RADIO RECEIVER DESIGN

By

K. R. STURLEY

Ph.D., B.Sc., A.M.I.E.E.
Marconi School of Wireless Communication

Part I

RADIO FREQUENCY AMPLIFICATION AND DETECTION

NEW YORK
JOHN WILEY & SONS, INC.
LONDON: CHAPMAN & HALL, LIMITED

First Published . . . 1943

IN THE REPRINTING OF THIS BOOK, THE RECOMMENDATIONS OF THE WAR PRODUCTION BOARD HAVE BEEN OBSERVED FOR THE CONSERVATION OF PAPER AND OTHER IMPORTANT WAR MATERIALS. THE CONTENT REMAINS COMPLETE AND UNABRIDGED.

3/45

PRINTED IN THE UNITED STATES OF AMERICA

AUTHOR'S PREFACE

AN attempt has been made in this book to bring together the fundamentals of radio receiver design. Difficulties were experienced in determining the order of treatment, and it was finally decided to follow introductory chapters on general considerations and valves by a detailed examination of the receiver, stage by stage, starting from the aerial. There are objections to this method from the teaching point of view ; for example, the chapter on aerials is better considered after that on R.F. amplifiers, whilst the chapter on I.F. amplifiers should be read before the latter half of that on R.F. amplifiers. To meet possible criticism a fairly detailed table of contents is given, so that the reader can develop his own plan of campaign.

Owing to war conditions the book has had to be divided into two parts, the first ending at the detector stage, leaving Part II to deal with audio frequency amplifiers, power supplies, receiver measurements, television and frequency modulated receiver design, etc.

The cosine expression, $\hat{E} \cos \omega t$, for a voltage of sinusoidal wave shape is used in preference to the sine expression because it is considered that it leads to a simpler mathematical analysis. For the same reason the grid bias voltage is written as $-E_b$, i.e., E_b represents a numerical and not algebraical value of bias. The advantage of so doing is most evident in Chapter 8.

Part I is practically self-contained, though there are one or two cross-references to sections in Part II. To facilitate cross-reference all sections, figures and expressions are prefixed by the chapter number.

No claim is made to an exhaustive bibliography, and reference is made, at an appropriate point in the text, only to those articles which have proved helpful in its preparation.

The author is indebted to his wife for help in reading the proofs, to Mr. R. M. Mitchell, B.Sc., for checking many of the calculations, to Mr. R. P. Shipway, B.A., for helpful discussion on parts of Chapter 8, and to Marconi's Wireless Telegraph Company for permission to publish material originally used by the author in lectures at the Marconi School of Wireless Communication.

August 1942.

ACKNOWLEDGMENTS

ACKOWLEDGMENTS are gratefully made to the following for permission to use figures and drawings taken from their publications.

Name of Journal	*Figure Numbers*
Electronics	3.22, 3.25 5.9 6.21
Journal of the Institution of Electrical Engineers	5.11*a*
Marconi Review	7.13*a* and *b*, 7.14, 7.15
Mullard Technical Bulletin	6.20
Proceedings of the Institute of Radio Engineers	2.17 5.23*a*, *b* and *c* 8.9
Wireless and Electrical Trader (*Pye Radio*)	5.24*a*
R.C.A. Review	5.8*b*
Wireless Engineer	3.8*a* to 3.19*b* 4.3, 4.11, 4.13, 4.14 5.8*a* 7.7, 7.9, 7.10, 7.11*a* 8.12, 8.13*a*, 8.17, 8.24
Wireless World	5.24*b* 6.13, 6.14 8.21*a* and *b*, 8.22

PART I

CONTENTS

PART I

CHAPTER 1

GENERAL CONSIDERATIONS

1.1. Introduction. The direct transmission and reception of speech or music over long distances, though not impossible, is impracticable and propagation of audio frequencies is usually accomplished by using them to modulate an R.F. wave acting as a carrier, i.e., the audio frequencies are used to control one of the three characteristics, amplitude, frequency, or phase, of the carrier. The most common method is by modulation of the carrier amplitude, the rate at which the amplitude is changed being directly proportional to the frequency of the original sound and the magnitude of the amplitude change being directly proportional to the intensity (a low intensity producing a small change of amplitude). This is illustrated in Figs. 1.1*a* and 1.1*b* for an unmodulated carrier amplitude of 1 volt peak. Fig. 1.1*a* corresponds to a low-intensity sound, the carrier amplitude varying between 0·9 and 1·1 volts, whilst Fig. 1.1*b* corresponds to a high-intensity sound, the carrier being 100% modulated, its amplitude varying from 0 to 2 volts. In frequency modulation, the carrier amplitude remains constant and its frequency is varied at a rate corresponding to the modulation frequency (50 times per second if f mod. = 50 c.p.s.), and the frequency deviation—rise and fall from the central unmodulated carrier frequency—is controlled by the intensity of the audio frequency. For example, suppose a 1,000 c.p.s. note is being transmitted on a carrier frequency of unmodulated value 1,000 kc/s, the variation of the carrier frequency takes place at the rate of 1,000 changes per second and the frequency limits may be

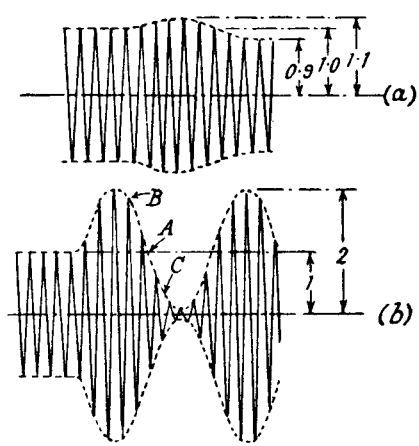


FIG. 1.1.—An Amplitude Modulated Wave.

± 100 c.p.s. (the carrier frequency changes between 1,000·1 and 999·9 kc/s) for a low intensity to $\pm 100{,}000$ c.p.s. for a high intensity note. These two conditions are illustrated in Figs. 1.2*a* and 1.2*b*. The frequency change of carrier is exaggerated in the figures for

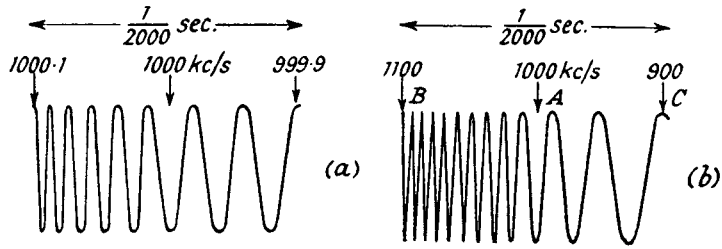


FIG. 1.2.—A Frequency Modulated Wave.

the purpose of illustration. With phase modulation the amplitude of the carrier remains constant and its phase angle with respect to its unmodulated condition is advanced and retarded at the frequency of the audio signal. The magnitude of the phase change is determined by the intensity of the audio frequency. Phase modulation has an effect on the carrier akin to frequency modulation, and whilst the phase of the carrier is varying, its frequency also is varying. There is a difference, but this is discussed later in Section 1.4.

Apart from the advantage of transmission over long distances the use of carriers enables many programmes to be transmitted simultaneously, any one of which may be selected by the listener with suitable receiving apparatus. This apparatus must abstract energy from the radiated electromagnetic carrier wave, separate desired from undesired signals, extract the audio frequency signal from the carrier, and amplify and reproduce the signal to any desired level. An aerial system is necessary to collect energy from the modulated carrier wave, the flux from which induces a voltage in the aerial as it passes. The function of the aerial is usually to act only as a collector, but it is sometimes employed as a discriminator against undesired signals, viz., in directional aerial systems. The separation of these signals is carried out by means of tuned circuits in the anode and grid circuits of the radio frequency valves, which amplify the desired modulated carrier. The audio frequency signal is extracted from the carrier by a suitable form of detector and is amplified to a level sufficient to operate the apparatus (telephones or loudspeaker) for converting electrical audio frequency energy into acoustical energy.

1.2. Amplitude Modulation. For understanding the function of any type of modulation a vector diagram is particularly useful, and amplitude modulation is quite simply illustrated if the envelope is sinusoidal as in Figs. 1.1*a* and 1.1*b*. The modulated signal is represented by

$$\hat{E} \cos \omega_c t(1+M \cos pt) \quad . \quad . \quad . \quad . \quad . \quad . \quad 1.1$$

where $\hat{E}$ = the unmodulated carrier peak voltage.
$\omega_c = 2\pi f_c$, the carrier pulsance.
$p = 2\pi f_m$, the modulating frequency pulsance.
M = modulation ratio, the maximum value of which is 1.

The factor 1 inside the bracket in expression 1.1 is necessarily included because the carrier still exists when there is no modulation, i.e., when $M = 0$. The vector representing expression 1.1 rotates at a speed of f_c c.p.s. and varies in amplitude f_m times per second between the limits $E(1+M)$ and $\hat{E}(1-M)$ as shown in Fig. 1.3*a*.

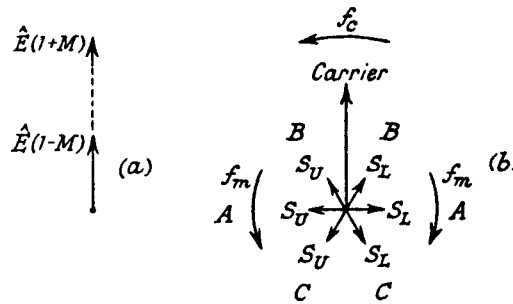


FIG. 1.3.—Alternative Vector Diagrams for an Amplitude Modulated Wave.

Expression 1.1 may be expanded as follows:

$$\hat{E} \cos 2\pi f_c t + \frac{\hat{E}M}{2} \cos 2\pi(f_c+f_m)t + \frac{\hat{E}M}{2} \cos 2\pi(f_c-f_m)t \quad . \quad 1.2$$

which consists of a carrier vector E rotating at f_c c.p.s. and two equal sideband vectors $\frac{\hat{E}M}{2}$ rotating at (f_c+f_m) and (f_c-f_m) c.p.s. about the same point as the carrier vector. Since these two sideband vectors only influence the carrier amplitude, it is clear that their resultant must always be in line with the carrier, and this is shown in Fig. 1.3*b*, for successive instants of time corresponding to points *A*, *B* and *C* on Fig. 1.1*b* The upper frequency sideband vector S_u is rotating round the carrier vector in an anti-clockwise direction (the accepted positive direction of frequency) at f_m c.p.s., whilst the lower frequency sideband vector S_L rotates in a clockwise direction round

the carrier vector at the same frequency. The carrier vector in Fig. 1.3*b* is, for convenience, shown as stationary, though actually it moves round in an anti-clockwise direction at f_c c.p.s.

