at Bellmore. W2XWG was used as the undesired station. The field intensity ratios were adjusted as desired by varying the power of W2XWG. Observations were made of the special field-test receiver and also of two commercial receivers with the results shown.

The figure shows that the undesired carrier level should be not more than about 10 db greater than the desired carrier level to prevent objectionable cross-talk in the commercial receivers. The RCA special field-test receiver will give equivalent performance at a carrier level ratio of 17 db, but this receiver is more elaborately built and is superior in certain respects to commercial models.

Intolerable cross-talk occurs on all three receivers at carrier level ratios of from 17 to 21 db.

The Noise Threshold in Frequency Modulation

Crosby points out that an interesting series of events takes place in a frequency-modulated system when the noise peaks equal or exceed the peaks of the carrier.[4] The result is a rapid increase of the noise

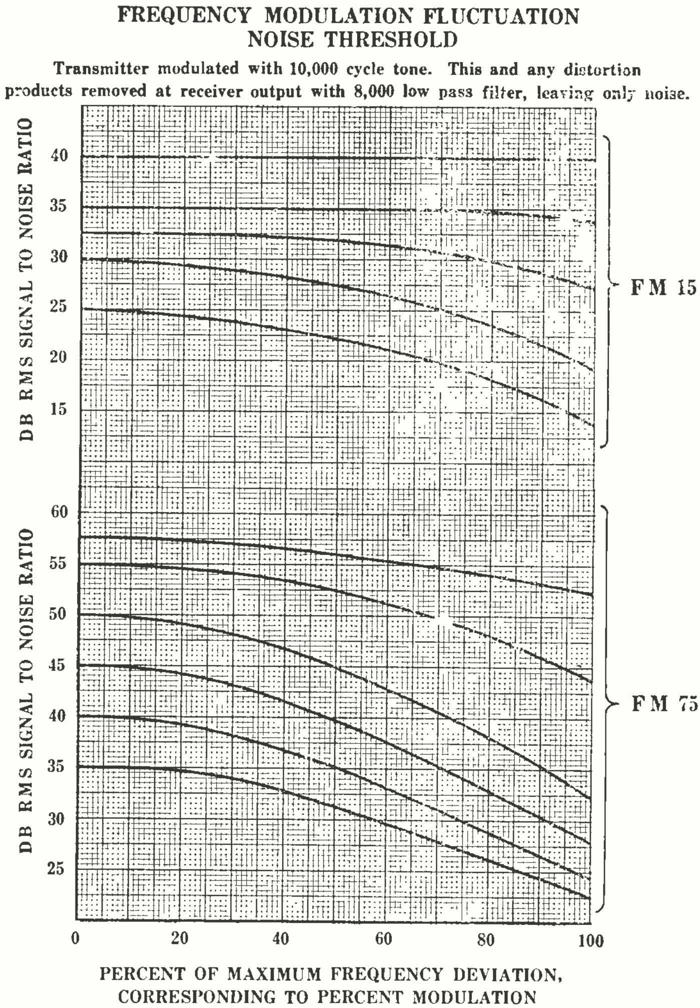


Fig. 24

level or decrease of the signal-to-noise ratio, with modulation. In frequency modulation wherein the maximum swing is 150 kc the point where this begins to occur is reached when the unmodulated signal-to-noise ratio is about 60 db. When the unmodulated signal-to-noise ratio is less than about 60 db, or 1,000 to 1, the noise level rises with modulation, and, as the noise peaks exceed the carrier peaks by a considerable amount, this noise level may increase almost 20 db, or 10 times. When operating above the threshold limit the noise changes little as the

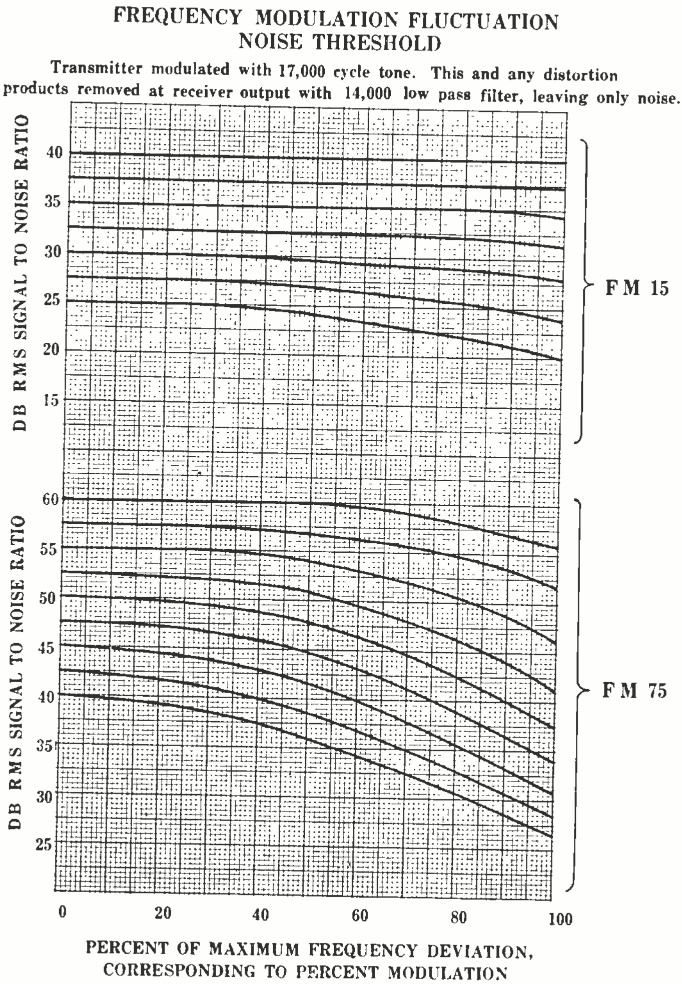


Fig. 25

station is modulated. Below the threshold limit the effect is not unlike severe harmonic distortion in an overloaded A-M transmitter.

In frequency modulation of a lesser swing, such as 30 kc, the same effect occurs. In this case, however, the threshold limit occurs at about 35 db signal-to-noise ratio. Figure 24 shows the results of some of the measurements made at Bellmore. In order that the noise would not be confused with any small amount of inherent distortion in a man-made system, the measurements were made in such a manner that the effects of distortion were eliminated. This was done by modulating the trans-

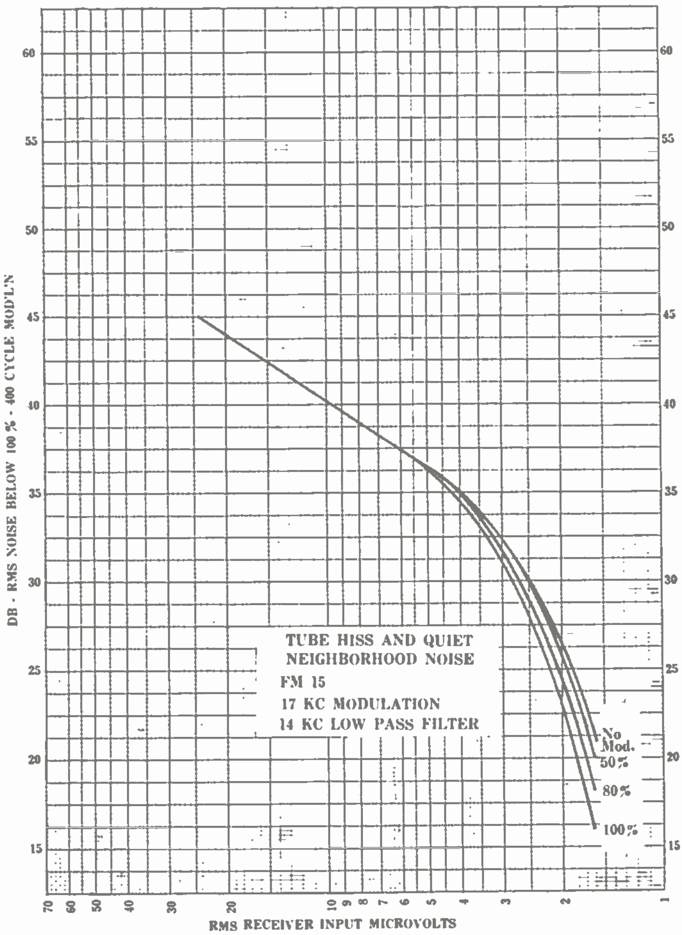


Fig. 26

mitter with a 10,000-cycle tone and eliminating at the output of the receiver, with an 8-kc low-pass filter, not only the fundamental modulating tone but also the distortion products, leaving only the noise. Figure 25 shows the results of another set of measurements made with a 17-kc modulating frequency, and a 14-kc low-pass filter to eliminate the fundamental tone and distortion products.

This effect has no doubt been observed by many without being understood. It is an inherent characteristic of a frequency-modulation system. The noise threshold in the case of an FM-40 system having a

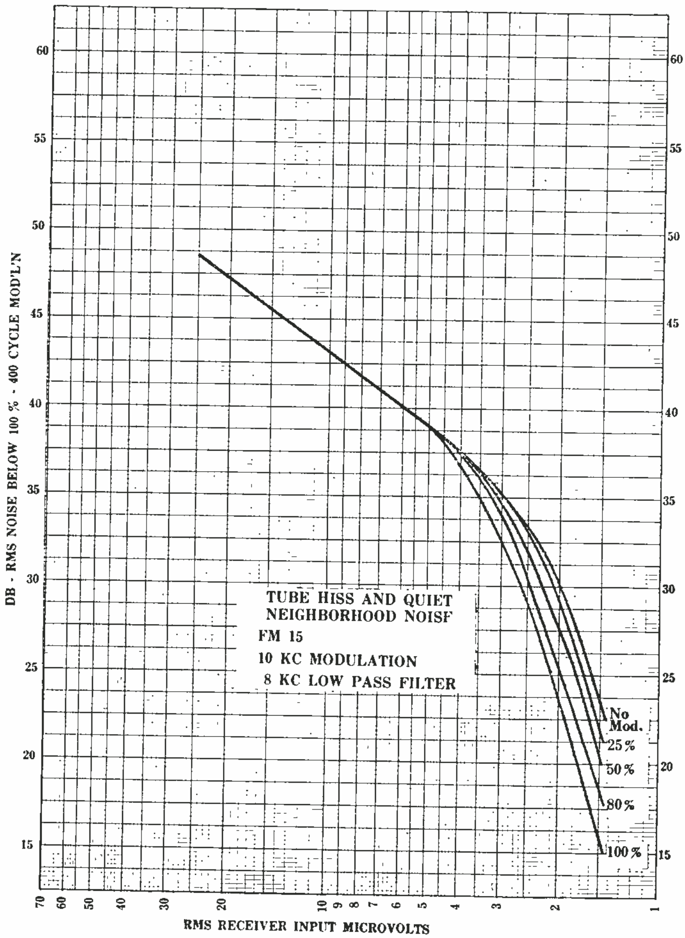


Fig. 27

total band width of 100 kc occurs at about 43 db. This provides a very good signal-to-noise ratio.

Figures 26-27 give additional results of threshold r-m-s measurements showing noise levels plotted against receiver input microvolts, with various percentages of modulation. The signal-to-noise ratio (ordinate) is the ratio of maximum 400-cycle modulation to noise.

Measurements of Peak Ignition Noise

Because of the peculiar wave shape and large crest factor of ignition noise it is preferable to measure the peak signal and peak ignition noise rather than the r-m-s values in order to establish, for one thing, the

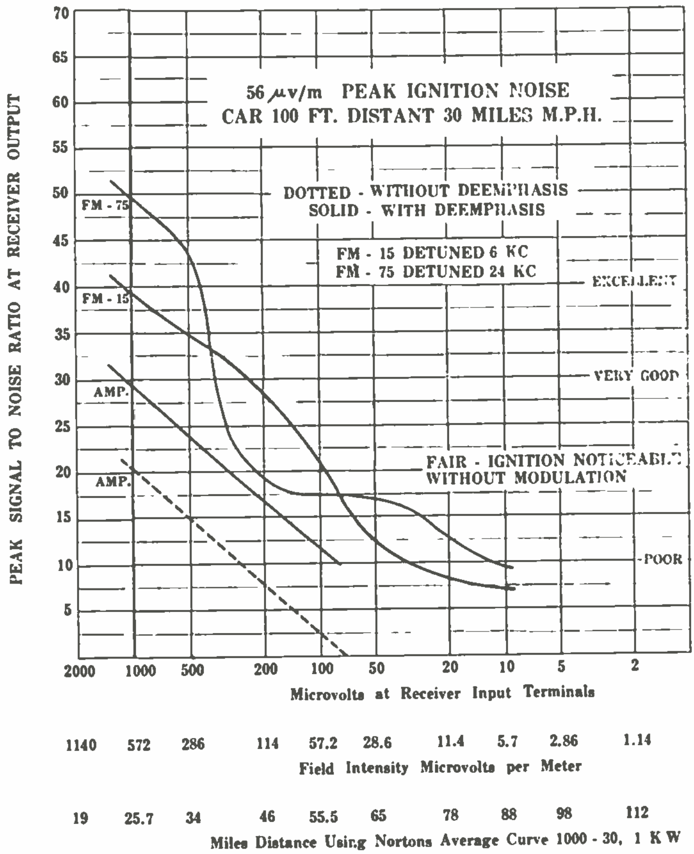


Fig. 28

threshold where they become equal. Because of the infrequent number of peaks, compared with hiss noise, much higher values of peak noise levels can be tolerated from ignition systems. In making measurements of ignition noise an actual automobile was used. Making such measurements is quite difficult because of the variation of peak noise amplitudes from an automobile system over short periods of time. Also, to show what the dynamic noise characteristics of a system would be without actually modulating it, it becomes necessary to de-tune the receiver or resort to some other expedient. This is necessary because high noise peaks, if synchronized with an unmodulated carrier, do not show the existence of the noise threshold. Modulation, in effect, de-tunes the receiver and the threshold becomes evident. De-tuning the receiver in the absence of modulation is one expedient which produces a similar result and was the method used in obtaining the data shown on Figure 28.