Amplitude modulation of a carrier is achieved by controlling the gain of a carrier frequency amplifier in the transmitter by means of the A.F. signal. This can be realized by using the latter to vary the control grid or suppressor grid bias (of a pentode amplifier valve) or the anode voltage of a triode valve amplifier. Anode modulation requires a much larger A.F. signal than grid modulation, but the modulation envelope is much less distorted at high modulation percentages. To obtain the A.F. signal from the carrier, the amplitude variations must be detected at the receiver, i.e., a device is required to produce mean voltage variations in accordance with the carrier amplitude variations. This is achieved either by suppressing or reducing one-half of the amplitude modulated carrier, for this disturbs the symmetry of the wave shape (Fig. 1.4) with regard to the centre line and so produces a changing

FIG. 1.4.—A Detected Amplitude Modulated Wave.

mean voltage. A detector suppressing completely one-half of the modulated carrier and passing the other without changing its shape is known as a linear detector.

1.3. Frequency Modulation.[1, 4, 9] The problem of representing frequency modulation by a single vector is complicated by the fact that the frequency of the carrier is varying in accordance with the amplitude of the modulating frequency, and this means that the carrier vector rotates at varying speeds. Taking the example given above, for a frequency deviation of $\pm 100{,}000$ c.p.s. the carrier varies from 1,100 kc/s to 900 kc/s, and if we, acting as observers of the carrier vector, were to rotate at a frequency f_c c.p.s. (the carrier unmodulated value) in the same direction as the carrier vector, the latter would appear to oscillate backwards and forwards like a metronome, at the frequency of the modulating wave, i.e., 1,000 times per second. This condition is illustrated in Fig. 1.5*a*; like the metronome, the vector is stationary at the extremes, positions X' and X'', of its stroke so that the carrier frequency is instantaneously at its unmodulated value f_c (this is point A in Fig. 1.2*b*), whilst it is moving at its fastest, backwards

or forwards at the centre X of its swing. Movement of the vector in an anti-clockwise direction means that the frequency is greater than f_c and in a clockwise direction the reverse. Hence point X when the vector moves anti-clockwise corresponds to B in Fig. 1.2*b*, but when the vector is travelling clockwise X corresponds to C. In frequency modulation the carrier deviation is fixed for a given amplitude of modulating input, so that the speed of the vector as it passes through X is constant and independent of the frequency of oscillation backwards and forwards (that of the modulating input). Treating the problem as a mechanical one, the initial velocity at X' or X'' is zero, and the final velocity v at X is constant, so that the distance travelled from X' or X'' to X is proportional to velocity $\times$ time, but time is proportional inversely to the modulating frequency so that

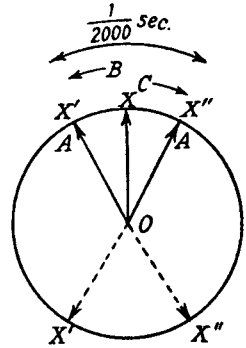


FIG. 1.5*a*.—Vector Diagram for a Frequency Modulated Wave.

$$X'X = X''X = Kvt = K_1 t = \frac{K_2}{f_m} \propto \frac{1}{f_m}$$

Hence the angle swept out by the carrier vector is large for low modulating frequencies and small for high modulating frequencies; dotted positions X' and X'' correspond to a lower modulating frequency than the full line positions. This again is analogous to the metronome, which increases its angle of sweep as the frequency decreases. The mathematical expression for a frequency modulated wave is

$$\hat{E} \cos\left[2\pi f_c t + \frac{Mf_c}{f_m} \cdot \sin 2\pi f_m t\right] \quad . \quad . \quad . \quad 1.3$$

where $$M = \frac{\text{deviation frequency}}{\text{carrier frequency}}$$

and the angle swept out by the vector is $\frac{Mf_c}{f_m}$ radians or $57{\cdot}3\ Mf_c/f_m$ degrees, which is seen to be inversely proportional to f_m. In the same way that expression 1.1 for amplitude modulation can be turned into a carrier and sidebands, so can expression 1.3 be treated, but instead of only two sidebands per fundamental modulating frequency it is found that there is a large number [1, 2] of sidebands spaced from the carrier by frequencies of $\pm f_m$, $\pm 2f_m$, $\pm 3f_m$, etc.

The expansion of expression 1.3 gives a carrier and infinite

series of sidebands having amplitudes corresponding to Bessel functions of ascending order, i.e.,

$$\hat{E}\cos\left[2\pi f_c t+\frac{Mf_c}{f_m}\sin 2\pi f_m t\right]$$
$$=\hat{E}\Big\{J_0\left(\frac{Mf_c}{f_m}\right)\cos\omega_c t$$
$$-J_1\left(\frac{Mf_c}{f_m}\right)[\cos(\omega_c-p)t-\cos(\omega_c+p)t]$$
$$+J_2\left(\frac{Mf_c}{f_m}\right)[\cos(\omega_c-2p)t-\cos(\omega_c+2p)t]$$
$$-J_3\left(\frac{Mf_c}{f_m}\right)[\cos(\omega_c-3p)t-\cos(\omega_c+3p)t]$$
$$+\ \ldots\ \text{etc.}\Big\}\qquad\ldots\qquad 1.4$$

where $\omega_c = 2\pi f_c$ and $p = 2\pi f_m$.

$\frac{Mf_c}{f_m} = \frac{\text{deviation frequency}}{\text{modulating frequency}}$, sometimes designated as the modulation index,[2] may have any value from 1 upwards. For example, if the deviation frequency is ± 75 kc/s and the modulating frequency 10 kc/s, $\frac{Mf_c}{f_m} = 7{\cdot}5$. The amplitude of the carrier and sidebands can be obtained from tables of Bessel functions.[10, 11]

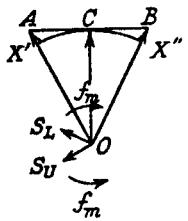


FIG. 1.5*b*. — Addition of Two Sidebands to Carrier to give Frequency Modulation.

The theoretically infinite number of sidebands is for all practical purposes fortunately limited, as the amplitudes of the sidebands more distant from the carrier normally decrease very rapidly. It is not only in the number of sidebands that frequency modulation differs from amplitude, but also in their position with respect to the carrier. All odd-numbered sideband vectors $\pm f_m$, $\pm 3f_m$, etc., from the carrier are so placed that their resultant falls on a line at right angles to the carrier [3] vector and all even-numbered sidebands have a resultant in line with the carrier vector. The addition of a carrier and two sideband vectors (spaced $\pm f_m$ from the carrier), having a resultant 90° displaced from the carrier vector, is shown in Fig. 1.5*b*. These two sideband vectors rotate round the carrier vector point 0 at f_m c.p.s., and their resultant adds to the carrier to cause the latter to oscillate backwards and forwards along the line AB. In fact, it is the same as Fig. 1.5*a*, apart from the ampli-

tude variation of the vector, and by adding suitable amplitudes of wider-spaced sidebands we can neutralize the amplitude variation, causing the point of the vector to describe an arc $X'X''$ giving only frequency modulation. Another important point to note with regard to frequency modulation sidebands is that the individual amplitudes are not directly proportional to the amplitude of the original modulating frequency as in amplitude modulation, but actually vary widely (sometimes becoming zero) as the angle swept out by the carrier vector changes. When this angle is small (f_m high) all sidebands except those nearest to the carrier are so small in amplitude that they can be neglected, but for a large angle of oscillation (f_m low) there are many sidebands of appreciable amplitude. For example, if $f_m = 50$ c.p.s. and its amplitude such as to give a carrier frequency change of $\pm 1{,}000$ c.p.s., the angle swept out by the vector moving from X' to X is $\frac{1000}{50} = 20$ radians (1,146°), and there are 23 sidebands of importance with amplitudes exceeding 1% of the fundamental sideband. If, however, $f_m = 1{,}000$ c.p.s., and the carrier frequency change is as before, the angle swept out by the vector is 1 radian (57·3°) and there are now only three sidebands of importance. It should be noted that the carrier vector may make a number of revolutions for low modulating frequencies.

Frequency modulation of a transmitter may be accomplished by varying the equivalent inductance or capacitance of the tuned circuit of the oscillator producing the carrier frequency (or a submultiple of this frequency). Since greater frequency stability is obtained at low frequencies it is quite common for the " master " oscillator driving the transmitter to generate a submultiple of the output carrier frequency, the submultiple being stepped up to the required output frequency by passing it through a series of multiplier stages, producing 2nd and 3rd harmonics of the input frequency. When the required frequency deviation of carrier is small compared with the carrier frequency—probable values are ± 75 kc/s at a carrier frequency of 45 Mc/s, representing a frequency deviation of $\pm 0{\cdot}001666 f_c$—the required change of tuning inductance or capacitance is linearly proportional to the frequency deviation. Hence a linear relationship between the amplitude of the A.F. signal and the change of inductance or capacitance results in frequency modulation having a deviation frequency linearly proportional to the amplitude of the A.F. signal. The variable reactance [8] valve (see Chapter 13 in Part II), as used in automatic frequency correction circuits, is particularly suitable for converting the A.F. amplitude into a change

of inductance or capacitance, since it acts as a reactance to any voltage source connected between its anode and cathode, the value of the reactance depending on the grid bias of the valve. If the relationship between the grid bias of the valve and its mutual conductance is linear, change of the former produces a linearly proportional change of reactance. Thus if the A.F. modulating voltage is applied to the grid circuit of the reactance valve it controls the mutual conductance in such a way as to produce a change of equivalent inductance or capacitance between the anode and cathode, which is linearly proportional to the A.F. voltage amplitude.

1.4. Phase Modulation. The vector representation of phase modulation is similar to that of frequency modulation with one important difference ; the angle swept out by the vector is constant for all modulating frequencies and is dependent only on the amplitude

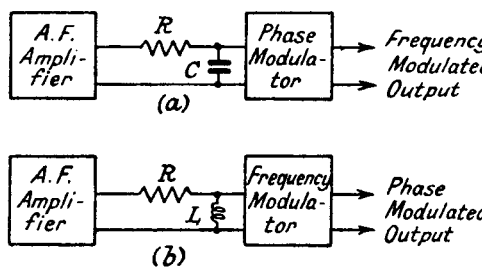


FIG. 1.6.
(*a*) Frequency Modulation from a Phase Modulator.
(*b*) Phase Modulation from a Frequency Modulator.

of the latter. This essentially means that the velocity of the vector at the centre X (Fig. 1.5*a*) increases as the speed of oscillation increases, i.e., the rise and fall in carrier frequency (frequency deviation) is directly proportional to the modulating frequency. It is possible to turn phase into frequency modulation by inserting in the modulation frequency amplifier a network having an amplification characteristic inversely proportional to frequency, giving a low frequency a large amplitude and a high frequency a small amplitude, i.e., the phase angle is no longer constant but increases as the modulation frequency falls. The reverse is also true ; frequency modulation can be converted to phase modulation by inserting in the modulation amplifier a network having an amplification characteristic directly proportional to frequency giving a low frequency a small and a high frequency a large amplitude. This is shown in Figs. 1.6*a* and 1.6*b* ; the first figure has a resistance–capacitance network in the modulation amplifier, thus producing

frequency modulation at the output of the phase modulator. In the second figure a resistance–inductance network produces phase modulation from a frequency modulator.

The modulated carrier expression is

$$\hat{E} \cos [2\pi f_c t + \phi M \cos 2\pi f_m t] \quad . \quad . \quad . \quad 1.5$$

and the phase angle swept out is equal to ϕM where ϕ is a constant and M is proportional to the amplitude of f_m but independent of its frequency. Phase modulation has an infinite number of sidebands spaced $\pm f_m$, $\pm 2f_m$, etc., from the carrier, but since the phase angle change is constant (unlike frequency modulation) the wider spaced sidebands for low as well as high modulating frequencies can usually be neglected. The resultants of all odd-numbered sidebands are 90° out of line with the carrier and those of the even-numbered ones are in line with the carrier vector. The amplitudes of the carrier and sidebands are Bessel Coefficients similar to those for frequency modulation, but the amplitude of the sidebands is dependent only on the amplitude of the modulating frequency. Thus

$$\begin{aligned} &\hat{E} \cos [2\pi f_c t + \phi M \cos 2\pi f_m t] \\ &= \hat{E} \{J_0(\phi M) \cos \omega_c t \\ &\quad - J_1(\phi M)[\cos (\omega_c - p)t - \cos (\omega_c + p)t] \\ &\quad + J_2(\phi M)[\cos (\omega_c - 2p)t - \cos (\omega_c + 2p)t] \\ &\quad - J_3(\phi M)[\cos (\omega_c - 3p)t - \cos (\omega_c + 3p)t] \\ &\quad + \ldots \text{ etc.}\} \quad . \quad . \quad . \quad . \quad . \quad . \quad 1.6 \end{aligned}$$

where ϕM is measured in radians.

The clearest distinction between frequency and phase modulation is indicated by considering a square-shaped modulating wave [7] as in Fig. 1.7. With frequency modulation the carrier frequency varies above and below its unmodulated value in accordance with the amplitude of the modulation. Phase modulation, however, shows that apart from the sudden instantaneous changes of phase when the carrier frequency becomes $+\infty$ or $-\infty$ the latter is constant at its unmodulated value.

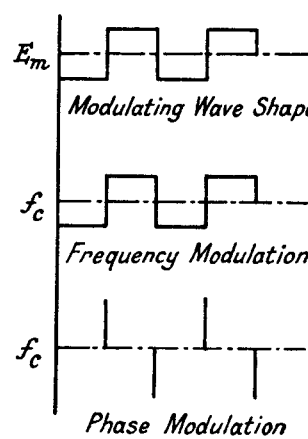


FIG. 1.7.—Illustration of the Difference between Frequency and Phase Modulation.

Phase modulation [5] can be achieved by separating the two sidebands of

an amplitude modulated carrier from the carrier and passing them through a phase adjusting network, which places their resultant at 90° to the carrier vector, as shown in Fig. 1.5*b*, and then recombines them with the carrier. Fig. 1.5*b* shows that amplitude modulation is also produced, but it can be suppressed by passing the phase modulated wave through a limiter (an amplitude distorting device giving almost constant carrier output for large changes of input). Phase change is not exactly proportional to the modulating voltage amplitude, but if its maximum value is limited to $\pm 25°$ deviation, distortion of the modulation does not exceed 3%.[6] Such a small phase change as $\pm 25°$ would represent a low modulation percentage, but if the original carrier is a submultiple of the required output frequency the phase angle is multiplied at the same time as the carrier in the frequency multiplier stages of the transmitter. For example, if the original carrier is one-tenth of the required, the phase angle is increased to $\pm 25° \times 10 = \pm 250°$ after passing through the multiplier stages. The extra sidebands essential to the larger phase angle are produced in the multiplier stages.

As far as the receiver is concerned, the main features of design are unaffected by the type of modulation. The frequency or phase modulation receiver requires two extra stages over its amplitude modulation counterpart; these are the amplitude limiter and frequency-amplitude converter. The latter is essential because the character of the original A.F. voltage modulating the carrier is that of amplitude variation, and it converts the frequency change of carrier into an amplitude change, which is then detected in the normal way by a diode or similar device. The converter does not remove the frequency variation, but this does not matter because the diode is sensitive only to amplitude change. The limiter stage is necessary to suppress amplitude change (due to transmission variations or noise) of the carrier before the converter, since the detector will respond to these undesired changes. Apart from these two modifications there is no vital difference between a frequency or phase modulation and an amplitude modulation receiver operating over the same range of carrier frequencies. In the chapters which follow, the amplitude modulation receiver is treated in detail, whilst the modifications necessary for reception of frequency or phase modulation are discussed in Chapter 15 of Part II.

1.5. Types of Amplitude Modulation Receivers. The simplest type of receiver is that consisting of a tuned circuit, connected to an aerial, and a detector, operating telephones. Selectivity

and output are low, though the latter may be increased to a value sufficient to operate a loud speaker by adding an audio frequency amplifier. Selectivity can be increased, but at the expense of output, by interposing more tuned circuits between the aerial and detector. With certain types of valve detectors selectivity and output may be increased by applying reaction, i.e., feeding radio frequency modulated carrier energy back from the anode to the grid circuit. The disadvantage of this method is that it produces a sharply peaked resonance curve, which discriminates against the higher audio frequencies modulating the carrier.

For adequate selectivity and low detector distortion, radio frequency amplification is essential. The selective property of a tuned circuit cannot be fully utilized if it is coupled to another tuned circuit, for each tends to damp the other. By separating tuned circuits with valves, maximum selectivity is achieved and the signal is also amplified. Most detectors are only linear over a limited range and above a given value of input signal, and R.F. amplification is necessary to raise the modulated carrier to this level.

A possible type of receiver is therefore one having R.F. and A.F. amplification. Owing to the difficulty of maintaining stability and the mechanical limitations imposed by the ganged variable capacitor, two stages of R.F. amplification are seldom exceeded for medium- and long-wave band operation. The amplification of an R.F. stage is proportional to $\frac{L}{CR}$, where L, C and R are the constants of the tuned circuit and, since L is constant and C decreases as frequency increases, amplification is not constant over the wave band but tends to increase with increasing frequency (in spite of the rise in R). Selectivity tends to fall as the frequency is increased. The tuning range of the capacitor is usually the same for each wave band, so that L is reduced as higher frequency wave ranges are selected. The value of L is so small on the short wave range that amplification is very low, and receivers for short wave operation are practically never dependent on the R.F. stages for amplification.

The difficulties of obtaining adequate carrier frequency amplification and constant selectivity over a given tuning range can largely be overcome by using the superheterodyne principle and converting all incoming R.F. carriers to a constant (generally) low R.F. frequency, at which high amplification and good selectivity can be obtained; the modulation sidebands are, of course, transferred to the new

fixed frequency carrier, which is called the intermediate frequency. The frequency conversion is carried out in a frequency changer valve by mixing the incoming modulated carrier with a local oscillation of frequency equal to the sum of the carrier and intermediate frequencies. The superheterodyne receiver is the most usual form of receiver for operation over all ranges from long to short waves. It tends to reduce the amplification and selectivity variations over the wave ranges by concentrating most of the amplification and selectivity in the intermediate frequency stage. There are disadvantages as well as advantages to the use of the superheterodyne principle, but they can be overcome by careful design. The local oscillator output may reach the aerial and be reradiated, causing interference in nearby receivers, and whistles may be produced by interaction between undesired carriers and the local oscillator and its harmonics. Reradiation can be reduced by

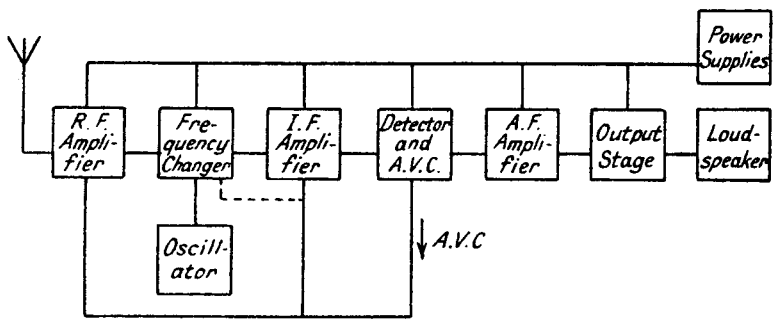


FIG. 1.8.—Schematic Diagram for a Typical Superheterodyne Receiver.

reducing the stray coupling in the frequency changer between the local oscillator and the aerial circuit. The inclusion of an R.F. amplifier valve between the aerial and signal circuit of the frequency changer almost entirely eliminates reradiation and has two other important advantages; it increases the signal-to-noise ratio when the noise is mostly produced in the first valve of the receiver, and it increases R.F. selectivity, so reducing undesired signals and whistle interference. The latter interference can be reduced by careful design of the frequency changer anode current-oscillator grid voltage characteristic.

A schematic diagram of a typical general purpose broadcast receiver is shown in Fig. 1.8. The aerial is coupled to the R.F. amplifier stage (usually a single valve, but in some communication type receivers two). This is followed by a frequency changer with

local oscillator, the output from which is fed to the intermediate frequency amplifier, consisting of one or two stages of amplification. A three-valve I.F. amplifier is usually necessary on receivers for television and frequency modulation. A diode detector is normally employed because of its linear detection characteristic with large signals ; carrier voltages exceeding 2 volts peak and modulated to 90% can be detected by the diode with very little distortion of the audio voltage output. An audio frequency amplifier usually precedes the power output stage. The amplification of the receiver should be such as to make it capable of reproducing with adequate output all signals above the noise level. Any increase in amplification beyond this is of little value except in so far as it improves automatic volume control action.