Of particular interest in this figure is the rating of the signal-to-noise ratio as shown at the right. A 30 db signal-to-noise ratio, when measured with peak values, is equivalent to a 40 db r-m-s signal-to-noise ratio. Even with as low a signal-to-ignition-noise ratio as 20 db the service is quite fair, although the ignition would be noticeable without modulation. The relative infrequency of ignition peaks produces an audible result which is very deceiving. With signal-to-ignition-noise ratios of 10 or 12 db, service is still not completely ruined but could be tolerated if there were a special interest in the program material.

In general, ignition noise is transient, lasting for only a matter of seconds as a car passes by a residence. During that period the noise is not distressing. Furthermore, over a period of years, it may be expected that automobile ignition systems will be provided with suppressors which will reduce the u-h-f interference by at least 20 db. Ignition noise is of particular concern and is the predominant noise in suburban areas. In urban areas, the field intensity from a F-M station or an A-M station will ordinarily be high enough to over-ride the higher noise levels experienced. Figure 28 represents some of the results obtained with 56 peak microvolts per meter noise. The method of making these measurements was not ideal in all respects but the data is indicative of the results obtained in the presence of ignition noise. The curve of Figure 29 shows r-m-s ignition noise measurements made without modulation and illustrates that no threshold is found under such conditions. Since a system is of no value until modulated, this curve is shown only to illustrate the point that the noise threshold must be associated with modulation.

Figure 30 presents interesting data showing peak ignition noise measurements made with an 8-kc audio band width. Peak noise input microvolts are plotted against peak S/N ratio, the field intensity of the station remaining unchanged. The signal with which the noise was compared was 100 per cent 400-cycle tone modulation. The noise source in this case was an automobile ignition system built up and mounted on a lathe at Bellmore. The battery power was not varied to produce different field intensities of noise because this method changed the character of the noise. The field intensity was varied over a very wide range, without changing its normal characteristic, by orienting and changing the length of the connected noise transmitting antenna, which was located about 50 feet from the receiving dipole. FM-15 represents a deviation ratio of 1.875 when the audio band width is 8 kc. As Crosby has shown[5], the noise spectrum advantage of FM is 6 db for impulse noise. The FM-15 threshold is shown. The FM-75 threshold is not shown because at the time the measurements were made a-c hum within

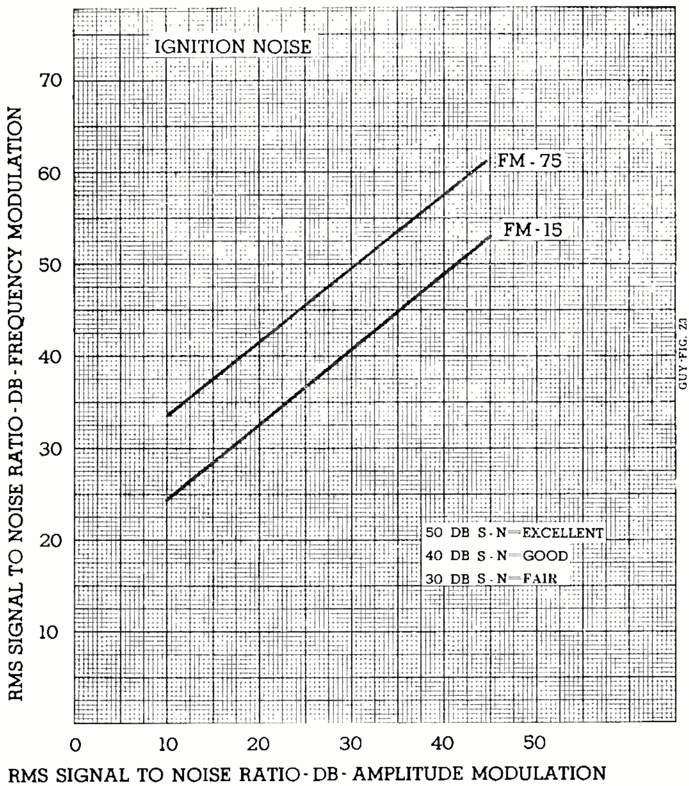


Fig. 29

the system made the accuracy of S/N measurements in the 60-db region uncertain. Of particular interest is the shape of the F-M curves at the lower right side of the figure. The explanation for it is given in the references.

This is to be expected from the character of ignition noise. The impulses are very short in duration, very high in amplitude and relatively widely separated. They literally blank out only small portions of the signal waves, without impairing the remainder. The short blanked out intervals of the signal change little over a wide range in noise peak amplitude. Once an ignition peak has risen to the value required to control the receiver and blank out the signal a further rise in the noise level will not occur until the peak increases in breadth, or duration, or until there is a sufficient rise in certain low amplitude components of ignition noise having fluctuation noise characteristics. The peculiar shapes of such curves below the threshold values are due to the wave shapes and crest factors of ignition noise.

Results of Observations at Temporary Receiving Stations

The observations at the temporary receiving stations confirmed the measurements made at Bellmore. In going to progressively greater distances the FM-75 noise threshold distance was passed and FM-15 became superior. Then the FM-15 threshold distance was passed and AM became superior, although it was very noisy. At the limit of A-M intelligibility both FM-75 and FM-15 were completely smothered by noise. The service limits were all in excellent agreement with those shown by the bar chart, Figure 20, subject to spasmodic interference from passing automobiles. This type of interference was at all stations the predominating one, but was only intermittent and not as troublesome as might be expected.

Observations of Diathermal Interference in AM and FM

Diathermal machines vary considerably in their characteristics, some types using raw a.c. and others using partially filtered power supplies. Observations were made of interference on the three types of modulation using raw a-c machines. It was concluded from these observations that this type of interference is characterized by the transmission of a band of frequencies about 15 kc wide. With amplitude modulation the background interference is essentially equal to the carrier-to-noise ratio. With fairly weak interference from diathermy exactly centered on the desired carrier, FM-75 reception is 20 to 25 db superior to AM and FM-15 is 10 to 12 db superior. However, with the diathermy 5 kc off the desired carrier, AM was approximately equal to FM-15 and was in some cases superior to FM-75. With the diathermy carrier at the edge of the AM and FM-15 passband, the interference is highly attenuated. Under these conditions the FM-75 interference was extremely severe. With the diathermy carrier well outside of the passband of FM-15 or AM, the interference is noticeable only when the diathermy amplitude is extremely high. Under normal receiving conditions it would not be heard. However, if the diathermy under these conditions is within the FM-75 passband, the interference is extremely severe. Thus, under such conditions, FM-75 was the worst of the three types of modulation. It was concluded that in locations having strong diathermy interference, narrow band receiving systems

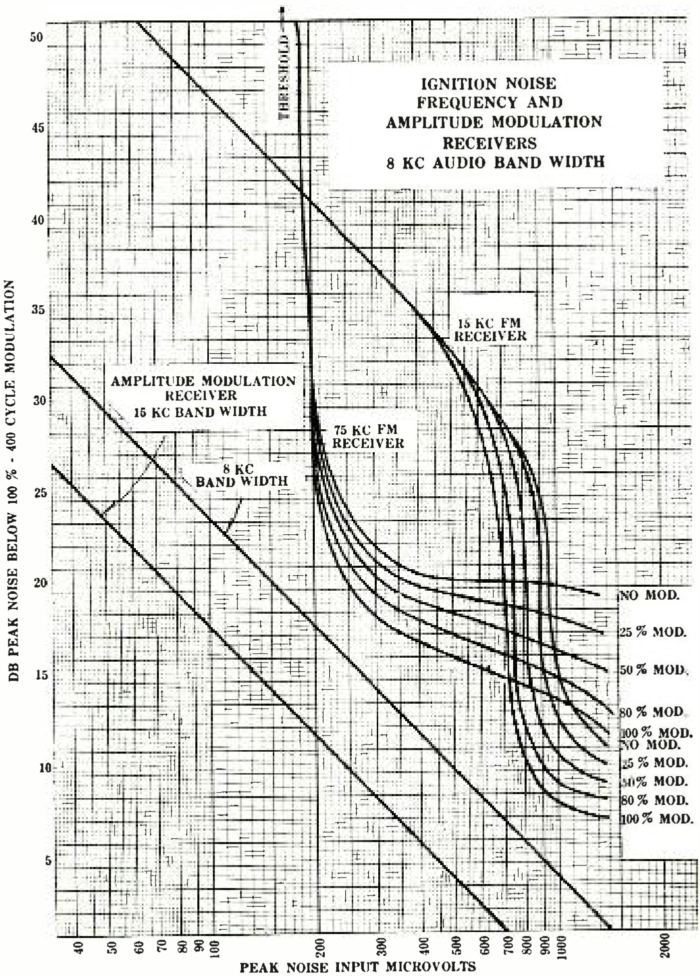


Fig. 30

would be far superior to others. It was also concluded that in locations having weak diathermy interference the wide band system would be much superior.

Noise Levels

In the foregoing, considerable data has been shown to indicate the field intensities which will provide good service for F-M and A-M systems. Figure 31 is shown because it is of particular interest in connection with this field test. It represents an accumulation of noise measurements made over a period of years by various RCA groups. These data were assembled by Dr. H. H. Beverage.

Noise levels vary considerably from time to time and it is not possible to give fixed values for any time or place. However, the informa-

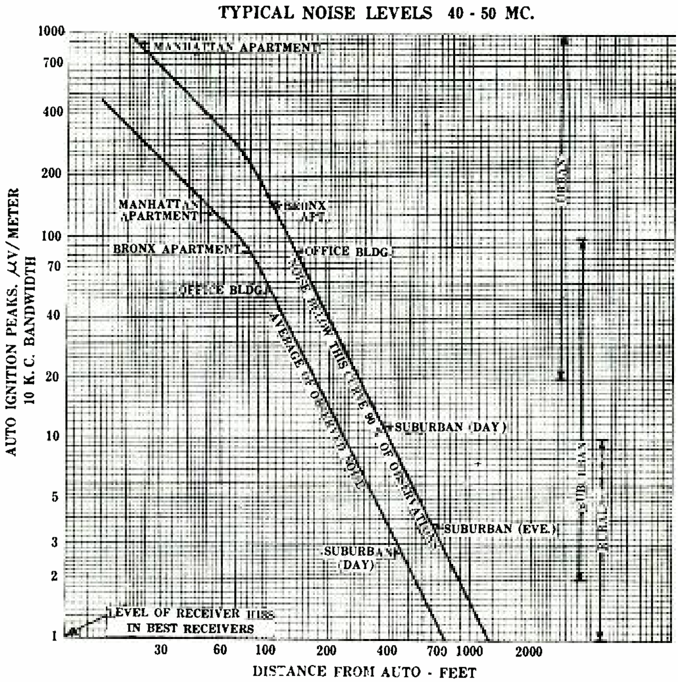


Fig. 31

tion shown on Figure 31 is indicative of the noise levels which may be encountered under a wide range of conditions with a 10-kc audio band width. For peak noise the amplitude varies directly with the frequency band. For fluctuation noise the amplitude varies as the square root of the band width.

Acknowledgment

The authors express their appreciation to Messrs. R. R. Beal and O. B. Hanson for their assistance in making this project possible. Grateful appreciation is also extended to many other NBC and RCA engineers, particularly Mr. L. A. Looney of NBC.

SOME NOTES ON COUPLED CIRCUITS

BY

W. R. FERRIS

Research and Engineering Department, RCA Manufacturing Company, Inc., Harrison, N. J.

Summary—A conventional analysis of the performance of a constant current-fed pair of coupled circuits is made. No claims as to fundamental aspects are made, but the derivation is direct and the formulas are directly applicable to practical circuits. The condition for matched impedances is derived and the expression for the resulting gain is given as

$$\left(\frac{E_2}{E_1}\right) = g_m \frac{\sqrt{r_p r_g}}{2}$$

which is the maximum possible amplification with a tube and circuit having a shunt input resistance r_g, a shunt output impedance r_p, and a transconductance g_m. A curve is included showing the advantage of proper matching as opposed to direct impedance or one-to-one transformer coupling.

THE action of a pair of coupled circuits containing the elements shown in Figure 1 depends upon so many independent variables, eight in the general case, that any attempt to simplify the calculation must involve either approximations or a reduction of the number of independent variables by assuming certain relations between them. The following calculations involve no approximations and the equations are, therefore, in a form suitable for further manipulation and subsequent approximation.

The coupled circuits will be taken as two ideal coils with pure magnetic coupling shunted by ideal capacitors and resistors* as shown in Figure 1. It is assumed that a vacuum tube supplying a constant current, $-g_m E_1$, is connected across the primary. Combining the capacitance and resistance elements, for the moment, as Z_1 and Z_2 in the primary and secondary circuits respectively, we find that the secondary voltage is

$$E_2 = \frac{-g_m E_1 j\omega M}{\left(1 + \dfrac{j\omega L_1}{Z_1}\right)\left(1 + \dfrac{j\omega L_2}{Z_2}\right) + \dfrac{\omega^2 M^2}{Z_1 Z_2}} \qquad (1)$$

* The resistances r_p and r_g are essentially the shunt high-frequency internal plate and grid resistances of the tube. If there are other losses in the tuned circuits they may be lumped as equivalent shunt resistances R_p and R_g and combined with r_p and r_g to give effective primary and secondary shunt resistances R'_p and R'_g. The formulas will then hold if R'_p and R'_g are substituted for r_p and r_g.

where $Z_1 = \dfrac{r_p}{1 + j\omega C_1 r_p}$ $\quad Z_2 = \dfrac{r_g}{1 + j\omega C_2 r_g}$ $\quad$ and $M = k\sqrt{L_1 L_2}$

Substituting these values gives the secondary voltage in the form

$$E_2 = \frac{-g_m E_1 k j\omega \sqrt{L_1 L_2}}{\left[1 - \omega^2 L_1 C_1 - \omega^2 L_2 C_2 + \omega^2 L_1 L_2 (1-k^2)\left(\omega^2 C_1 C_2 - \dfrac{1}{r_p r_g}\right)\right] + j\left[\dfrac{\omega L_1}{r_p} + \dfrac{\omega L_2}{r_g} - \omega^2 L_1 L_2 (1-k^2)\left(\dfrac{\omega C_2}{r_p} + \dfrac{\omega C_1}{r_g}\right)\right]} \tag{2}$$

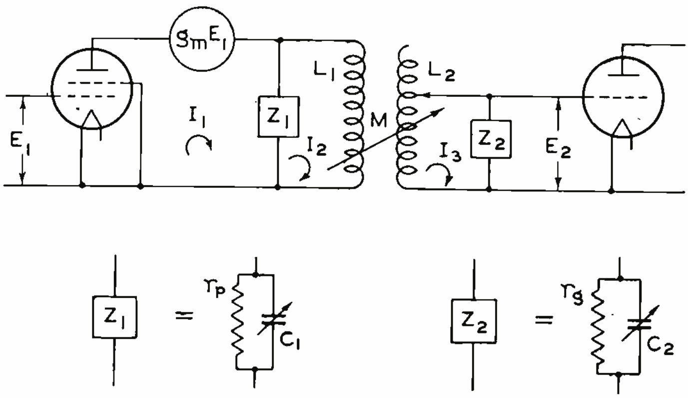


Fig. 1—Transformer-coupled circuit.

The absolute value of E_2 is

$$|E_2| = \frac{g_m E_1 k\omega\sqrt{L_1 L_2}}{\sqrt{\left[1 - \omega^2 L_1 C_1 - \omega^2 L_2 C_2 + \omega^2 L_1 L_2 (1-k^2)\left(\omega^2 C_1 C_2 - \dfrac{1}{r_p r_g}\right)\right]^2 + \left[\dfrac{\omega L_1}{r_p} + \dfrac{\omega L_2}{r_g} - \omega^2 L_1 L_2 (1-k^2)\left(\dfrac{\omega C_2}{r_p} + \dfrac{\omega C_1}{r_g}\right)\right]^2}} \tag{2a}$$

Equation (2a) may be maximized on several parameters as follows:

Differentiating with respect to C_1 and C_2 and setting the derivative equal to zero gives, respectively,

$$\left[1-\omega^2L_1C_1-\omega^2L_2C_2+\omega^2L_1L_2(1-k^2)\left(\omega^2C_1C_2-\frac{1}{r_pr_g}\right)\right]$$
$$\left[1-\omega^2L_2C_2(1-k^2)\right]+\left[\frac{\omega L_1}{r_p}+\frac{\omega L_2}{r_g}-\omega^2L_1L_2(1-k^2)\right.$$
$$\left.\left(\frac{\omega C_2}{r_p}+\frac{\omega C_1}{r_g}\right)\right]\left[\frac{\omega L_2}{r_g}(1-k^2)\right]=0 \qquad \text{(3a)}$$

$$\left[1-\omega^2L_1C_1-\omega^2L_2C_2+\omega^2L_1L_2(1-k^2)\left(\omega^2C_1C_2-\frac{1}{r_pr_g}\right)\right]$$
$$\left[1-\omega^2L_1C_1(1-k^2)\right]+\left[\frac{\omega L_1}{r_p}+\frac{\omega L_2}{r_g}-\omega^2L_1L_2(1-k^2)\right.$$
$$\left.\left(\frac{\omega C_2}{r_p}+\frac{\omega C_1}{r_g}\right)\right]\left[\frac{\omega L_1}{r_p}(1-k^2)\right]=0 \qquad \text{(3b)}$$

These homogeneous equations are satisfied by the following relation obtained by dividing Equation (3b) by Equation (3a).

$$\frac{1-\omega^2L_1C_1(1-k^2)}{1-\omega^2L_2C_2(1-k^2)}=\frac{L_1r_g}{L_2r_p} \qquad \text{(4)}$$

Equation (4) is quite general and may be satisfied in an infinite variety of ways. A common choice is to make $L_1C_1=L_2C_2$, that is, to tune both circuits to the same frequency. Then, Equation (4) becomes $1=L_1r_g/L_2r_p$ and these relations may be written in the forms

$$\frac{L_1}{L_2}=\frac{r_p}{r_g}=\frac{C_2}{C_1} \qquad \text{(4a)}$$

and
$$\frac{\omega C_1r_p}{\omega C_2r_g}\equiv\frac{Q_1}{Q_2}=1 \qquad \text{(4b)}$$

Peak Frequencies

Going back to (2a) and differentiating it with respect to ω and setting the derivative equal to zero, we have

$$\left\{1 - \omega^2 L_1 C_1 - \omega^2 L_2 C_2 + \omega^2 L_1 L_2 (1-k^2)\left(\omega^2 C_1 C_2 - \frac{1}{r_p r_g}\right)\right\}$$
$$\left\{1 + \omega^2 L_1 C_1 + \omega^2 L_2 C_2 - \omega^2 L_1 L_2 (1-k^2)\left(3C_1 C_2 - \frac{1}{r_p r_g}\right)\right\}$$
$$- 2\omega^2 (1-k^2) L_1 L_2 \left(\frac{\omega C_2}{r_p} + \frac{\omega C_1}{r_g}\right)\left\{\frac{\omega L_1}{r_p}\left[1-(1-k^2)\omega^2 L_2 C_2\right]\right.$$
$$\left. + \frac{\omega L_2}{r_g}\left[1-(1-k^2)\omega^2 L_1 C_1\right]\right\} = 0 \qquad (5)$$

Equation (5) is of the fourth degree in ω^2 and therefore Equation (2a) can have at most two maximum and two minimum values when plotted as a function of ω^2.

Taking the special case represented by Equation (4a), we can factor Equation (5) as

$$\left\{\omega^4 (1-k^2) L_1^2 C_1^2 - \omega^2 \left[2L_1 C_1 - \frac{L_1^2}{r_p^2}(1-k^2)\right] + 1\right\}$$
$$\left\{3\omega^4 (1-k^2) L_1^2 C_1^2 - \omega^2 \left[2L_1 C_1 - \frac{L_1^2}{r_p^2}(1-k^2)\right] - 1\right\} = 0 \qquad (5a)$$

The first factor gives the two maximum values and the second a single minimum, one root of the second factor being negative for ω^2, i.e., imaginary for ω.

$$\omega^2 \max = \frac{-\dfrac{L_1^2}{r_p^2}(1-k^2) + 2L_1 C_1 \pm \sqrt{\left[\dfrac{L_1^2}{r_p^2}(1-k^2) - 2L_1 C_1\right]^2 - 4(1-k^2) L_1^2 C_1^2}}{2(1-k^2) L_1^2 C_1^2} \qquad (6a)$$

$$\omega^2 \min = \frac{-\dfrac{L_1^2}{r_p^2}(1-k^2) + 2L_1 C_1 + \sqrt{\left[\dfrac{L_1^2}{r_p^2}(1-k^2) - 2L_1 C_1\right]^2 + 12(1-k^2) L_1^2 C_1^2}}{6(1-k^2) L_1^2 C_1^2} \qquad (6b)$$

Critical Coupling

We may adjust the parameters so that these three values become coincident, that is, a single peak of maximum height may be obtained. This requires that

$$\sqrt{\left[\frac{L_1^2}{r_p^2}(1-k^2)-2L_1C_1\right]^2-4(1-k^2)L_1^2C_1^2}=0 \qquad (7)$$

This equation gives $\omega^2 \max \equiv \omega^2 \min = \dfrac{1}{L_1C_1\sqrt{1-k^2}}$ as the frequency for which the circuits are critically coupled if the conditions in Equation (4a) are satisfied. The critical coupling value is $k = 1/\omega C_1 r_p$ in this case.

Maximum Gain

By substituting Equation (4a) and the first factor of Equation (5a) in Equation (2a) we obtain the significant result

$$|E| = g_m E_1 \frac{\sqrt{r_p r_g}}{2} \qquad (8)$$

This value is independent of the coupling so long as the first factor of Equation (5a) has real roots, in other words, so long as there are two peaks. This is also true for the critical case in which the two peaks coincide. Equation (7) gives the minimum value of coupling which makes this possible and it is the critical coupling for the circuit. Any greater value of coupling gives a double-humped curve, each peak having the height given by Equation (8). Usually the double-humped curve is avoided and the coupling is made such that Equation (7) is satisfied.

Equation (8) takes the place of the usual equation for a straight impedance coupling as, for example, a single-tuned circuit between a plate load r_p and a grid resistance r_g. The usual equation is the familiar

$$|E| = \frac{-g_m E_1}{\left(\dfrac{1}{r_p}+\dfrac{1}{r_g}\right)} \qquad (9)$$

In case $r_p = r_g$, Equations (8) and (9) become identical, but otherwise Equation (9) gives a higher value of $|E|$. The higher value is

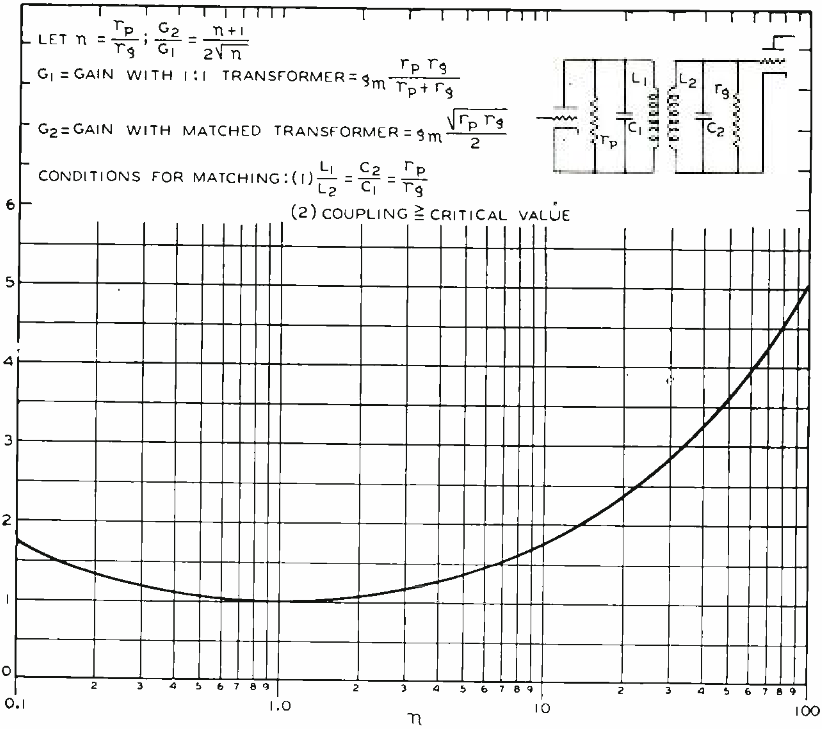


Fig. 2—Curve showing advantage of proper matching in transformer-coupled circuits.

obtained by satisfying the relations in Equation (4a). An inspection of Equation (4a) shows that a "step-down transformer" will give better gain than a straight impedance or a one-to-one transformer if the grid resistance is lower than the plate resistance. If the grid resistance is the higher, a step-up ratio should be used. If the two resistances are not widely different, however, the advantage is not appreciable as may be seen in Figure 2.

A NEW ELECTRON MICROSCOPE

By

L. Marton, M. C. Banca, and J. F. Bender

Research Laboratories, RCA Manufacturing Company, Inc., Camden, N. J.

Summary—*The paper describes a magnetic electron microscope which was built in the Research Laboratories of the RCA Manufacturing Company. This instrument, which is useful in the fields of bacteriology, colloidal chemistry, and other industrial research, incorporates a number of new features as well as others previously described by one of the authors. A resolving power considerably better than 100 Å has been shown in preliminary tests. The "stage" carrying the specimens, which can be moved vertically and horizontally independently of the alignment of the electron optical system, is introduced between the pole pieces of the objective coil in order to take full advantage of the improvement in performance obtained by having the specimen as close as possible to the objective lens. Provision is also made for cooling specimens, when necessary, with liquid air or other refrigerant. The object and photographic chambers are provided with air locks which do not require greased joints, facilitating the rapid changing of specimens and photographic plates.*

THE fundamental discovery of the wave nature of electrons by de Broglie in 1924 started a very intense research development in atomic and molecular dimensions. Of the many fields covered by modern electronic research, three branches seem to the authors to

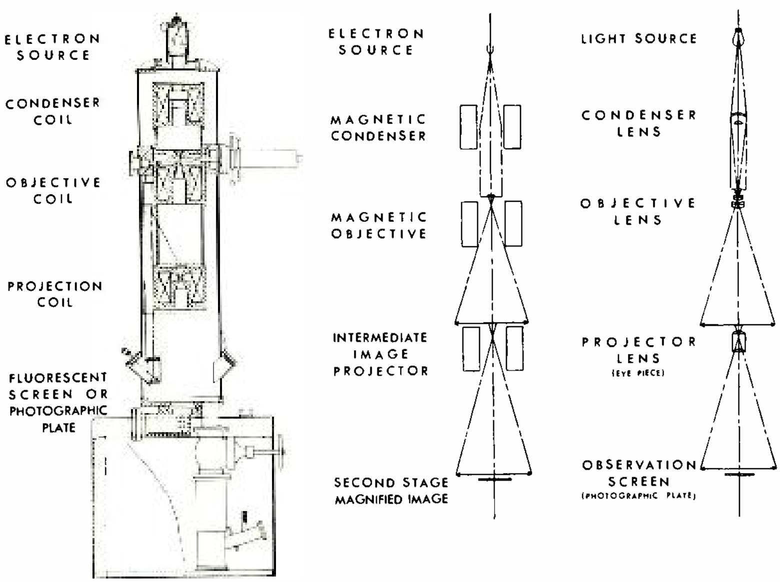


Fig. 1—Diagram of Electron Microscope

be the most important: first, the application of electrons to the exploring of the structures of atoms and of nuclei; second, electron diffraction methods for exploring molecular structure; and last, the electron microscope for filling the gap between the limit of light microscopy and methods of exploring molecular structure.