Most receivers incorporate accessory circuits aimed at making control less complicated, and two important ones are those providing automatic volume control, more correctly termed automatic gain control, and automatic frequency correction. Transmission of an electromagnetic wave through space is accomplished by direct and " sky reflected " waves, and beyond the service or local area much of the transmission is carried by the indirect ray reflected from the ionosphere surrounding the earth. The strength of the reflected ray is subject to slow or rapid fluctuations due to changes in the condition of the ionosphere, and the received input signal variations cause a corresponding change in audio output unless special measures are taken to vary the receiver gain in an inverse relationship to the signal strength. The gain of a receiver may be automatically controlled by using the D.C. component, produced by detection of the carrier, to vary the bias on the radio frequency, frequency changer (not always) and intermediate frequency stages, reducing the gain of these stages when the input signal increases, and vice versa. This is the principle involved in automatic gain control. A further advantage of A.G.C. is that it prevents " blasting " of the loudspeaker in tuning from a weak to a strong signal. Sometimes incorporated in an A.G.C. system is a device for silencing the receiver when the input signal is below a predetermined level—generally one governed by the noise level. Such a device, termed quiet automatic gain control or interstation noise suppression, removes one of the disadvantages of tuning with A.G.C., viz., that in the absence of a signal the gain of the receiver is maximum, thus amplifying receiver noise to an objectionable level.

Automatic frequency correction of the local oscillator in a superheterodyne receiver can be used to compensate for inaccuracy in

tuning, or drift of the oscillator, due to temperature changes, from its originally correct setting. It is particularly useful in remotely controlled and push-button operated receivers.

Push-button tuning is incorporated in order to facilitate rapid selection of, or changeover from, one station to another. The push-button switches may operate to switch in separate circuits tuned to the selected station or to control mechanically by cams, or electrically by a motor, the travel of the main tuning capacitor. Electrical motor control is readily adaptable to remote control.

In most receivers provision is made for the reproduction of gramophone records by switching the volume control of the receiver to the pick-up from the gramophone.

A third type is the superregenerative receiver, which may be used for ultra short wave communication purposes when quality of reproduction is unimportant and high sensitivity is required from simple apparatus. The receiver usually consists of a quenching valve, and an A.F. amplifier preceded by a detector, in which regeneration has been carried to the point of oscillation. The smallest signal is capable of initiating oscillation of the detector, and in the absence of control the oscillation will build up to a value determined by the valve characteristics. The rate of rise of oscillation amplitude, which is exponential, is proportional to the amplitude of the initiating signal and inversely to the inductance of the circuit tuning the signal. To make the valve useful as a detector it is necessary to stop periodically the oscillation and reset it to the sensitive condition. The interrupting frequency is supplied from a separate oscillator (in the simplest case the detector may be made self-setting by allowing it to develop a squegger oscillation) in series with the anode or grid circuit of the detector, so that with each negative half-cycle the latter is taken to cut-off and oscillation ceases. It is then ready for " triggering " by the signal once again. Since the rate of rise of oscillation is determined by the amplitude of the initiating signal, the average current taken by the valve follows approximately the modulation envelope of the signal. Owing to the very sensitive condition of the detector the smallest signal is capable of initiating oscillation and the noise level of the receiver is consequently high. In addition selectivity is low due to damping from the detector valve and it cannot be increased by including additional stages of R.F. amplification because positive feedback from the detector causes damped oscillations to occur in the added tuned circuits, thus masking the desired signal. For high sensitivity the ratio of initiating signal to quench frequency (about 40 kc/s) must be high ; this and

low selectivity confines the use of the superregenerative receiver to the ultra short wave band. Distortion of the modulation envelope of the received signal and high noise level make the receiver suitable only for communication for which reasonable intelligibility and not quality is the criterion.

1.6. Design Considerations Based on the Power Supply. The power supply has an important influence upon the receiver. For example, the H.T. battery in a battery receiver has a maximum voltage of 120 and a " life " capacity of approximately 0·2 kw. hrs., hence the power output of the receiver must be limited, the permissible distortion limit raised, and current consumption reduced to the lowest possible value. Since the output stage absorbs most of the current, economical working must be achieved by the use of special types of overbiased push-pull circuits, such as Class B and quiescent push-pull, which adjust their current consumption to the volume of the transmitted programme. Valves in the other stages of the receiver must be designed for maximum gain with minimum current and this can only be achieved in the R.F., frequency changer and I.F. stages at the expense of the variable-*mu* characteristic, the curvature of which must be greater, leading to higher distortion and cross-modulation effects than occur with their counterparts in the mains receiver.

In a mains-driven receiver for A.C./D.C. operation the H.T. voltage is limited by the mains connection, but large power can be obtained by high current consumption, and the output stage must be designed for low voltage and high current. The heaters of the valves are connected in series for economical operation.

A receiver for A.C. mains operation imposes practically no design limitations since with a suitable transformer the power source can be made to supply any desired current or voltage.

Cost also dictates the design of the receiver. High valve cost limits the number of stages and demands the highest amplification per stage. It also leads to the use of a pentode or beam tetrode valve in the output stage because of its high power sensitivity.

In the following chapters is described design procedure for each stage of the receiver, beginning at the aerial and progressing to the output. Part I ends with detectors and Part II, which begins with audio frequency amplifiers, includes chapters on television and frequency modulation receiver design.

BIBLIOGRAPHY

1. Notes on the Theory of Modulation. J. R. Carson, *Proc. I.R.E.*, Feb. 1922, p. 27.
2. Frequency Modulation. B. Van der Pol, *Proc. I.R.E.*, July 1930, p. 1194.
3. Note on the Relationship existing between Radio Waves Modulated in Frequency and Amplitude. C. H. Smith, *Wireless Engineer*, Nov. 1930, p. 609.
4. Amplitude, Phase and Frequency Modulation. H. Roder, *Proc. I.R.E.*, Dec. 1931, p. 2145.
5. A Method of Reducing Disturbances in Radio Signaling by a System of Frequency Modulation. E. H. Armstrong, *Proc. I.R.E.*, May 1936, p. 689.
6. Armstrong's Frequency Modulator. D. L. Jaffe, *Proc. I.R.E.*, April 1938, p. 475.
7. Frequency or Phase Modulation. Editorial Note, *Wireless Engineer*, Nov. 1939, p. 547.
8. Reactance Tube Frequency Modulators. M. G. Crosby, *R.C.A. Review*, July 1940, p. 89.
9. Amplitude, Frequency and Phase Modulation. Editorial, *Wireless Engineer*, Aug. 1940, p. 339.
10. *A Treatise on Bessel Functions.* Gray, Mathews and MacRobert. Text-book.
11. *Bessel Functions for Engineers.* N. W. McLachlan. Oxford University Press. Text-book.

CHAPTER 2

VALVES

2.1. Introduction.[9, 10] The all-important role played by the valve in receiver circuit design necessitates some knowledge of its constructional details and the principles, chief of which is thermionic emission, involved in its operation. An understanding of thermionic emission is helped by reference to the solar system hypothesis of the basic construction of matter. According to this theory, which, though superseded by later theories giving a more adequate explanation of certain phenomena, is very satisfactory in interpreting some of the simple actions, the atoms of all elements have a more-or-less compact central positively charged nucleus surrounded by electrons revolving in orbits rather like planets round the sun. The nucleus carries most of the mass of the atom and has a gross positive charge equal to the atomic weight of the element being considered ; part of this charge is neutralized by electrons in the nucleus itself and a net positive charge numerically equal to the atomic number is left. The atom is rendered electrically neutral by the orbital electrons, which equal in number the net positive charge on the nucleus. The maximum number of electrons, which can be maintained in each orbit, is limited, the first to 2 (2×1^2), the second and third to 8 (2×2^2), the third and fourth to 18 (2×3^2), and so on. The number of outer orbit electrons largely determines the activity of the element, inert elements like Helium, Argon, etc., having complete outer orbits and the most chemically active, the alkaline group (Lithium, Sodium, Potassium, etc.) headed by Hydrogen, having only one electron. The Potassium atom, fourth in the alkaline group, is illustrated in Fig. 2.1. Its atomic weight (total positive charge on the nucleus) is 39 and its atomic number (electrons in outer orbits) is 19, leaving 20 electrons in the nucleus. There is no evidence to fix the actual shapes of the orbits, which for convenience are shown as circular.

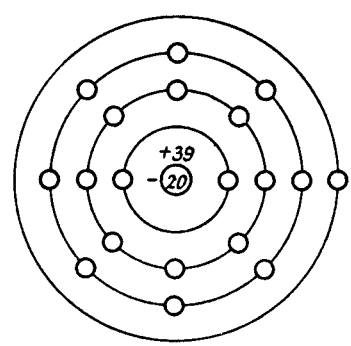


FIG. 2.1.—A Representation of the Potassium Atom.
(Atomic Weight 39, Atomic Number 19.)

In solid substances the atomic structure may allow the detachment of electrons from the atoms and the movement of these " free " electrons and the resultant positively charged atoms within the substance. Other solid media may, on the other hand, not allow the formation of these free electrons. Metals generally have many free electrons, and their motion is random except under the action of a driving potential, which causes them to move in one direction, so constituting a conduction current through the medium. The number of free electrons is indicative of the conductivity of the medium, a large number implying a good conductor. Insulating substances have practically no free electrons and an applied potential causes displacement of the orbits but no general drift of electrons. Electrical energy can be transferred across an insulator, if an alternating potential is applied, much in the same way that mechanical energy can be transferred through an elastic coupling.

In the metallic atom the velocity of the free electrons is largely determined by the temperature. At normal room temperatures the velocity is insufficient to carry the free electron beyond the boundary of the metallic surface, but as the temperature is raised its velocity increases and may reach a high enough value to allow it to break through and escape. This characteristic of liberating electrons under the action of heat is known as thermionic emission. A positive electrode placed near the metal will collect the electrons and show a small current. The number of electrons collected by this electrode depends on

(1) the metallic material forming the boundary of the emitting surface ;
(2) the temperature of the emitter ;
(3) the attractive power of the collector electrode, i.e., the positive voltage difference between it and the electron emitting surface ;
(4) the nature and density of the gaseous medium between the metal and electrode. When the density of the gas is high, the number of atoms per unit volume is large and the electrons collide with them, losing velocity so rapidly that many fail to reach the collector. If the gas density is very low, electron collisions are much less frequent and many more reach the collector anode.

This is the principle underlying the operation of the thermionic valve, which in its simplest form, a diode, consists of a metal filament or cathode heated to a high temperature, and a collector anode at a

suitable positive voltage. Earliest receiving valves used Tungsten for the heated filament since this could be raised to a high temperature without serious evaporation of the metal occurring. It was found, however, that the ease with which an electron escaped from its metallic barrier plane varied with different metals, and certain ones, notably the alkaline earth metals like Barium and Strontium, emitted electrons more readily and at a much lower temperature than Tungsten. Barium and Strontium oxides form the coating for the high efficiency low temperature valve, and during manufacture a thin film of the metals is distilled through to the surface of the cathode to act as an emitter, giving a copious flow of electrons.

We will now briefly touch upon the characteristics of the different types of valves, starting with the simplest, the diode.