Electron microscopy is the logical development of the known principles of electron optics, which have been discussed in many papers.[1] It was early realized that by the use of electron optics, a compound microscope could be built in a way very similar to that employed in the construction of a light microscope. Figure 1 illustrates the general principles applied in the construction of an electron microscope of the transmission type, to which the discussion in this paper will be

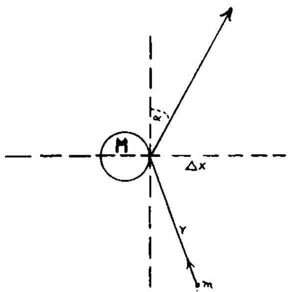


Fig. 2—Resolving Power Calculation

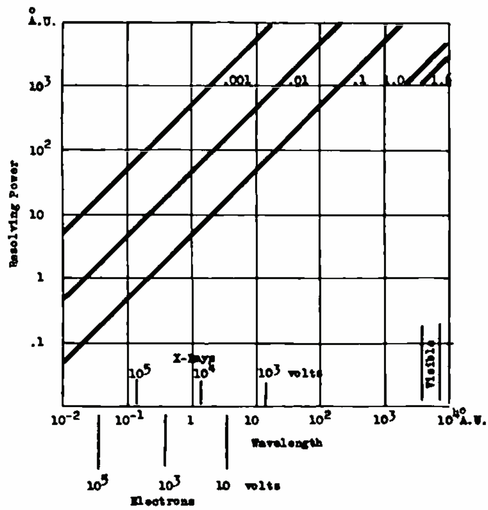


Fig. 3—Relation between Resolving Power and Wavelength

restricted, although it is obvious that this is only one among several types of microscopes.

The main reason for building electron microscopes is their increased resolving power as compared with light microscopes. It was assumed quite early that by substituting the de Broglie wavelength into the Abbe formula for resolving power, the theoretical limits for such a device would be obtained. It was later proved by Henneberg,[2] by optical considerations, that such an assumption was right. Henriot, in a private communication, came to the same conclusion by starting from the indetermination principle. Henriot's deduction is repeated here because of its beautiful simplicty (Figure 2):

If an electron of mass m and velocity v, moving toward the axis of

[1] For example: O. Klemperer, "Electron Optics", Cambridge University Press, 1939; V. K. Zworykin and G. A. Morton, "Television", John Wiley & Sons, Inc., 1940, pp. 69-127.

[2] W. Henneberg, Zeits. f. Instrumentenkunde, 55, 300, 1935.

the microscope at an angle α, strikes an object (e.g., a molecule) of mass M, the transversal indetermination (resolving power) is $\triangle x$, the indetermination of momentum is $\triangle p_x$, and h is Planck's constant. Their relationship is given by the Heisenberg formula:

$$\triangle x \triangle p_x = h$$

The momentum $\triangle p_x$ transmitted to M by the electronic encounter is $2mv \sin \alpha$.

Thus

$$\triangle x \triangle p_x = 2mv \sin \alpha \triangle x = h$$

$$\triangle x = \frac{h}{2mv \sin \alpha}$$

On the other hand, the wavelength is:

$$\lambda = \frac{h}{mv},$$

and therefore

$$\triangle x = \frac{\lambda}{2 \sin \alpha} \quad \text{as for light.}$$

The limits of resolving power calculated this way can be represented in functions of the wavelength and for different numerical apertures (Figure 3). If the calculations are right, we should be able to form images of atoms. At present, however, this is not possible; not because the calculations are incorrect, but because some other limitations affect the results. It would take too long to discuss all the limitations here, but they can be briefly enumerated.

A very important limitation is spherical aberration, which limits the useful apertures of the electron microscope to about $1/1000^{\text{th}}$ of those used in light microscopy.

Considerable attention must be given to the chromatic aberration. The focal length of a magnetic lens is directly proportional to v^2, or, what is equivalent to V, the velocity expressed in electron-volts. This statement means that practically "monochromatic" beams are needed, and that no chromatic correction is possible. The smallest resolvable distance depends, therefore, on the variation of beam speed. The requirements in this connection are extremely rigorous. The beam velocity and field strength of the lenses must be maintained to an extremely high degree of accuracy. The technical requirements for the necessary power supplies for these purposes will be discussed in a paper by A. W. Vance.[3]

[3] A. W. Vance, "Stable Power Supplies for Electron Microscope", presented at I.R.E. Meetings, June, 1940.

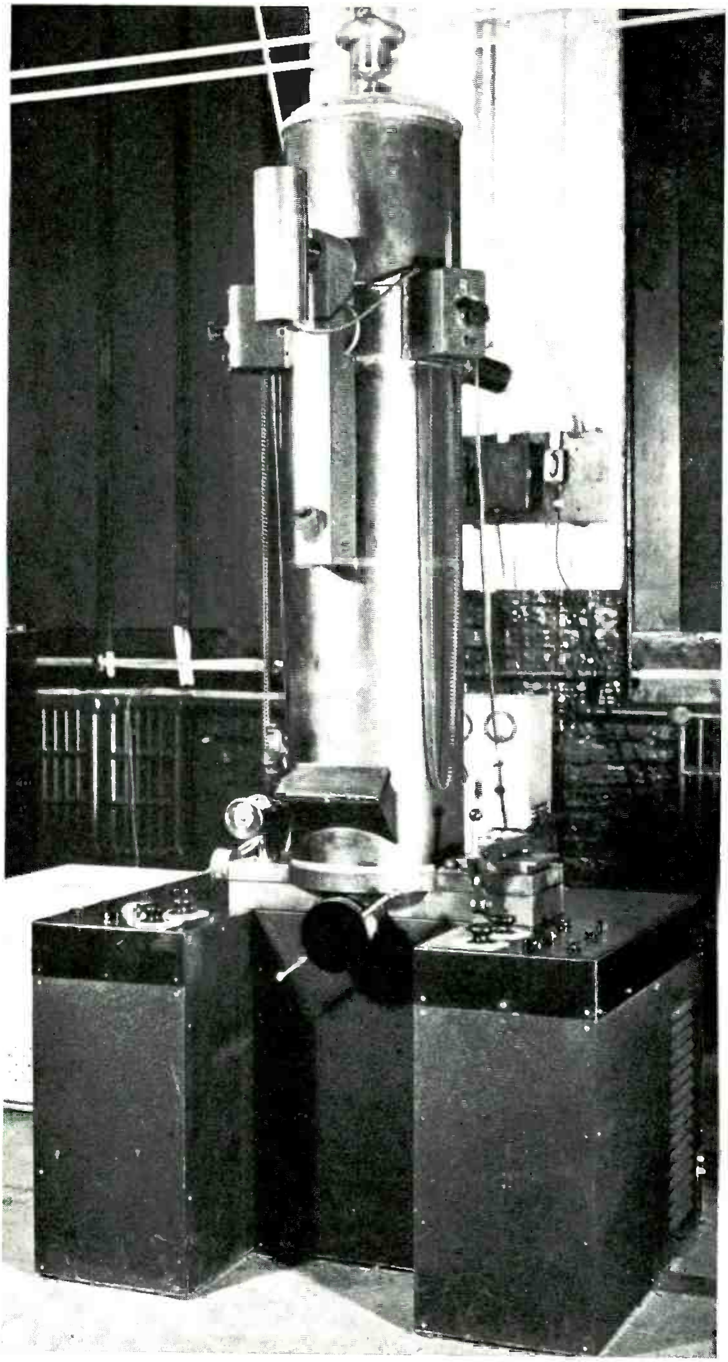


Fig. 4—Photograph of RCA Electron Microscope

Another limitation is the result of the special mechanism of image formation, which is due to electron scattering,[4] instead of absorption or refraction as in the case of light. Varying external stray fields and vibrations of the instrument may blur the image; the more these disturbing factors can be reduced, the better will be the images which can be obtained. All these limitations affect the design principles which underlie the construction of an electron microscope.

Figure 4 shows a photograph of the microscope described in the present paper. It was built in the Research Laboratories of the RCA Manufacturing Company, and incorporates a number of new features, as well as others previously described by one of the authors.[5] The chief objectives in its design were the creation of an instrument for the practical microscopist and avoidance of all external disturbances to a very high degree. Improved designs of air locks for the introduction into and withdrawal from vacuum of the specimens and photograph plates, provide easy and rapid operation of the microscope.

DESCRIPTION OF MICROSCOPE

Electron Source

The electron gun, mainly inspired by the constructional principles of X-ray tube design, creates the electron beam necessary for image formation. The coaxial filament carrier is easily removable for changing the filament, and is made vacuum tight by means of a rubber gasket. A flexible metal bellows is provided for the adjustment of the filament in both horizontal and vertical directions. Additional adjustments are provided for moving the complete gun relative to the optical axis of the microscope. The entire gun can be easily dismantled and cleaned if necessary.

Microscope Body

The optical system of the microscope, i.e., the three coils forming the "magnetic lenses" of the microscope, is completely enclosed in a brass tube. This brass tube carries the object chamber, somewhat above the center, and is lined with a magnetic shield. At its lower end the tube has two large windows for binocular observation of the fluorescent screen on which is produced the highly magnified image. Three rods inside the brass tube connect the different parts of the instrument, and hold them so rigidly that the image quality from a material standpoint remains entirely unaffected by external vibration.

A fluorescent screen, provided with a small aperture, is inserted between the objective and the projector coil. The magnified image in the first stage can be viewed on this fluorescent screen by means of a

[4] L. Marton, Physica, 3, 959, 1936.

[5] L. Marton, Bull. Acad. Belg., Bruxelles, 20, 439, (1934); 21, 553, (1935); 21, 606, (1935); 22, 1336, (1936); L. Marton, "A New Electron Microscope", Phys. Rev., 58, No. 1, 57 (1940).

periscope-like optical system, built into the main tube of the microscope. An aluminized surface mirror projects the image downwards in

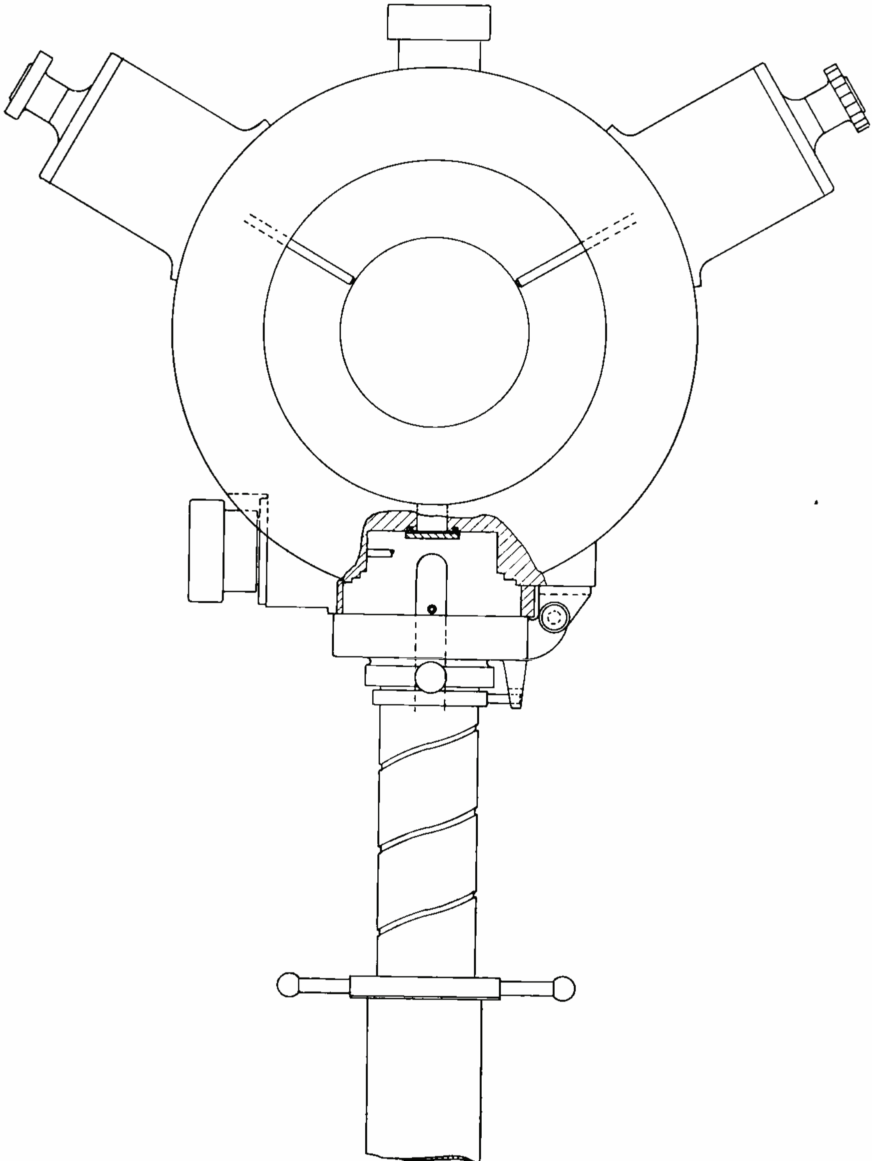


Fig. 5—Diagram of Object Chamber

the direction of an achromatic lens, and a second aluminized mirror is provided near one of the windows through which the image can be viewed by means of an eyepiece. With this arrangement, not only is is unnecessary to cut a hole in the magnetic shield, but the observer is also enabled to view both the intermediate and highly magnified images from the same position.

At about the same height at which the windows are placed on the outside of the main tube, two sprockets provided with handles are connected by means of chains to the "stage" adjustment of the object chamber. This arrangement, together with the enclosed, built-in control panel for the electrical adjustments, makes it possible for the observer to manipulate each important adjustment of the instrument while seated in the normal observation position.

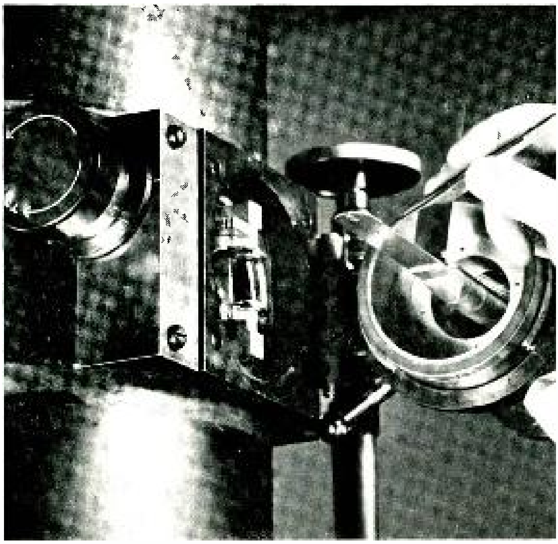


Fig. 6—Photograph of Object Chamber

Electron Optical System

Referring again to Figure 1, the electron optical system, consisting of a condenser coil, an objective coil, and a projection coil, is entirely within the main tube. Following previous designs,[5] the coils are completely ironclad with pole pieces screwed into the inner threaded part of the iron enclosures. The coils themselves are enclosed in sealed copper shields with glass-metal seals for the leads. This latter provision is necessary in order to avoid excessive pumping time due to gas between the coil windings. With the present set of pole pieces, the magnification of the objective is of the order of 100 diameters and that of the projection coils varies between 20 and 200 diameters, giving, therefore, a maximum total of magnification of 20,000, at an electron speed of 80 ekv. For lower speeds (around 30 ekv) the magnification can reach about 25,000.

Object Chamber

The air lock for introducing and withdrawing the specimens, and the mechanical "stage" are the main parts of the object chamber. Introduction and withdrawal are carried out by means of a lateral compartment which can be separated from the main body of the microscope by means of a small gate operated from the outside. When the specimen is withdrawn into the lateral compartment (Figure 5) and the gate is closed, the whole tubular part can be filled with air and swung back on its hinges, giving access to the specimen, which is clamped

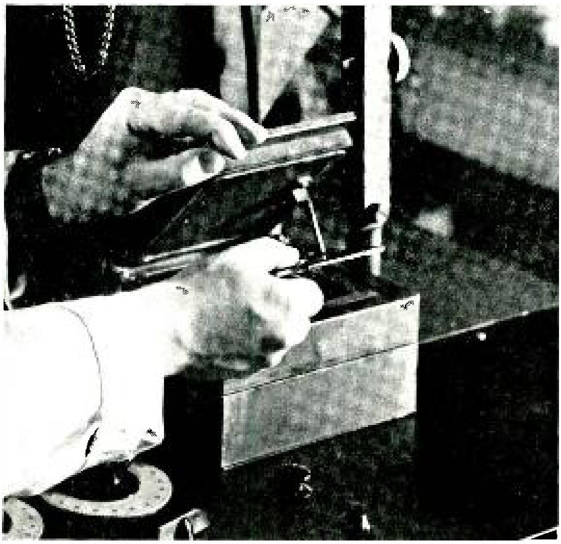


Fig. 7—Photograph of Photographic Chamber

between two "blades" (Figure 6). When the specimen is replaced, the compartment can be closed by means of a rubber gasket and pumped out separately, after which the gate is removed and the specimen is introduced into the microscope. In the microscope the blades are inserted into a ring shaped holder which can be moved by means of a micrometer adjustment operated through flexible bellows. Outside scales are provided for reading the movement of the specimen with an accuracy of one micron. An additional, externally operated gear arrangement is provided for raising and lowering the "stage".

Photographic Chamber

The photographic chamber (Figure 7) is also built on the air lock principle. A gate separates the photographic chamber from the main body of the microscope. This gate, which closes the opening at the bottom part of the main body by means of a gasket, is mounted on a

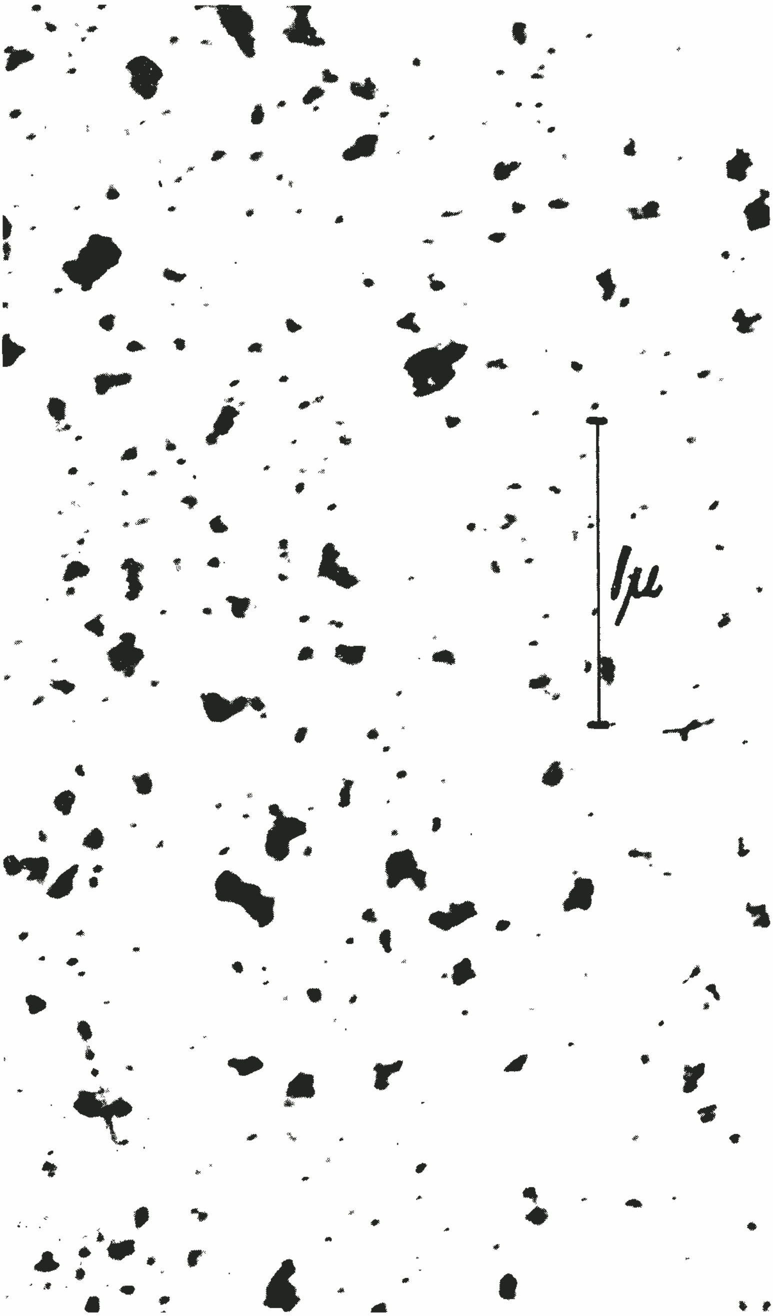


Fig. 8—Colloidal Gold

carriage, and can be lowered and moved sideways from the aperture by means of sloping guides. The same carriage holds the photographic plate holder. An externally operated gear is provided for moving the carriage back and forth, and for placing either the plate holder or the gate in front of the opening. Stops are provided for removing the cover of the plate holder for exposure of the plate and for replacing the cover after exposure. The top of the gate is coated with fluorescent material for visual observation of the highly magnified image. Another fluorescent screen is provided, hinged above the gate opening, which also can be operated externally by means of a flexible bellows. Series

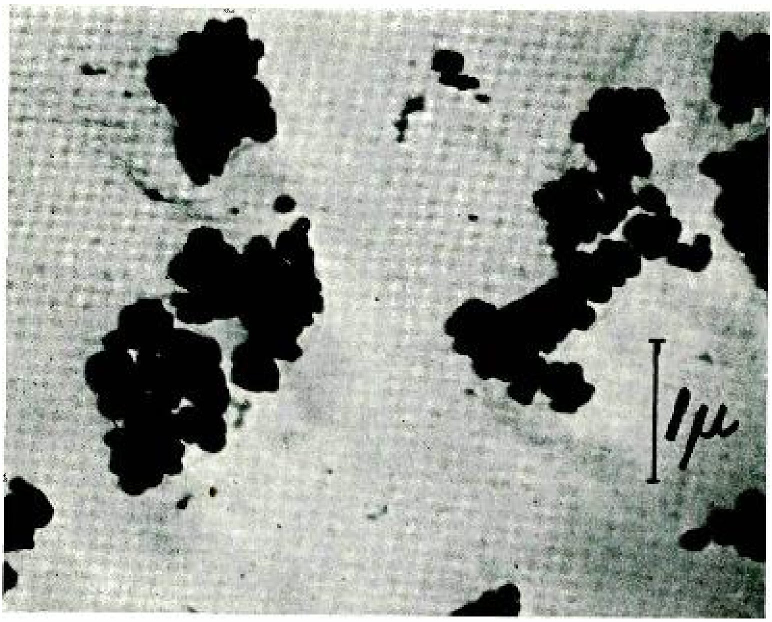


Fig. 9—Titanium Dioxide

exposures can be taken on one photographic plate by partially covering the field with the hinged fluorescent screen and moving the carriage by fractions of its motion. An auxiliary pumping system permits operation of the air lock in a manner similar to that described for the object chamber.

Pumping

The main pumping system for the microscope consists of a mechanical pump and a three-stage, oil diffusion pump. The latter is suspended from the bottom of the microscope and can be shut off by means of a vacuum valve having a large cross section. Thermocouple and ionization gauges are provided for controlling the vacuum, which is generally maintained at 10^{-5} millimeters of mercury. The auxiliary pumping system consists of a mechanical pump connected to both the object and photographic chambers by means of special bellows-operated two-way valves.

Results

The main applications of the transmission type electron microscope are the investigation of minute particles beyond the limit of visibility of the light microscope, and of details of larger particles, which, them-

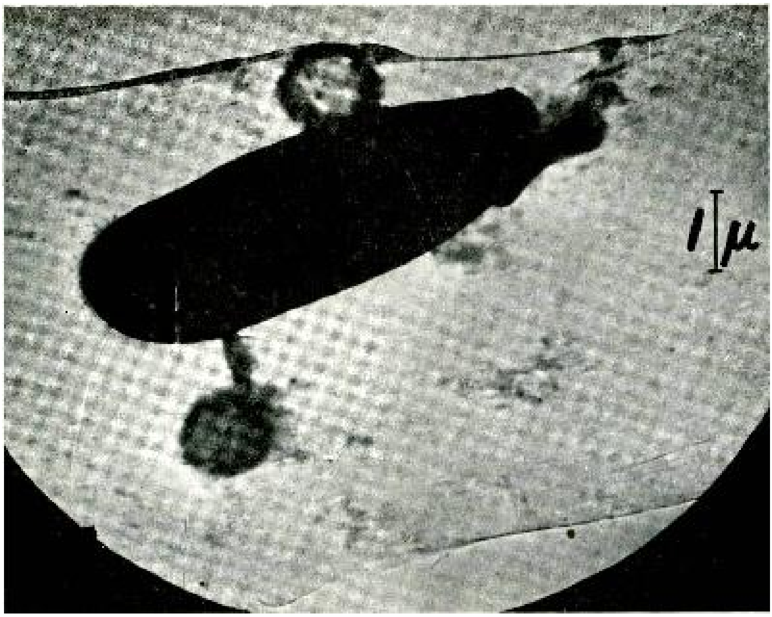


Fig. 10—Spores of Trychophyton Mentagrophytes (Athletes Foot)

selves, may be visible.[6] In the first group belongs all the colloids (Figure 8), industrial and biological; virus particles; fine fibers; pigments; etc. (Figure 9). In the second group we can mention the large fields of bacteriological research (Figure 10) with the possibility of exploring the internal structures of bacteria (Figure 11) and structures of animal and vegetable cells.

The present resolving power of the instrument can best be estimated from Figure 12, which shows a hole in a nitrocellulose film which developed a crack, the edges of which are at a distance of 100 Angstrom Units. Both edges are sufficiently sharp to estimate from them a resolving power of about 50 Angstrom Units.

[6] L. Marton, Proc. of the International Union Against Cancer, Paris, 1938.

ACKNOWLEDGMENTS

The authors wish to express their gratitude to Dr. V. K. Zworykin, Associate Director of the Research Laboratories of the RCA Manu-

Fig. 11—Whooping Cough Bacteria

facturing Company, for his constant interest in this work; to Dr. E. G. Ramberg for his help on the optical parts; to Mr. V. Goltzoff, who

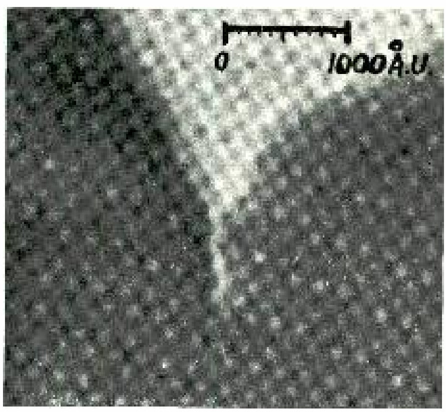


Fig 12—Crack in Nitrocellulose Film

helped with tests on the instrument; and to Dr. Stuart Mudd, and his staff, of the Department of Bacteriology, University of Pennsylvania, who supplied specimens.

FLUCTUATIONS IN SPACE-CHARGE-LIMITED CURRENTS AT MODERATELY HIGH FREQUENCIES

By

B. J. THOMPSON, D. O. NORTH AND W. A. HARRIS

RCA Manufacturing Company, Inc., Harrison, N. J.

Continued from July, 1940, RCA REVIEW

PART III—MULTI-COLLECTORS

BY DWIGHT O. NORTH

Summary—*The fluctuations in cathode current (I_a) due to true shot effect in cathode emission are*

$$\overline{i_k^2} = \Gamma^2 \; (I_a \cdot 2e\triangle f).$$

Γ *is a factor, evaluated in Part II, equal to unity when the current is temperature-limited, but ordinarily only about 0.2 under normal space-charge-limited conditions.*

Theory here presented shows that fluctuations in the current (I_n) to the n^{th} collector of a device possessing several collectors for the cathode current are, in general, improperly represented by the expression,

$$\overline{i_n^2} = \frac{I_n}{I_a} \overline{i_k^2},$$

but correctly by the following formula:

$$\overline{i_n^2} = \left[1 - \frac{I_n}{I_a} (1 - \Gamma^2) \right] (I_n \cdot 2e\triangle f).$$

In connection with suppressor-grid pentodes, for which it is generally true that $\Gamma^2 << I_{c2}/I_b$, the formula for shot effect in either screen or plate lead (with input grounded) is

$$\overline{i_{g2}^2} \approx \overline{i_p^2} \approx \frac{I_{c2} I_b}{I_a} \cdot 2e\triangle f.$$

The noise in the plate lead is, then, usually higher than that in the cathode lead, and this fact accounts in large measure for prevalent observations that pentodes are usually inferior to triodes from purely a noise standpoint.