2.2. The Diode. The diode consists of an electron-emitting hot cathode and a collector anode. For electrons to be collected, the anode generally must have a positive voltage applied between it and the cathode, though, owing to the initial velocity of the electrons, some may be collected at low negative anode voltages (between — 1 and 0 volt). The form of the anode current-anode voltage curves is shown in Fig. 2.2. Owing to the negative charge from the cloud of liberated electrons surrounding the cathode, maximum anode current does not flow immediately the anode becomes slightly positive. At low positive anode voltages this electronic negative space charge neutralizes the positive attractive field from the anode, and the anode current builds up to such a value that the number of electrons in transit produces a negative charge, just neutralizing the positive field from the anode. Increase of E_a requires a higher negative space charge for equilibrium, and so more electrons are drawn across to the anode to establish this equilibrium. The process continues as E_a is increased until all available electrons are being drawn off from the cathode; for higher values of E_a anode current remains constant. The maximum or saturation value of anode current is determined by the temperature of the cathode, i.e., the heater voltage, and increasing heater voltage raises the saturation current as shown in Fig. 2.2, though it has little effect on the current in the non-saturated region, since

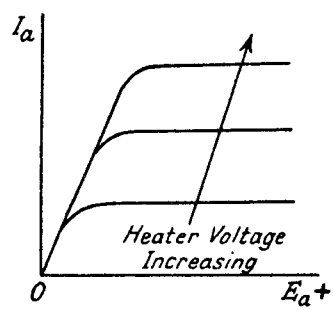


FIG. 2.2.—Diode I_aE_a Curves for Increasing Heater Voltage.

it is equivalent to increasing the reservoir of electrons without increasing the propulsive force from cathode to anode.

It can be shown that for concentric cathode and anode surfaces the anode current is

$$I_a = \frac{KE_a^{\frac{3}{2}}l}{r_a B^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.1$$

where K = constant depending on the emitting surface

l = length of cathode

r_a = radius of anode

B = a function of $\frac{r_a}{r_k}$

r_k' = radius of cathode.

Curve 1 in Fig. 2.3 shows an I_aE_a curve for expression 2.1.

This expression neglects the effect of the initial velocity of the electrons, which produces more initial curvature of the I_aE_a characteristic than that indicated by the $\frac{3}{2}$ power relationship and causes anode current to flow even for negative values of E_a (curve 2 in Fig. 2.3). Another factor controlling the I_aE_a curve is contact potential; this is due to the materials forming the cathode and anode, and it causes a shift of the curve in a positive anode voltage direction (curve 3 in Fig. 2.3). The use to which the valve is put determines the relative importance of these effects. A diode valve is employed either for rectification or for detection. As a rectifier it is required to convert A.C. to D.C. power, and for high efficiency it must allow large currents to flow for small positive voltage differences between anode and cathode. The current should be zero for negative anode voltages, and special precautions to prevent emission from the anode must be taken during manufacture. It is clear that a rectifier diode needs a high saturation current, i.e., the cathode should be long, for the temperature of an oxide-coated cathode, unlike that of the tungsten filament valve, must be maintained within fairly narrow limits if a long life is desired. Too high

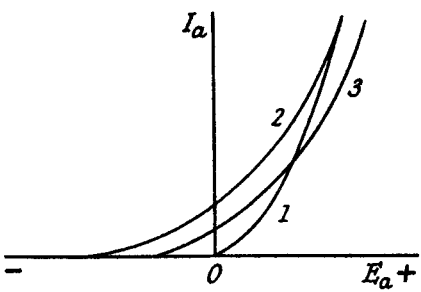


FIG. 2.3.—Effect of Initial Velocity and Contact Potential on the I_aE_a Curves of a Diode.

(Curve 1 : $I_a = kE_a^{\frac{3}{2}}$. Curve 2: Curve 1 with Initial Velocity Effect. Curve 3: Curve 2 with Contact Potential Effect.)

a temperature causes evaporation of the coating and loss of emission, and too low a value causes " poisoning " due to absorption of gas. The rate of rise of I_a with increasing E_a must be as steep as possible since this reduces the power loss in the valve and raises rectification efficiency. Hence from expression 2.1 we require a valve with a long cathode of large diameter, and minimum cathode-anode clearance. A small effective anode-cathode clearance is often obtained by including a grid (joined to the anode) close to the cathode. Since the A.C. peak voltages to be rectified are large, start of anode current at a small negative anode voltage and initial curvature of the I_aE_a characteristic at low anode voltages are not so important.

For detection purposes, linearity of I_aE_a characteristic is very necessary, and as the applied A.C. peak voltage will generally be small, the starting-point for anode current becomes important (Sections 8.2.6 and 7). On the other hand, detection current is small and a high saturation value is not essential.

2.3. The Triode. A triode valve has three electrodes, a grid being included between cathode and anode to control the flow of electrons. Its particular advantage is that small changes of grid voltage can produce large changes of anode current with negligible

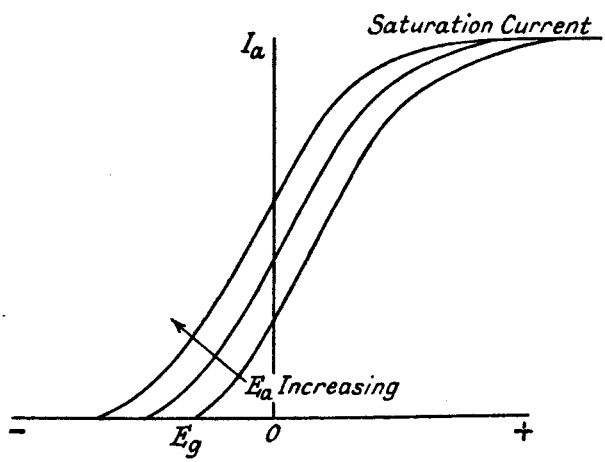


FIG. 2.4.—I_aE_a Curves for a Triode Valve.

expenditure of power in the grid circuit. The flow of electrons is now a result of the combination (at the cathode) of the electric fields from the grid and anode. Typical curves showing the effect of grid voltage variation on anode current are given in Fig. 2.4 for

different anode voltages. If the grid is made sufficiently negative it can completely neutralize the attractive power of the anode and stop the flow of electrons. As the negative grid voltage is reduced, I_a begins and increases until it finally reaches saturation at some positive value of grid voltage. When the grid voltage becomes positive (in mains valves with equipotential cathodes it may occur at a slightly negative value) electrons are collected by this electrode and grid current flows. The condition is rarely allowed to occur in practice because, as soon as grid current flows, power is taken from the driving source and this may cause distortion of the applied controlling voltage. Increase of E_a causes the I_aE_g curve to move to the left, but it does not appreciably affect the saturation current value.

The relationship between I_a, E_g and E_a can be more usefully expressed in the form of I_aE_a curves for different values of E_g, as shown in Fig. 2.11*a*. These curves are particularly helpful for indicating the performance of a valve in relation to its external anode circuit, and they also give three very important valve parameters, amplification factor (μ), mutual conductance (g_m), and internal resistance (R_a).

Amplification factor μ

$$= \frac{\text{change of anode voltage}}{\text{change of grid voltage to preserve constant anode current}}$$

$$= \frac{\Delta E_a}{\Delta E_g}(I_a = \text{constant}) \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.2.$$

Thus if the grid voltage lines in Fig. 2.11*a* are separated by 1 volt intervals, amplification factor at a particular value of I_a is given by the horizontal distance between the intercepts of a horizontal line through this I_a value with the particular grid voltage lines. Hence the amplification factor for $I_a = 3\ mA$ and E_g from -3 to -4 volts is

$$\mu = \frac{XX'}{1} = 30.$$

Mutual conductance g_m

$$= \frac{\text{change of anode current}}{\text{change of grid voltage for constant anode voltage}}$$

$$= \frac{\Delta I_a}{\Delta E_g}(E_a = \text{constant}) \text{ expressed in mA/volt} \quad . \quad . \quad 2.3.$$

Mutual conductance is obtained in Fig. 2.11*a* by drawing a vertical line through a given E_a and measuring the vertical distance (in

milliamps) between any given E_g intercepts. Thus g_m at $E_a = 200$, E_g from -3 to -4 is

$$g_m = \frac{YY'}{1} = 3 \text{ mA/volt.}$$

Internal or Slope Resistance R_a

$$= \frac{\text{change of anode voltage}}{\text{change of anode current for a constant grid voltage}}$$

$$= \frac{\Delta E_a}{\Delta I_a}(E_g = \text{constant}) \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.4.$$

This is measured by the inverse slope of the tangent to a given E_g curve at a particular value of E_a or I_a; R_a at $E_a = 200$, $E_g = -3$ is the cotangent of the angle that the tangent to the $E_g = -3$ volt curve at Y makes with the E_a axis,

$$\text{i.e., } R_a = 10{,}000\Omega.$$

By combining expressions 2.2, 3 and 4 we note that

$$\mu = \frac{\Delta E_a}{\Delta E_g} = \frac{\Delta E_a}{\Delta I_a} \times \frac{\Delta I_a}{\Delta E_g} = R_a.g_m$$

A point worth observing is that μ is a good guide to the grid base, i.e., the negative voltage required to stop anode current for

$$E_g(\text{cut-off}) \simeq \frac{E_a}{\mu}.$$

Actually E_g (cut-off) tends to be slightly greater than this because of the initial curvature of the I_aE_a characteristics.

The electrode construction of the valve determines the values of the three parameters. From Fig. 2.11*a* we see that decrease of the slope (increase of R_a) of the I_aE_a lines for the same perpendicular separation must increase XX' and hence increase μ. This can be achieved by reducing the effect of E_a on I_a, that is, by increasing the shielding power of the grid. Increase of the turns per inch, the thickness of the wire, or the distance between the grid and anode all increase μ.

Mutual conductance is raised by increasing the emitting surface and the control action of the grid. The former is realized by having a long cathode, and the latter by reducing the grid-cathode clearance. Increasing the turns per inch of the grid wire also tends to raise g_m.

The chief characteristics of a triode valve are that over the normal working ranges of I_a, E_a and E_g, μ, g_m and R_a are almost constant and only change as cut-off grid bias is approached, when μ and g_m fall and R_a rises.

2.4. The Tetrode. Between the control grid and anode, the tetrode valve has an additional grid positively biased with respect to the cathode. Two advantages are obtained; the anode-grid capacitance is very much reduced — 0·005 $\mu\mu$F in an R.F. tetrode as compared with about 3 $\mu\mu$F for a triode—so that coupling between anode and grid circuits is decreased, and the internal slope resistance becomes very high (not less than 0·5 MΩ). Both these effects are to be expected because the additional electrode screens the anode from the grid and the cathode, and very much reduces the influence of changes of anode voltage on the anode current.