Experimental work is described, supporting the theory.

Combination of this theory with that of Part II yields a practical working formula for the apparent input shot effect of a conventional pentode amplifier with coated cathode and negative control grid. Expressed in terms of an ohmic resistance (R_{eff}) at room temperature (300°K), the thermal agitation of which is equal to the apparent input shot effect, the formula is

$$R_{eff} \text{ (pentode)} = \left[1 + 8.7\, \sigma \frac{I_{c2}}{g_m} \cdot \frac{1000}{T} \right] \cdot R_{eff} \text{ (triode)}$$

where $$R_{eff}\ (\text{triode}) = \frac{2.2}{\sigma} \cdot \frac{T(°K)}{1000} \cdot \frac{I_b}{I_a} \cdot \frac{1}{g_m}.$$

T is cathode tempature in °K, g_m is pentode transconductance in micromhos, I_{c2} is the screen current in microamperes, and σ is the ratio of total transconductance to conductance of the "equivalent diode."

THEORY

THE phenomena peculiar to shot-effect fluctuations in a space-charge-limited device which possesses more than one collector have been outlined in Part I. If a pentode, for example, is space-charge-limited, the shot effect in the anode lead is often much in excess of an estimate based upon a simple division of cathode current fluctuations similar to the division of the cathode current itself. In *tetrodes* such a behavior might *ab initio* be ascribed, at least in part, to additional noise associated with secondary-emission current circulating among the collecting electrodes. The analysis and experiment presented below are confined to arrangements in which secondary emision plays no role. In fact we shall be concerned with only an extension of the concepts and formulas of Parts I and II.

Any electron in transit alters the potential of the virtual cathode as explained earlier. This slight depression of the potential minimum is by no means localized but extends over a region the linear dimensions of which are, loosely speaking, comparable to the distance between cathode and anode of the "effective diode". It is beyond the purposes of this paper to examine the precise geometrical scope of the potential variation at the virtual cathode, but it should be evident that the concomitant variation in the steady-state cathode current I possesses a comparable geometrical spread. In this way it comes about that, at times, the compensatory flow of current which accompanies and tends to reduce an initial fluctuation in emission (yielding the Γ of Part II) does not all arrive at the same collecting electrode to which the initial fluctuation may be consigned, but is divided among the several collectors. Consequently, the fluctuation in the electrode named is not fully reduced, while in the other electrodes there flows simultaneously a "compensating" current for which there is no antecedent. The mean-square fluctuations in each electrode are, therefore, in excess of an estimate based upon a simple apportioning of the cathode current fluctuations. At any instant the sum of the linear fluctuations in all collector currents is still, naturally, equal to the fluctuation in cathode current. But the sum of the mean-square current fluctuations in the former is *greater than* in the latter.

Let us consider in detail two extreme examples of cathode-current division. The behavior of these will set natural limits encompassing the performance of all others. In Figure 1a the collectors are to be considered so close to the cathode (in comparison with their linear dimensions) that there is virtually no region of the cathode from which current is delivered to both collectors. Except in the immediate vicinity of the boundary common to the collectors this will be strictly true. Hence, by extending the cathode and collectors laterally the ideal can be approached as closely as desired. These collectors now operate independently and their performance will be altered in no way whatsoever if the cathode be split also and the two diode systems sep-

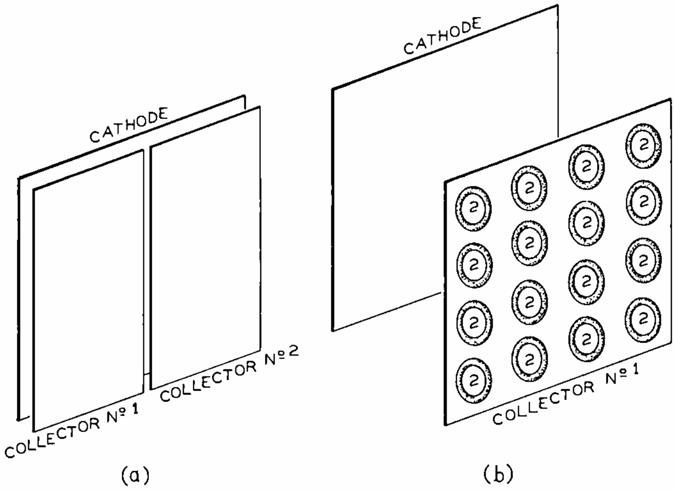


Fig. 1—Extreme examples of a multi-collector.

arated. Employing the notation of Part II, we may therefore write the shot effect in each collector as follows:

$$\left.\begin{aligned} \overline{i_1^2} &= \Gamma_1^2 \cdot 2eI_1\Delta f = \theta_1 \cdot 4kTg_1\Delta f \\ \overline{i_2^2} &= \Gamma_2^2 \cdot 2eI_2\Delta f = \theta_2 \cdot 4kTg_2\Delta f\,. \end{aligned}\right\} \quad (1)$$

In short, there is nothing novel here; each collector current shows a shot effect prescribed by the analysis of Part II, and the sum of the mean-square fluctuations in the two anode currents, measured separately, is equal to the mean-square fluctuation in the anode current of the device when the two collectors are tied together.

Proceeding to the opposite extreme, we find in Figure 1b a scheme in which all the litle discs (2) constitute one collector, the other consisting of the remanent sheet. If now the number of discs is thought to become exceedingly large while, at the same time, the area of each is proportionately diminished so that the currents I_1 and I_2 are unaltered, another ideal situation is approached. It has the following significant properties:

1. The current to either collector is composed of electrons which, at the virtual cathode, exhibit a *precise* Maxwell-Boltzmann distribution of velocities.

2. The depression of virtual-cathode potential which accompanies the passage of any electron (from whatever point on the cathode it emerges) gives rise to a "compensating" current which is *always* divided between the collectors in proportion to their steady-state currents I_1 and I_2.

The scheme can be generalized to N collectors, and this we shall analyze. The cathode current will be

$$I = \sum_{1}^{N} I_n . \tag{2}$$

Assuming that all electrodes are grounded with respect to noise frequencies, let us determine the net current fluctuations in the n^{th} collector. It is composed of three parts: A, fluctuations associated with true shot effect in those β-electrons[1] which are collected at the n^{th} electrode; B, fluctuations associated with true shot effect in those β-electrons which are collected at all other electrodes; C, fluctuations associated with true shot effect in the α-electrons[1].

A. Were it not for compensating action at the virtual cathode, the shot effect in I_n would be true shot effect originating wholly in random fluctuations in the emission of those β-electrons which reach I_n, namely,

$$\overline{i_n^2} = I_n \cdot 2e\Delta f . \tag{4}$$

In the presence of a virtual cathode, it was shown in Part II that for each emission velocity v_s, there is a compensating action such that each fluctuation which contributes to (4) is *linearly* reduced by a factor γ_β (itself a function of v_s). However, in contrast with the simple diode, it cannot now be said that an initial fluctuation of unit amplitude (1) gives rise to a "compensating" current *in the n^{th} electrode* of amplitude—$(1-\gamma_\beta)$. For the compensating flow is divided in proportion to the steady-state current distribution. Rather, each

[1] See Part II, pp. 444 and 451. The α-electrons are those emitted with velocity insufficient to carry them over the virtual cathode. The remainder, the β-electrons, constitute the normal cathode current which is divided and collected at the N electrodes.

initial fluctuation which contributes to (4) is here linearly reduced by the factor $\left[1 - \frac{I_n}{I}(1-\gamma_\beta)\right]$. The net fluctuations in I_n, due to true shot fluctuations in I_n, and averaged over the appropriate range of emission velocities,* are, therefore,

$$\overline{i_n^2} = \left[1 - 2\frac{I_n}{I}(1-\overline{\gamma_\beta}) + \left(\frac{I_n}{I}\right)^2 (1 - 2\overline{\gamma_\beta} + \overline{\gamma_\beta^2})\right] I_n \cdot 2e\Delta f. \quad \text{(4a)}$$

B. A similar treatment starts with initial true shot fluctuations in the current to all collectors but the n^{th}, namely,

$$(I - I_n) \cdot 2e\Delta f. \quad (5)$$

Again, for each fluctuation which contributes to (5), there is a compensating flow which, for unit amplitude of the initial fluctuation, produces a current in the n^{th} electrode of amount $-\frac{I_n}{I}(1-\gamma_\beta)$. Averaged over the same range of emission velocities, the contribution from this source to fluctuations in the n^{th} electrode is

$$\overline{i_n^2} = \left[\left(\frac{I_n}{I}\right)^2 (1 - 2\overline{\gamma_\beta} + \overline{\gamma_\beta^2})\right](I - I_n) \cdot 2e\Delta f. \quad \text{(5a)}$$

C. The initial true shot effect in the α-electron current is

$$(I_s - I) \cdot 2e\Delta f, \quad (6)$$

where I_s is the total emission, so that $(I_s - I)$ is simply that portion of the emission current which fails to cross the virtual cathode. And once more, for each fluctuation which contributes to (6) there is a disturbance of the virtual cathode potential and a consequent flow of current to the collectors, each collector receiving its share in proportion to the steady-state current it carries. Averaged over the appropriate range of emission velocities, the contribution from this source to fluctuations in the n^{th} electrode is

* The averaging process will not be elaborated here. It is explained in detail in Part II, p. 450, from which it will be recognized that the $\overline{\gamma_\beta^2}$ of (4a) is simply Γ_β^2, and that the $\overline{\gamma_\alpha^2}$ of (6a) is similarly Γ_α^2.

$$\overline{i_n^2} = \left[\left(\frac{I_n}{I} \right)^2 \overline{\gamma_\alpha^2} \right] I \cdot 2e\Delta f . \tag{6a}$$

Provided there is no secondary emission, nor elastic reflection of the sort described in Part II, the total shot effect in the n^{th} collector is the sum of these three contributions, (4a), (5a), (6a). Remembering that

$$\Gamma^2 = \Gamma_\alpha^2 + \Gamma_\beta^2 ,$$

we find for the total fluctuation:

$$\overline{i_n^2} = \left[1 - \frac{I_n}{I} (1 - \Gamma^2) \right] I_n \cdot 2e\Delta f . \tag{7}$$

The fluctuation in the cathode current is

$$\overline{i^2} = \Gamma^2 I \cdot 2e\Delta f . \tag{8}$$

Part II was devoted to a determination of Γ^2. At this point one needs only to recall that it runs from unity in the absence of a virtual cathode to the neighborhood of 1/20 for normal space-charge-limited operation of modern medium-gain receiving tubes.

The sum of the separately measured shot effects in all N collectors is

$$\sum_1^N \overline{i_n^2} = \left[\Gamma^2 I + \frac{(1 - \Gamma^2)}{I} \sum_1^N I_n (I - I_n) \right] 2e\Delta f ,$$

substantiating the remark above that the sum of the mean-square current fluctuations in the collector leads is greater than that in the cathode lead—unless either $\Gamma = 1$ (no virtual cathode, no compensating action), or one collector takes the entire cathode current (no problem).

Inspection of (7) leads to the following observations:

1. No fluctuation is greater than the true shot effect for the current considered.

2. The smaller fraction of the total current an electrode collects the more nearly the noise in that current approaches true shot effect.

3. For vanishingly small Γ, the mean-square fluctuation in the current collected at any electrode is equal to the product of the true shot effect for said current and the fraction of the cathode current *not* collected at said electrode.

4. The ratio of actual noise to true shot effect in a divided portion of the cathode current *exceeds* the corresponding ratio for the total cathode current.

5. The noise in a divided portion I_n of the cathode current *exceeds* the noise in the total cathode current provided

$$\frac{\Gamma^2}{1-\Gamma^2} < \frac{I_n}{I} < 1 \, .$$

In conventional tubes this is usually true for all collectors.

6. With constant Γ, the noise in a given collector current I_n is a maximum (against variations in I_n) when

$$\frac{I_n}{I} = \frac{1}{2(1-\Gamma^2)} \, .$$

In other words, provided $\Gamma^2 < 1/2$, the noise in *no* collector lead should exceed

$$\frac{1}{4(1-\Gamma^2)} I \cdot 2e\triangle f \, .$$

Application to Pentodes

The conventional suppressor-grid pentode is expressly designed to inhibit exchange of secondaries between screen grid and plate. Its performance can, therefore, be expected to lie within the extremes just discussed, provided there is no appreciable secondary emission from insulating surfaces, e.g., the envelope. The suppressor grid will be supposed to collect no current. It will also be supposed that the control grid is given a negative bias; from the viewpoint of this section a positive control grid would, of course, be regarded as just another collector, but it will generally introduce further complexities because of possible secondary emission to other electrodes, and, more important yet, the likelihood of elastic reflection resulting in an indeterminate increase in Γ of the sort described at length in Part II.

Tubes designed with no intent to align grids so as to produce electron beams, or, more specifically, to reduce the screen-grid current, would be expected *a priori* to perform very nearly like the idealized scheme of Figure 1b. For, scattering of electrons under the influence of the control-grid field, together with the lack of alignment, are sufficient reasons for supposing that the two important properties (p. 247) of the idealized structure are closely approximated. Even some tubes of beam-forming design can be placed in the same category, but, in general, each device of this class should be accorded an individual examination.

Since subsequent experiments support these conjectures, it will be useful to cast (7) into pentode nomenclature. With the cathode current divided as follows,

$$I_a = I_{c2} + I_b \,, \tag{9}$$

and with all electrodes grounded with respect to noise frequencies, the shot effect in the cathode, the screen-grid, and the plate leads,[1] respectively, is

$$\left.\begin{aligned} \overline{i_k^2} &= \Gamma^2 (I_a \cdot 2e\Delta f) \\ \overline{i_{g2}^2} &= \frac{\Gamma^2 I_{c2} + I_b}{I_a} (I_{c2} \cdot 2e\Delta f) \\ \overline{i_p^2} &= \frac{\Gamma^2 I_b + I_{c2}}{I_a} (I_b \cdot 2e\Delta f) \,. \end{aligned}\right\} \tag{10}$$

It should be noted carefully that these formulas will serve directly to predict the noise voltage appearing across an impedance in any lead, but only provided no noise voltages appear on other electrodes. For example, if the screen is grounded for radio frequencies, the noise voltage appearing across an impedance Z in the plate lead is simply

$$\overline{e_p^2} = \overline{i_p^2} \left| \frac{r_p Z}{r_p + Z} \right|^2 , \tag{11}$$

where r_p is the plate resistance of the tube. If, however, there is impedance in both screen and plate circuits, the screen-plate transconductance (and, to a lesser extent the plate-screen transconductance), together with the fact that the current fluctuations in screen and plate are partly in phase, partly random, necessitates a return to fundamentals and construction of entirely new expressions to take the place of (10) and (11). These are not difficult to obtain but lie beyond the aim of this paper which refers specifically to normal pentode amplifiers only. To these the expressions above are usually applicable.

[1] Different methods of attack have led to the same formula, e.g., W. Schottky, "On the Theory of Electron Noise in Multiple-Grid Tubes," *Ann. d. Physik*, Vol. 32, p. 195, May, (1938). C. J. Bakker, "Current Distribution Fluctuations in Multi-Electrode Radio Valves," *Physica*, Vol. 5, No. 7, p. 581, July, (1938).

Now it will ordinarily be found that $\Gamma^2 << \dfrac{I_{c2}}{I_b}$. The formulas (10) then approach the following limiting forms:

$$\left.\begin{aligned} \overline{i_k^2} &= \Gamma^2 (I_a \cdot 2e\Delta f) \\ \overline{i_{g2}^2} = \overline{i_p^2} &= \frac{I_{c2} I_b}{I_a} \cdot 2e\Delta f\,. \end{aligned}\right\} \qquad (10a)$$

If it is also true that $I_{c2} << I_b$, *the noise in both screen and plate circuits is approximately equal to the true shot effect for a current equal to the screen current* I_{c2}. The noise in the plate circuit may then exceed that in the cathode lead by an order of magnitude. This undesirable behavior is illustrated by the measurements below, and is in accord with the prevalent impression that conventional pentodes are usually inferior to triodes from purely a noise standpoint.

The most valuable figure of merit is the effective input noise, namely, that voltage applied between grid and cathode which produces output current fluctuations equal in magnitude to those actually generated by shot effect. The effective input noise is expressed by

$$\overline{e_i^2} = \frac{\overline{i_p^2}}{g_m^2},$$

where g_m is used here to denote control grid-to-plate transconductance. The symbol g_t will be used to denote total (or cathode) transconductance. In brief,

$$g_m \equiv \frac{\partial I_b}{\partial E_{c1}}, \quad g_t \equiv \frac{\partial I_a}{\partial E_{c1}},$$

and, since it is normally true, we shall *assume*

$$g_t = \frac{I_a}{I_b} g_m \,.$$

Using (10), we then have

$$\overline{e_i^2} = \left[1 + \frac{I_{c2}}{\Gamma^2 I_b} \right] \frac{\Gamma^2 I_a \cdot 2e\Delta f}{g_t^2}\,. \qquad (11)$$

The term without brackets will be recognized as the equivalent input noise of the tube operated as a triode (both screen and plate working into the same radio-frequency circuit). The term within brackets is, therefore, the factor by which the effective input noise is increased when the tube is employed as a pentode.

For engineering purposes a more valuable formula expresses the input noise simply as the effective resistance R_{eff} which, at room temperature T_o, exhibits a thermal-agitation voltage equal to the effective input shot voltage of the tube. This development parallels that of p. 471, Part II, to which reference may be made for details. We also need (43b) of Part II, namely,

$$\Gamma^2 = 2\frac{\theta}{\sigma}\,\frac{g_t V_e}{I_a}. \tag{12}$$

In this expression θ is a pure number, practically equal to 2/3 over the whole range of space-charge-limited operation, provided the ratio of cathode current to emission is small; σ is the ratio of total transconductance g_t to the conductance g of the "equivalent diode", and is given more explicit definition in (51) of Part II; $V_e = \frac{kT}{e}$, where T is cathode temperature, k is Boltzmann's constant, e is electron charge. Assuming that $\theta = 2/3$ and $T_o = 300°K$, we find

$$\left.\begin{array}{l} R_{eff}(\text{pentode}) = \left[1 + 8.7\sigma\dfrac{I_{c2}}{g_m}\cdot\dfrac{1000}{T}\right]\cdot R_{eff}\ (\text{triode}) \\ \text{where } I_{c2} \text{ is in microamperes, } g_m \text{ in micromhos, and } T \text{ in } °K, \\ \text{and where} \\ R_{eff}\ (\text{triode}) = \dfrac{\theta}{\sigma}\cdot\dfrac{T}{T_o}\cdot\dfrac{1}{g_t} = \dfrac{2.2}{\sigma}\cdot\dfrac{T}{1000}\cdot\dfrac{I_b}{I_a}\cdot\dfrac{1}{g_m}. \end{array}\right\} \tag{13}$$

For tubes with oxide-coated cathodes, it has been shown in Part II that $\theta = 2/3$ for practical purposes, so that (13) may be considered valid. However, whenever there is any question as to the validity of (12), (11) should be used in place of (13), and Γ^2 should be found experimentally from a measurement of noise in the cathode lead.

As an example of the magnitude of pentode shot noise as predicted by (13), consider a typical radio-frequency pentode for which

$$T = 1000\,°K$$

$$I_k = \frac{4}{5} I_a = 2 \text{ milliamperes}$$

$$g_m = 1200 \text{ micromhos}$$

$$\sigma = 0.83$$

Then R_{eff} (triode) = 1770 ohms

R_{eff} (pentode) = 4.0 R_{eff} (triode) = 7100 ohms.

In concluding the theory it is worth observing that, although (10) and (11) are developed with the understanding that $0 < \Gamma < 1$, and that Γ is simply a manifestation of fluctuations in virtual-cathode potential elicited by and tending to compensate true shot fluctuations in electron emission, they are equally valid in at least two further instances in which Γ *may* be greater than unity. Whenever positive ions are evaporated from the cathode (discussed in Part II) or whenever they are formed within the tube by collisions between electrons and molecules of residual gas (the subject of Part IV), if the ion currents are small enough their sole contribution to the current fluctuations in the plate circuit is due simply to the fact that the ions alter the potential of the virtual cathode, so that fluctuations in ion current bring about fluctuations in cathode current.[1] The mechanics of this process is strictly parallel to that discussed at length in Part II. Fluctuations observed in the cathode lead may even here be expressed empirically in the usual fashion:

$$\overline{i_k^2} = \Gamma^2 I_a \cdot 2e\Delta f\,.$$

But now the former analytical expression,

$$\Gamma^2 = \Gamma_\alpha^2 + \Gamma_\beta^2\,,$$

must be extended to include, phenomenologically, the noise produced by positive ions, so that we have

$$\Gamma^2 = \Gamma_\alpha^2 + \Gamma_\beta^2 + \Gamma_{pos.}^2 \tag{14}$$

Exceedingly small ion currents can make $\Gamma > 1$. Nevertheless, when

[1] Assuming, as before, that the control grid is grounded with respect to noise frequencies. If not, gas current to the grid introduces additional noise, cf. Part IV.

Γ is defined as stated, the noise in screen and plate leads is still properly described by (10). An observed example of such a situation will be presented in the following section.

Experiment

All measurements were conducted at a frequency of about one megacycle. The apparatus and techniques are fully detailed under this heading in Part II, and the signal substitution method described therein was employed throughout.

Since the chief purpose of this work was a direct test of multicollector theory as enunciated in (10), there was no attempt to make use of the theoretical value of Γ. In brief the procedure was as follows. A two-collector tube was operated at any chosen point on its characteristic. The fluctuations in cathode current were measured to permit an experimental determination of Γ. At the same operating point the fluctuations in plate and screen current were successively measured. These two values were then compared with their corresponding theoretical estimates, using the experimental Γ in the calculation.

The control grid was biased to a negative potential in every instance, and the suppressor, if any, was connected directly to the cathode. Measurement always showed no significant current to either of these electrodes. When plate-current fluctuations were being measured, all electrodes but the plate were grounded for noise frequencies with blocking condensers, and a known impedance was placed in the plate lead; the voltage fluctuations appearing across this impedance were amplified and measured. A similar procedure yielded screen-current fluctuations. Cathode-current fluctuations were obtained in this way also whenever screen and plate were operated at the same potential so that they could be tied together. Otherwise, known impedances with conductive branches were inserted in both screen and plate leads, and the screen and plate were tied together (as regards noise frequencies) with a blocking condenser. Screen and plate, therefore, operated as a unitary electrode at one megacycle, and the voltage fluctuations at the plate were a direct measure of fluctuations in cathode current. (A schematic circuit of this type, used for the same purpose, is given in Part II, Figure 16.)

Voltage fluctuations are plotted here, but these can be interpreted immediately in terms of current fluctuations because all voltages refer to a common impedance (45.5×10^3 ohms) and band width (9.50 kilocycles)—these being the magnitudes of circuit impedance used, and amplifier pass-band, respectively. Thus the fundamental quantity, the mean-square current fluctuations per unit band width $\left(\frac{\overline{i^2}}{\Delta f}\right)$ corresponding to any rms noise voltage $\sqrt{\overline{e^2}}$ (μv) shown in Figures 2-4 can be obtained from

$$\frac{\overline{i^2}}{\Delta f}\left(\frac{\mu\mu a^2}{\text{kc}}\right) = \frac{\overline{e^2}}{Z^2 \Delta f} = 50.9\,\overline{e^2}\,(\mu v^2).$$

The three sets of data exhibited were selected from measurements on several types of tube, and may be considered truly representative

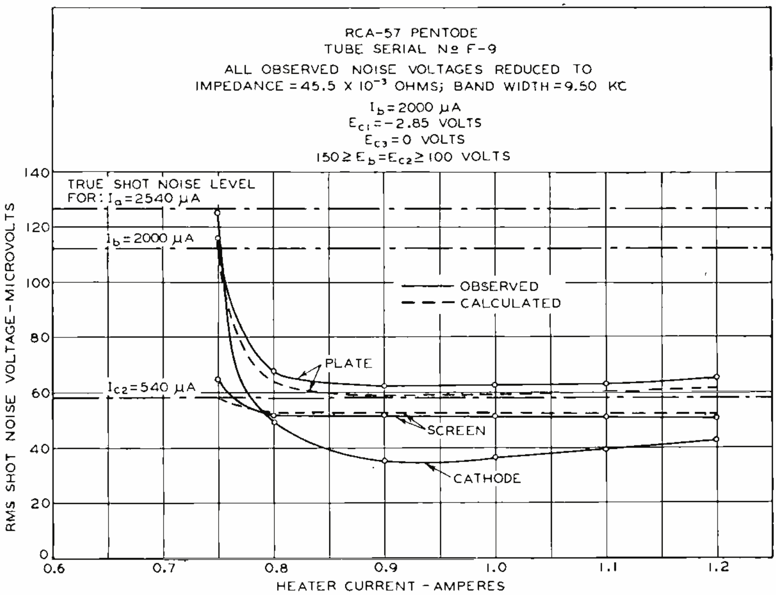


Fig. 2—Comparison between theory and observation of fluctuations in screen and plate leads of a typical radio-frequency pentode. Constant plate current, maintained by adjusting E_{c2} and E_b.

of the extent to which (10) may be employed in predicting noise behavior of similar tubes. In each figure the course of noise is traced while, with constant plate current and fixed control-grid bias, the heater temperature is altered so that the tube moves from an essentially temperature-limited condition to one in which the emission is considerably higher than the cathode current. Γ (measured), therefore, runs from unity rapidly to a minimum and then rises again with increasing temperature, in general accord with the theoretical representations of Part II. Γ is not plotted but may be evaluated in any instance by taking the ratio of cathode noise to the true shot noise level for the cathode current, plotted in each figure. This procedure implies that the cathode current was constant for each plot, which is not strictly true, but the variations are small enough that they cause no material error in this connection.

Figure 2 exhibits the behavior of an RCA-57, a typical radio-frequency pentode with coated cathode and a screen current equal to about one-fifth of the cathode current. The plate and screen were operated at the same potential which was varied from 150 to 100 volts in order to keep I_b constant as T was increased. I_{c2} ran simultaneously from 540 to 530 microamperes. The theory is seen to agree very well with the data. For minimum cathode current fluctuations,

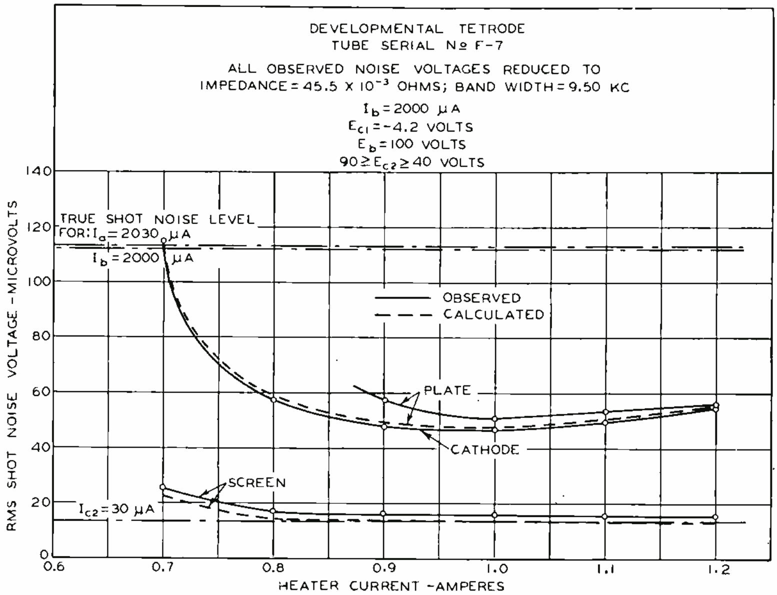


Fig. 3—Comparison between theory and observation of fluctuations in screen and plate leads of a special beam-type tetrode with aligned grids. Constant plate current maintained by adjusting E_{c2}.

$$\Gamma^2 = 0.077 \qquad \frac{I_{c2}}{I_b} = 0.27\,,$$

so that

$$\Gamma^2 << \frac{I_{c2}}{I_b}$$

and the limit (10a) is approached, both screen and plate noise approximating the true shot effect for a current equal to the screen current. Due simply to the presence of a screen grid, the rms plate noise in this tube has roughly twice the magnitude which would be observed if the screen grid were absent.