The purpose for which the tetrode is to be used determines the details of its electrode construction. In an A.F. output stage, a low anode-grid capacitance and high resistance are not so essential and a coarse mesh screen allowing high g_m is desirable. For detection and amplification, particularly at radio frequencies, low anode-grid capacitance and high R_a are essential because the frequency response of an amplifier with tuned anode and grid circuits can be appreciably modified by regeneration and degeneration produced by the undesired anode-grid coupling. Furthermore, a low R_a reduces the selectivity of the anode tuned circuit. Metal skirts are often included at the top and bottom of the screen in order to reduce the stray capacitance between grid and anode.

R_a and g_m are the all-important parameters in a tetrode valve. The former is determined by the mesh of the screen wires (a close mesh screening the anode from the electron reservoir round the cathode), and the clearance between anode and screen (a large clearance producing a high R_a). Secondary emission from the screen to anode reduces R_a. A compromise is essential as screen current usually increases as the mesh becomes finer because a greater area of the electron stream is covered by the screen wires, and a large screen-anode clearance necessitates a large glass envelope. Mutual conductance is governed largely by the dimensions of the cathode, grid and screen and the anode has little effect.

The I_aE_g characteristics of a tetrode are similar to those of Fig. 2.4, except that E_s takes the place of E_a. The I_aE_a characteristics of the earliest form of tetrode, a screened grid, are illustrated in Fig. 2.5. Change of grid voltage alters the magnitude but not the general shape of the curves, and increase of screen voltage increases screen and anode current and also moves the point of minimum I_a to a higher E_a. Taking a particular curve, we see that for small increasing positive values of E_a, I_a rises quite steeply until at about +10 volts the velocity of the electrons striking the

anode is sufficient to detach electrons from the anode material. These low velocity secondary electrons are attracted to, and collected by, the screen because of its higher positive voltage. The secondary emission from the anode causes the anode current

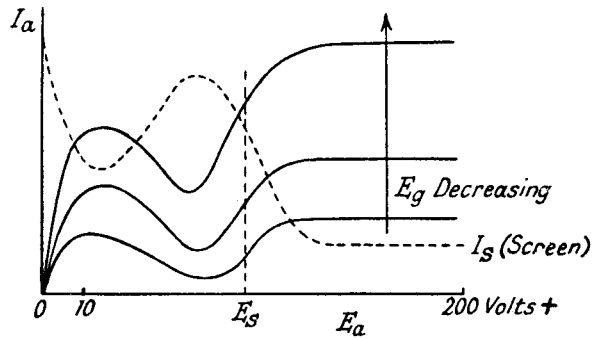


FIG. 2.5.—I_aE_a Curves for a Screened Grid Valve.

to fall and screen current to rise. Under suitable conditions the secondary electrons may exceed the primaries and I_a become negative. In the secondary emission region of falling I_a with increasing E_a, the anode presents a negative resistance to any external circuit and the valve can be made to act as an oscillator (known as a dynatron oscillator). When E_a approaches E_s, the attractive power of the screen lessens and the negative charge due to the electrons in the anode-screen space acts as a repulsive force driving the secondary electrons back to the anode. I_a therefore rises and I_s falls. As E_a is increased above E_s, secondary emission from the anode ceases and the anode collects all the primary electrons not intercepted by the screen wires. Thereafter I_a remains practically constant except for a small rise due to slight penetration of the anode field to the cathode, and to secondary emission from the screen to the anode. This secondary emission from the screen is much less effective than that from the anode since the secondaries are produced on the side away from the anode. The amount of secondary emission is dependent on the condition of the anode and screen surfaces, and if Barium and Strontium, volatilized from the cathode, are deposited on these the number of secondary electrons may be large. The emission from the anode can be greatly reduced by carbonising the surface.

Anode secondary emission is an undesirable characteristic in the tetrode because it limits the minimum working anode voltage to a little more than the screen voltage, i.e., it limits the maximum

output voltage from the valve. Its suppression can be achieved in a number of ways, all of which depend on the production of a voltage minimum in the anode and screen space, as shown in the curve of voltage distribution between the cathode and anode

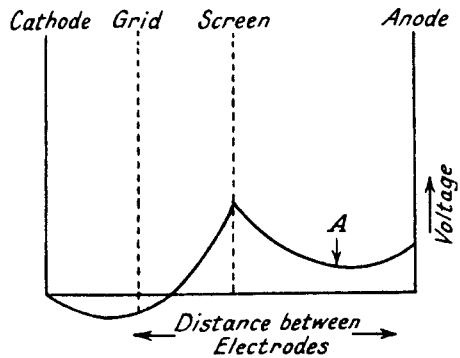


FIG. 2.6.—Voltage Distribution in a Tetrode Valve showing Suppression of Secondary Emission by a Voltage Minimum between Anode and Screen.

(Fig. 2.6). The secondary electrons are unable to pass this barrier (*A* in the figure) and so return to the anode. One of the simplest methods is to include an open-mesh grid (connected to the cathode) between the anode and screen. The pentode valve so formed has the I_aE_a characteristics of Fig. 2.7; the limitation of output voltage is now removed and an output voltage approaching 80% to 90% of the H.T. voltage can be obtained. The mesh of this grid

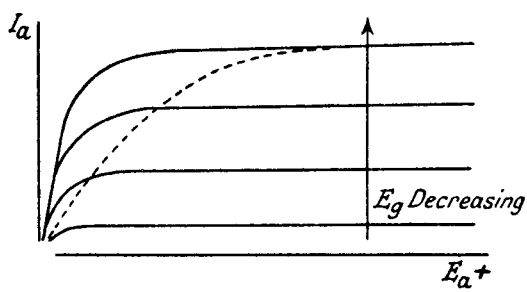


FIG. 2.7.—I_aE_a Curves for a Pentode Valve.

must be such as only just to suppress secondary emission, because if the mesh is made closer the flow of primary electrons to the anode at low E_a voltages is reduced, producing the dotted I_aE_a characteristic, which is demonstrably less efficient.

The potential minimum can be realized by fins [5] on the anode at right angles to the surface facing the screen, and also by increasing the anode-screen clearance.[4] In the latter instance the primary electrons in the anode-screen space provide the potential minimum by reason of their negative space charge. In a tetrode most of the anode secondary emission occurs in the region round the screen support wires where the primary electronic density is low. Plates, connected to the cathode and placed between the screen supports and the anode, return the secondary electrons to the anode and also tend to concentrate the primary electron stream into a beam, so producing the required potential minimum by the negative charge of the concentrated electrons, and preventing the return of secondaries in the main electron stream. This is the principle of secondary emission suppression employed in the beam tetrode, which often has the screen and control grids aligned so that the latter shades the former and reduces the pick-up of primary electrons by the screen.

In tetrodes required for R.F. amplification some form of gain control by variation of the D.C. electrode voltages is essential, and this is normally obtained by varying the grid bias of the control grid. This control would, without special grid construction, be much too rapid (curve 1 in Fig. 2.8) and would introduce serious distortion ; it is necessary to obtain a steady change of mutual conductance with grid bias as shown by curve 2 in Fig. 2.8. Curve 1 shows the variation of g_m against E_g for a normal grid construction of constant mesh. The curved characteristic 2 can be achieved by winding a grid of constantly variable pitch. The valve so produced is equivalent to a number of valves connected in parallel, each one having a different $g_m E_g$ characteristic, the close mesh part of the grid contributing mainly to the high g_m section (curve 3) and having a low negative cut-off voltage, and the open mesh giving a

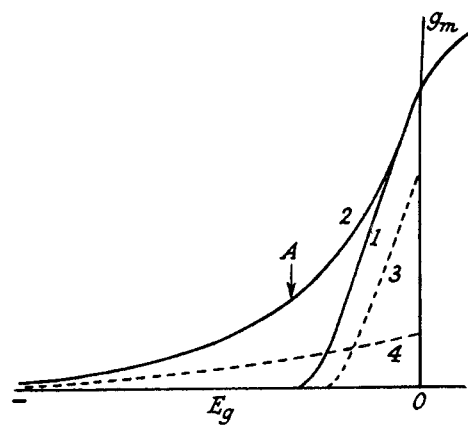


FIG. 2.8.—$g_m E_g$ Curves for a Tetrode or Pentode Valve.
(Curve 1 : Non-Variable *mu* Characteristic. Curve 2 : Variable *mu* Characteristic.)

low $g_m E_g$ curve (curve 4) with a high negative cut-off voltage. Generally, the mesh is fine at the ends of the grid electrode and coarse in the middle. Some distortion of the input voltage wave is inevitable with the variable *mu* (as it is called) valve, and it is a maximum at the point *A* (curve 2) of greatest curvature. A sudden change of mesh produces a sharp bend at this point, whilst a graded change of mesh produces a much smoother change.

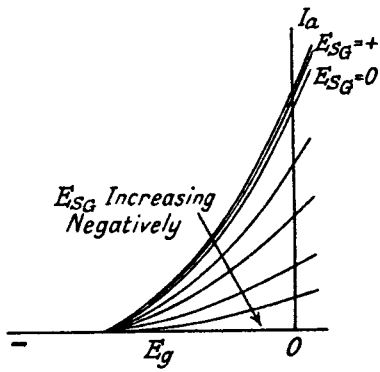


FIG. 2.9*a*.—$I_a E_g$ Curves of a Pentode Valve for Different Suppressor Grid Bias Voltages.

This variable *mu* or, more correctly, variable g_m effect can occur unintentionally in a non-variable *mu* valve if the grid electrode is out of line with the cathode or is conical in shape.

2.5. The Multi - electrode Valve. The multi-electrode valve, such as the pentagrid or heptode, octode, and hexode, is mainly used as a frequency changer. For an understanding of its action it is helpful to study the effect of negative bias on the suppressor grid of a pentode. Typical $I_a E_g$ curves showing the effect of negative bias on this grid are illustrated in Fig. 2.9*a*. As the bias increases, the electrons are repelled and the slope, g_m, of the characteristic decreases, and therefore the gain of the valve falls. Applying positive bias to the suppressor grid has little effect since the screen limits the flow of electrons from the cathode. As variation of suppressor grid negative bias controls gain, the valve may be used either as a modulator with the carrier applied to the control grid and modulation to the suppressor grid, or as a frequency changer with the signal applied to the control grid and the local oscillator output to the suppressor grid. There is, however, a disadvantage to the method and it can be seen from Fig. 2.9*b*. Increasing suppressor grid bias lowers R_a and eventually turns the $I_a E_a$ characteristic into that of a triode. This is due to the formation

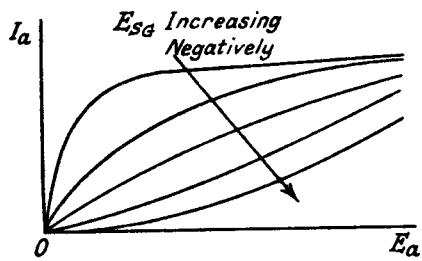


FIG. 2.9*b*.—$I_a E_a$ Curves for a Pentode Valve with Different Suppressor Grid Bias Voltages.

of a reservoir of electrons between the screen and suppressor grid ; the field from the anode can draw upon this virtual cathode and influence considerably the anode current. The reduction in R_a is a serious defect, particularly in a frequency changer, because it damps any highly selective anode circuit.

The difficulty can be overcome by converting the triode, formed by the virtual cathode, suppressor grid and anode, into a screened grid by interposing a positively biased grid between the suppressor grid and anode. Thus we have the hexode frequency changer of Fig. 2.10. The I_aE_{g1} characteristics for this valve are similar to those of Fig. 2.9*a* and the I_aE_{g3} characteristics for different values of E_{g1} are shown in Fig. 2.10. There is little increase in I_a for

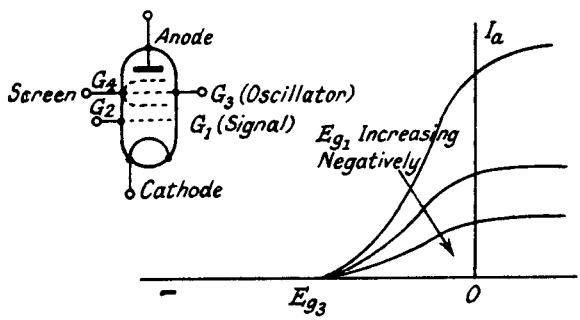


FIG. 2.10.—I_aE_{g3} Curves for a Hexode Valve with Different Signal Grid Bias Voltages.

positive values of E_{g3} since the first screen shields grid 3 from the cathode electron reservoir, and all available electrons between grids 2 and 4 have been collected by the anode. The I_aE_a characteristics are similar to those for a screened grid valve though the secondary emission " kink " is generally much less pronounced than that of Fig. 2.5.

The position of the signal and local oscillator grids may be interchanged, that is, the oscillator may be next to the cathode, and this occurs in pentagrid, heptode, and octode valve. In the heptode and octode valve another positively biased electrode is inserted between the oscillator grid and first screen to serve as an oscillator anode, and the valve acts as a combined oscillator and frequency changer. The octode has an extra suppressor grid between the anode and second screen to produce pentode I_aE_a characteristics. Details of these special valves are given in Section 5.3.

2.6. Representation of the External Anode Load Impedance on the I_aE_a Characteristic Curves. To illustrate the method of representing the external anode circuit on the I_aE_a curves we will consider a triode valve having the I_aE_a characteristic curves of Fig. 2.11*a*, with an external resistance R_0 of 10,000 ohms and an H.T. voltage $E_b = 200$ volts. This resistance is represented by a straight line AB passing through the points $I_a = 0$, $E_a = 200$

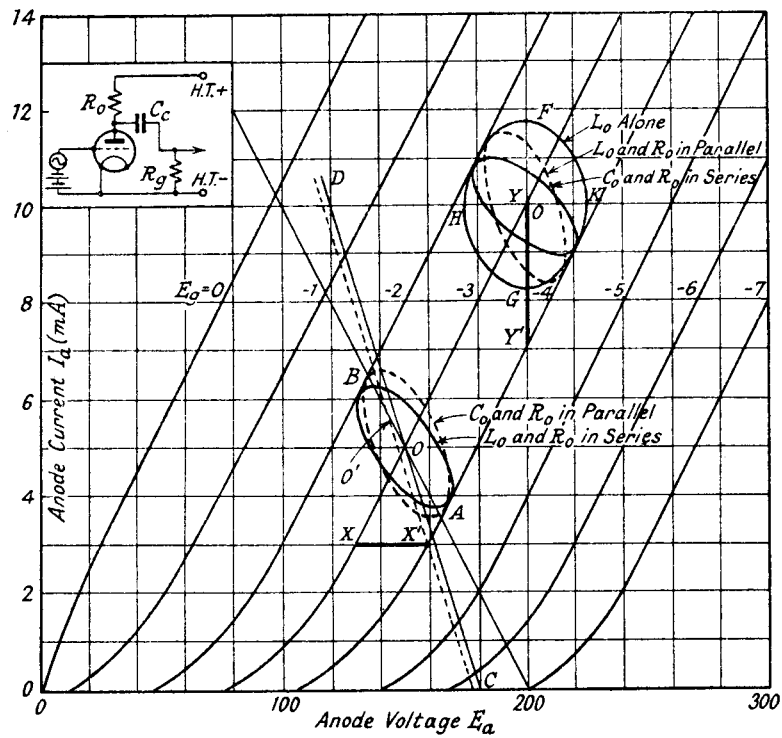


FIG. 2.11*a*.—Load Lines and Curves on the I_aE_a Characteristics.

(the anode voltage must equal E_b when $I_a = 0$ because no voltage drop occurs across R_0) and $I_a = 2$ *mA*, $E_a = 200 - I_aR_0 = 180$ volts. The resistance load line is straight because its equation is represented by

$$I_a = \frac{E_b - E_a}{R_0},$$

which shows that I_a is linearly proportional to E_a. We can use this line to determine the D.C. anode voltage and A.C. output voltage for any given D.C. grid bias and A.C. grid voltage. If the grid bias

is -3 volts and a sinusoidal A.C. voltage of peak value 1 volt is applied to the valve, the D.C. anode voltage is given by the intercept 0 of the R_0 line and the $E_g = -3$ volts curve measured on the E_a axis, i.e., it equals 150 volts. The A.C. anode peak voltage is 15, half the horizontal distance between the intercepts B and A of the R_0 line and the $E_g = -2$ and -4 volts curves. When a valve with a resistance load is used as a voltage amplifier the output from its anode circuit is coupled to the grid of the next valve through a capacitance C_c so as to isolate the D.C. voltage and allow only the A.C. component to be passed on. A grid leak R_g completes the D.C. path back to the cathode. The value of C_c is chosen so that at the lowest frequency its reactance is very much less than R_g. The effect of R_g is therefore to make the load resistance to A.C. of the first valve less than the D.C., i.e., the slope of the load line is $\frac{R_g R_0}{R_g + R_0}$, and it is shown on Fig. 2.11*a* as the line CD, which passes through the point O on the D.C. line. The A.C. voltage amplification is therefore reduced from

$$\frac{\mu R_0}{R_a + R_0} \text{ to } \frac{\mu \frac{R_g R_0}{R_g + R_0}}{R_a + \frac{R_g R_0}{R_g + R_0}},$$

but this is of much less importance than the fact that the bottom end of the A.C. line enters the cramped part of the $I_a E_a$ curves and distortion is produced with large outputs. R_g should therefore be chosen to have as high a value as possible so long as it does not produce softness in the succeeding valve. An interesting point to note is that the A.C. load line moves its position for large output voltages because rectification, and hence increased D.C. current, results from distortion. The line CD tends to move to a point O' corresponding to lower D.C. anode voltage in a manner similar to that described in Section 8.4.1. for anode bend detection.

To explain the construction of the locus load line for any form of load the simplest procedure is to start with the resistance case, dissociating the resistance locus load line from the $I_a E_a$ curves and drawing it as in Fig. 2.11*b* with reference to A.C. output voltage and current. To allow reference back to Fig. 2.11*a*, we will make the positive direction of output voltage to the left, because increasing positive load voltage means decreasing anode voltage E_a. The point O on the line AB in Fig. 2.11*a* becomes the origin for the line in Fig. 2.11*b*. Instantaneous values of E_0 and I_0 can be obtained

by projecting horizontally and vertically on to the two axes, for the voltage and current are in phase. Thus the A.C. cycle commences at O; 30° later E_0 and I_0 have values corresponding to OC and OD, which are half the maximum values of E_0 and I_0, OH and OP (at the 90° position of the cycle), for $\sin 30° = 0{\cdot}5$.

Now let us imagine that an inductive reactance of 15,000 ohms is connected in place of R_0 and that the peak voltage is arbitrarily

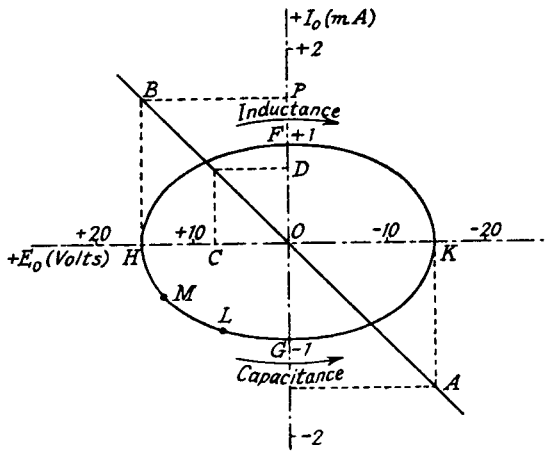


FIG. 2.11b.—Representation of the Locus Load Line for R_0, L_0, and C_0 Anode Loads.

chosen as 15, viz. OH, the same as that across R_0. Since the anode load is an inductive reactance, E_0 and I_0 are out of phase by 90°, I_0 lagging behind E_0. When E_0 is zero, I_0 has its maximum negative value of $\frac{-E_0}{X_0} = \frac{-15}{15{,}000} = -1$ mA; hence G is a point on the locus. 30° later E_0 has a value $15 \sin 30° = 7{\cdot}5$, and $I_0 = -1 \cos 30° = -0{\cdot}866$ mA. Therefore L is a point on the locus. Similarly after 60° of the cycle $E_0 = 13$ and $I_0 = -0{\cdot}5$, and point M is found. When $E_0 = 15$, its maximum value, $I_0 = 0$. If this procedure is repeated over the cycle the ellipse $GHFK$ is drawn out and this is the locus load line for an inductance L_0. The direction of progressive time instants is clockwise round the ellipse. A similar ellipse can be obtained for a capacitance C_0, but the point corresponding to progressive time instants now travels in an anti-clockwise direction because I_0 leads on E_0. The length of the I_0 axis of the ellipse is inversely proportional to the reactance X_0, being short for a large value of X_0.

In drawing out the reactive ellipse we assumed, for convenience,

that the output voltage was E_0, the same as that across R_0, but in actual fact the output voltage is dependent on the valve slope resistance R_a and on X_0. The shape of the locus curve is not affected, however, by the value of E_0 since I_0 is proportional to E_0, and a change of the former only means an increase or decrease in the size of the ellipse. In transferring the reactive ellipse to the I_aE_a curves we must adjust E_0 so that the ellipse is just tangential to the limit grid voltage curves, in this case $E_g = -2$ and -4 volts. Thus, if we have an H.T. voltage of 200 volts, a bias of -3 volts, an A.C. peak grid voltage of 1, and an inductance L_0 of reactance $15{,}000\Omega$ and no D.C. resistance, the load curve is the ellipse $GHFK$ drawn on Fig. 2.11*a* with its centre O at the point $I_a = 10$ mA, $E_a = 200$ volts.