The data shown were duplicated in tests wherein the plate was held at 250 volts and the screen-grid potential assigned whatever lower value (100 — 200 volts) would set $I_b = 2000$ microamperes.

The tube concerned in Figure 3 was a special developmental tetrode made by Mr. Otto Schade of these laboratories during studies culminating in the RCA-6L6, a beam power tube.[1] The exceptionally low

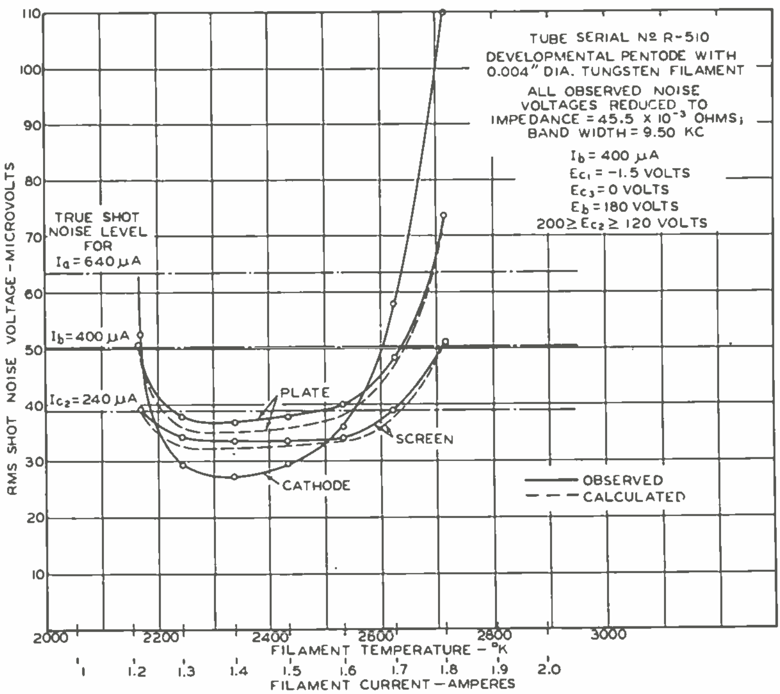


Fig. 4—Comparison between theory and observation of fluctuations in screen and plate leads of a special pentode with tungsten filament, illustrating region for which $\Gamma > 1$ as a result of positive-ion emission. Constant plate current, maintained by adjusting E_{c2}.

screen current, 1.5 per cent of the cathode current, is obtained through alignment of control-grid wires with those of the screen grid. Actually the screen-grid current was about 100 microamperes in the temperature-limited condition, and dropped to a nearly constant value of 30 microamperes for all other points. The increase in screen-grid noise above the true shot noise level may be due in part to exchange of secondaries between screen and plate. The course of plate noise is

[1] O. H. Schade, "Beam Power Tubes," *Proc. I.R.E.*, Vol. 26, No. 2, p. 137, February, (1938).

shown for only that portion of the total range in which the emission is far in excess of cathode current, i.e., the noise is independent of emission.[1] At the time plate noise was measured, the emission was increasing slowly with constant heater current; hence, measurements for nearly temperature-limited currents could not be expected to agree well with theoretical estimates based upon subsequent measurements of noise in the cathode current. In any event, only the region of "complete" space-charge limitation is practically important. And here the agreement between theory and observation is again good. In this region the screen potential is less than 50 volts, so that, in view of the extremely small screen current, noise originating in secondary emission cannot be serious.

The plate noise is practically equal to the cathode noise, simply because the screen current is negligible. At a casual inspection these data might be considered a none-too-convincing confirmation of (10), inasmuch as the plate noise would naturally approach that of the cathode if the screen current were to vanish. The proper focal point, however, is the screen noise, which agrees well with (10) but is, on the other hand, more than twice the magnitude predicted by (1).

Finally, Figure 4 is presented to demonstrate the extension of (10) to instances in which positive-ion production raises the noise level. The 4-mil tungsten filament of this pentode was formed from wire of ordinary commercial purity. Furthermore, there was never any detectable ion current, and yet at around 2600°K enough ions evaporated from the filament to raise Γ above unity. The behavior is typical of all cathodes but is especially marked in tungsten filaments merely because the operating temperature is high. The first ions evaporated (about 2500°K) are probably surface contaminations, largely potassium. Subsequently (about 3000°K) tungsten ions, themselves, predominate. Reasonably enough, the ion emission always falls off slowly with age at a fixed temperature, and is also greatly diminished by flashing at a higher temperature, even though the temperature increment be as little as 10° to 50°. This performance is reflected in fluctuation measurements. Consequently, the three plots of noise in Figure 4 do not represent precisely the same cathode conditions for the same temperature. The cathode noise, measured last, yields a value of Γ smaller than that prevailing in prior measurements; and this accounts for a portion of the small discrepancy between observed and calculated noise. The important point, however is the good agreement in the range for which $\Gamma > 1$.

On the whole there appears to be thorough experimental support for the hypothesis that shot effect in conventional multi-collectors

[1] See Part II.

adheres closely to the performance, summarized in (10), of the idealized structure depicted in Figure 1b. In conjunction with the evidence of Part II, and subject to the following summarized group of conditions:

a) $I/I_s << 1$

b) No positive ion emission

c) No secondary emission

d) No elastic reflection of electrons into the virtual cathode

e) Current collected at plate and screen only. The expressions (11) and (13) can be relied upon to give an accurate estimate of the effective input shot noise of modern amplifying tubes, and thus serve a useful purpose in analysis and design.

(To be continued)

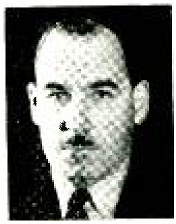
OUR CONTRIBUTORS

M. C. BANCA was born in Illinois, January 11, 1908. He obtained his B.S. degree at the University of Illinois, 1930, and took two years Electrical Engineering at University of Pennsylvania Evening School. Has filled the position of Research Engineer at RCA Manufacturing Company, 1930 to 1933; Philco, 1933 to 1935; RCA, 1935 to date.

J. F. BENDER was born in Lancaster, Pa., April 29, 1892. Educated in Central High School, Philadelphia, and engineering evening schools. Worked from 1912 to 1926 on mechanical design, New York shipyard; 1926-1928, plant engineering and safety engineering, RCA Manufacturing Company; 1934 to 1936, chemical plant equipment, Dupont and Hercules Powder Company; 1936 to date, RCA Manufacturing Company, plant engineering and mechanical design.

IRVING F. BYRNES entered the General Electric Test Department in 1918 and later engaged in radio development in their Engineering Laboratory. From 1920 on he was occupied in the development of radio equipment for commercial and military vessels, submarines and aircraft. He participated in the design and tests of the early ship-to-shore duplex radio telephone equipment used on the *SS. America* in 1922. Mr. Byrnes joined the Engineering Department of RCA Manufacturing in 1930, later transferring to the Radiomarine Corporation of America in charge of engineering activities.

WARREN ROBERT FERRIS was born May 14, 1904 in Vigo County, Indiana. He received his B.S. degree in electrical engineering at the Rose Polytechnic Institute in 1927, and his M.S. degree at Union College in 1932. He entered the service of the General Electric Company at Schenectady, as a member of the Research Laboratory staff during 1927-1930. In 1930 he transferred his services to the Research and Engineering Department of RCA Manufacturing Company at Harrison, New Jersey, where he is today. Mr. Ferris is a member of the Institute of Radio Engineers.



RAYMOND F. GUY entered the marine service of the Marconi Wireless Telegraph Company in 1916, resigning in 1918 to enlist in a regular army Signal Corps Replacement Company. After a year overseas he entered Pratt Institute, graduating in Electrical Engineering in 1921. After short periods as Inspector for the Shipowners Radio Service and The Independent Wireless Telegraph Company, he became, in 1921, one of the small pioneering group that built and operated WJZ, at that time licensed by the Westinghouse Company and located in the Newark plant. After three years of supervisory activities, Mr. Guy transferred to the RCA Research Department, heading the Broadcast Engineering Section, where he engineered the RCA stations, did consulting work for RCA clients, and directed the development of all broadcast transmitting apparatus used or sold by RCA. In 1929 he became the Radio Facilities Engineer of NBC. Since that time he has been responsible to the Vice President and Chief Engineer for the design, construction, and engineering of all of NBC's Standard Broadcasting, Television, UHF and International Radio Facilities, coverage studies and surveys, cost studies, frequency allocation, power tube engineering, development and design of antenna systems, propagation, etc. He is licensed to practise as a Professional Engineer in New York and New Jersey, is a Fellow of the Institute of Radio Engineers, a Fellow of the Radio Club of America, and a Member of the New York Electrical Society. During the last 10 years he has been active on technical committees of the IRE and other bodies.

HUMBOLDT W. LEVERENZ received his B.A. degree in chemistry from Leland Stanford Jr. University in 1930. Studied physics and chemistry as an Exchange Fellow of the Institute of International Education at the University of Münster, Westphalia, Germany from 1930-1931. From 1931 to the present time has been a chemico-physicist in the Research and Engineering Dept. of the RCA Manufacturing Co. Inc., engaged in synthesizing and applying materials used in the electronics art, particularly luminescent materials. Is a member of Phi Lambda Upsilon, The American Physical Society, The American Chemical Society, and The Franklin Institute of Philadelphia. Received an award as a "Modern Pioneer" from the National Association of Manufacturers for inventions in the television field.

DR. L. MARTON was born in Hungary in 1901. Received Ph.D. from University of Zurich, Switzerland, 1924. From 1924 to 1925, Research Assistant, University of Zurich. 1925 to 1928, Physical Chemist, Tungsram Lamp Company, Hungary. 1928 to 1938, Instructor, Research Fellow, and Assistant Professor at University of Brussels, Belgium. 1938, Lecturer, University of Pennsylvania. 1938 to date, Research Engineer, RCA Manufacturing Company.

DWIGHT O. NORTH received his B.S. degree from Wesleyan University in 1930 and his Ph.D. degree from the California Institute of Technology in 1933. Since 1934, Dr. North has been with the Research and Engineering Department of the RCA Manufacturing Company at Harrison, N. J., engaged principally in research studies of tube and circuit noise. He is a member of The Institute of Radio Engineers and a member of the American Physical Society.

TECHNICAL ARTICLES BY RCA ENGINEERS

Published Third Quarter, 1940

ALBURGER, J. R.—Mathematical Expression of Developer Behavior—*Journal of the Society of Motion Picture Engineers,* September.

ANDERSON, E. I.—see SEELEY and ANDERSON.

BARACKET, A. J.—Demonstrating Broadcasting Facsimile at the New York World's Fair—*RCA Review,* July.

BEDFORD, A. V.—see KELL, BEDFORD and FREDENDALL.

———and J. P. SMITH—A Precision Television Synchronizing Signal Generator—*RCA Review,* July.

BROWN, GEORGE H., and J. EPSTEIN—An Ultra-High-Frequency Antenna of Simple Construction—*Communications,* July.

CROSBY, M. G.—Reactance-Tube Frequency Modulators—*RCA Review,* July.

DREW, R. O., and E. W. KELLOGG—Starting Characteristics of Speech Sounds—*The Journal of the Acoustical Society of America,* July.

———and E. W. KELLOGG—Filtering Factors of the Magnetic Drive—*Journal of the Society of Motion Picture Engineers,* August.

EDDY, W. C.—Remote Control Television Lighting—*Journal of the Society of Motion Picture Engineers,* September.

EPSTEIN, J.—see BROWN and EPSTEIN.

FREDENDALL, G. L.—see KELL, BEDFORD and FREDENDALL.

GEORGE, R. W.—Field Strength Measuring Equipment at 500 Megacycles—*RCA Review,* July.

GLOVER, A. M., and R. B. JANES—A New High-Sensitivity Photosurface—*Electronics,* August.

GUY, RAYMOND F.—Moving and Rebuilding WEAF—*Association Technical Employees Journal,* August.

HOLLY, G. F.—Sound Screens: Structure and Function—*International Projectionist,* August.

JANES, R. B.— see GLOVER and JANES.

KELL, R. D., A. V. BEDFORD and G. L. FREDENDALL—A Determination of Optimum Number of Lines in a Television System—*RCA Review,* July.

KELLOGG, E. W.—Reversed Speech—*RCA Review,* July.

———see DREW and KELLOGG.

KOWALSKI, R. J.—Theatre Sound System Optical Data—*International Projectionist,* July.

KROGER, F. H., B. TREVOR and J. E. SMITH—A 500-Megacycle Radio-Relay Distribution System for Television—*RCA Review,* July.

LEVERENZ, H. W.—Optimum Efficiency Conditions for White Luminescent Screens in Kinescopes—*Journal of the Optical Society of America,* July.

MASSA, F.—see OLSON and MASSA.

NORTH, D. O.—Fluctuations in Space-Charge-Limited Currents at Moderately High Frequencies, Part II—Diodes and Negative-Grid Triodes—*RCA Review,* July.

OLSON, H. F. and F. MASSA—Applied Acoustics (second edition)—Book, P. Blakiston's Son & Co., Philadelphia.

PHELPS, WILLIAM D.—Power Transmission Loss in Exponential Horns and Pipes with Wall Absorption—*The Journal of the Acoustical Society of America,* July.

PREISMAN, ALBERT—Some Notes on Diode Detection—*Communications,* August.

SEELEY, S. W., and E. I. ANDERSON—U-H-F Oscillator Frequency Stability Considerations—*RCA Review,* July.

SMITH, J. E.—see KROGER, TREVOR and SMITH.

SMITH, J. P.—see BEDFORD and SMITH.

TREVOR, B.—Ultra-High-Frequency Propagation Through Woods and Underbrush—*RCA Review,* July.

———see KROGER, TREVOR and SMITH.

USSELMAN, G. L.—A V-T Voltmeter for Coaxial Line Measurements—*Electronics,* July.

WALLER, L. C.—An Efficient U-H-F Unit for the Amateur Television Transmitter—*QST,* July.