The size of the ellipse is adjusted to make it tangential to the $E_g = -2$ and -4 volt lines, and we find that E_0 has been increased to approximately 26 volts as compared with 15 volts arbitrarily chosen for constructing the ellipse in Fig. 2.11*b*. The major axis of the ellipse is vertical in Fig. 2.11*a* instead of horizontal as in Fig. 2.11*b*, because the scale relationship of I_a to E_a is changed.

Let us now consider the practical condition of a combined resistive and reactive load. Taking first a series circuit of resistance R_0 ($10{,}000\Omega$) and reactance ωL_0 ($15{,}000\Omega$), let us assume some arbitrary value of maximum current such as 1 mA, OF in Fig. 2.11*c*. We begin with current because a series circuit is being considered. The maximum voltage E_1 across R_0 is $IR_0 = 10$ volts (FB in Fig. 2.11*c*) and the maximum E_2 across L_0 is $IX_0 = 15$ volts (OH in the figure). By following the procedure outlined above we can draw the resistance line AB and the reactance ellipse $GHFK$. Since R_0 and L_0 are in series, their instantaneous voltages must be added for the combined locus load line. Thus for maximum $I_0 = OF$, the voltages E_1 and E_2 across R_0 and L_0 are FB and zero, and B is therefore a point on the combined locus. Similarly for $I_0 = 0$, the voltages across R_0 and L_0 are zero and OH respectively, and H is another point. For I_0 rising positively from zero, at a value $I_0 = OM$ we have $E_1 = MN$ and $E_2 = MP$ respectively, giving point Q. When I_0 is decreasing towards zero, for $I_0 = OM$ we have the total voltage $= MP - MN$ and point R is found. Continuing this process produces the locus curve $HQBKA$ similar in shape to a sheared ellipse, the direction of progressive time instants being clockwise. This curve must now be magnified or reduced for transferring to the I_aE_a curve in Fig. 2.11*a*. The resistance R_0, which will be assumed to have a D.C. resistance of

10,000Ω, and E_b (200 volts) determines the position of the centre point O of the "sheared" ellipse, and its overall size is adjusted until it is tangential to the grid voltage lines $E_g = -2$ and -4 volts as shown in Fig. 2.11*a*. A similar combined locus curve is obtained

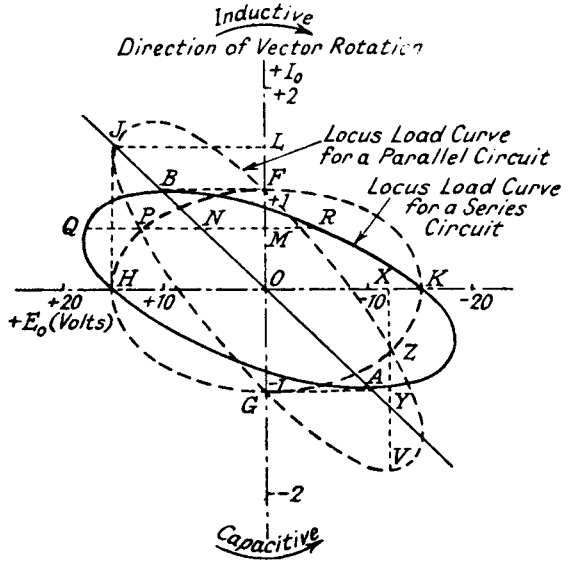


FIG. 2.11c.—Locus Load Curves for Series or Parallel Circuits of Resistance and Reactance.

for a series combination of R_0 and C_0, but the direction of vector rotation is anti-clockwise. We cannot transfer the series R_0C_0 circuit load curve directly to Fig. 2.11*a* because it has no D.C. path for the anode current. If, however, we assume the D.C. component to be carried by a choke of very high inductive reactance and zero D.C resistance, the combined locus is centred immediately above E_b (200 volts) at the intersection of $E_b = 200$ volts and $E_g = -3$ volts.

When the anode load is a parallel circuit of R_0 and L_0, we must take some arbitrary value of voltage E_0, OH in Fig. 2.11*c*, and draw the straight line for R_0 and ellipse for L_0. The maximum value of current I_1 through R_0 is $\frac{E_0}{R_0}$, OL in the figure, and the maximum value of current I_2 through L_0 is $\frac{E_0}{X_0}$, OF in the figure. Since the circuit is a parallel one we must add currents. When E_0 is zero, I_1 = zero and $I_2 = OG$, so that G is a point on the combined locus; for $E_0 = OH$ (its maximum positive value) $I_1 = OL$ but I_2 = zero,

hence J is a point on the combined locus. When E_0 is negative, equal to OX and decreasing to zero, $I_1 = XY$, $I_2 = XZ$ and V is a point on the combined locus where $XV = XY + XZ$. Continuing the process gives the " sheared " elliptical shape, $JFVG$, and this is increased or decreased in size, so that when transferred to Fig. 2.11a it becomes tangential to the $E_g = -2$ and -4 volt lines. The load curve for R_0 and C_0 in parallel is similar, with the direction of vector rotation in an anti-clockwise direction. The D.C. resistance component decides the position of the load curve above the E_a axis on Fig. 2.11a. The parallel R_0L_0 load curve is centred above the H.T. voltage E_b if L_0 has a negligible D.C. resistance, and the R_0C_0 curve about the intersection of the R_0 line with the $E_g = -3$ volt line if the D.C. resistance of R_0 is the same as its A.C. resistance.

2.7. Equivalent Circuits for a Valve. A valve may be represented by a constant voltage or a constant current generator [1]; the former representation is most suitable for a triode valve, which has a comparatively low resistance, whilst the latter is more useful in analysing the action of a tetrode or high resistance valve. The two circuits are shown in Figs. 2.12a and 12b; coupling between the anode and grid circuits has been assumed to be negligible.

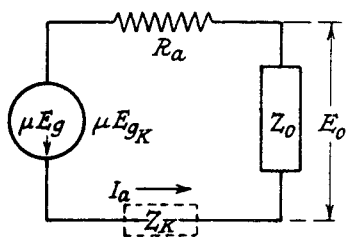


FIG. 2.12a.—The Valve as a Constant Voltage Generator.

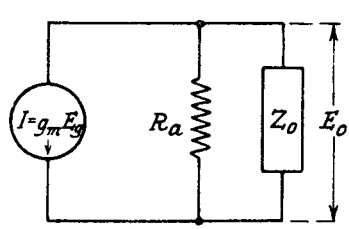


FIG. 2.12b.—The Valve as a Constant Current Generator.

The voltage developed across the external load Z_0 in Fig. 2.12a is

$$E_0 = \frac{\mu E_g Z_0}{R_a + Z_0} \quad . \quad . \quad . \quad . \quad 2.5a$$

whilst in Fig. 2.12b.

$$E_0 = \frac{IR_aZ_0}{R_a+Z_0} = \frac{g_mE_gR_aZ_0}{R_a+Z_0} = \frac{\mu E_gZ_0}{R_a+Z_0} \quad . \quad . \quad 2.5b.$$

The two circuits are therefore identical. Fig. 2.12b is very helpful when Z_0 is a tuned circuit, for it shows the desirability of having R_a as large as possible in order that the overall gain may be high and the maximum selectivity of Z_0 realized.

Sometimes an impedance is connected in the cathode-earth lead of a valve, notably for providing negative feedback ; the circuit of Fig. 2.12*a* is modified by the inclusion of the dotted impedance Z_k, and the effective grid input voltage becomes that between grid and cathode, E_{g_k}. The output voltage is

$$E_0 = \frac{\mu E_{gk} Z_0}{R_a + Z_0 + Z_k}$$

where $$E_{g_k} = E_g - E_k$$

but $$E_k = \frac{\mu E_{g_k} Z_k}{R_a + Z_0 + Z_k}$$

$$\therefore E_{g_k} = E_g - \frac{\mu E_{g_k} Z_k}{R_a + Z_0 + Z_k}$$

$$= \frac{E_g(R_a + Z_0 + Z_k)}{R_a + Z_0 + (\mu + 1)Z_k}$$

$$\therefore E_0 = \frac{\mu E_g Z_0}{R_a + Z_0 + (\mu + 1)Z_k} \quad . \quad . \quad . \quad 2.5c.$$

Hence the inclusion of the cathode impedance has had the effect of increasing the equivalent valve impedance by $(\mu+1)Z_k$. For high slope resistance valves, such as screened-grids or pentodes, g_m is the important parameter and expression 2.5*c* is modified as follows :

$$E_0 = \frac{g_m R_a E_g Z_0}{R_a + R_0 + (\mu + 1)Z_k} = \frac{g_m E_g Z_0}{1 + \frac{R_0}{R_a} + \left(g_m + \frac{1}{R_a}\right)Z_k} \simeq \frac{g_m E_g Z_0}{1 + g_m Z_k} \quad . \quad 2.5d$$

when $R_a \gg Z_0$. Thus the equivalent mutual conductance of the valve is reduced in the ratio $\frac{1}{1+g_m Z_k}$. Expression 2.5*d* assumes that the screen is decoupled to cathode by a large capacitance and that there are no voltage changes across Z_k due to change of screen current. When the screen is decoupled to earth, a voltage is developed across Z_k due to screen current changes, and the following current voltage equations result.

$$E_k = (I_a + I_s)Z_k$$

$$I_a = \frac{\mu E_{g_k}}{R_a + Z_0 + Z_k} \simeq g_m E_{g_k}$$

$$I_s = \frac{\mu_s E_{g_k}}{R_s + Z_0 + Z_k}$$

where $$\mu_s = \frac{\Delta E_s}{\Delta E_g} (I_s \text{ constant}).$$

R_s = slope resistance of the $I_s E_s$ curves.

Generally $R_s \gg Z_k$ and

$$I_s = g_s E_{g_k}$$

where $$g_s = \frac{\Delta I_s}{\Delta E_g} (E_s \text{ constant}).$$

$$\therefore E_k = (g_m + g_s) Z_k E_{g_k}.$$

But $$E_g = E_{g_k} + E_k = [1 + (g_m + g_s) Z_k] E_{g_k}.$$

$$\therefore E_0 = g_m E_{g_k} = \frac{g_m E_g}{[1 + (g_m + g_s) Z_k]} \quad . \quad . \quad 2.5e.$$

$$= \frac{g_m E_g}{(1 + g_k Z_k)}.$$

Usually $g_m \propto I_a$ and $g_s \propto I_s$ so that

$$g_k = g_m \left(1 + \frac{I_s}{I_a}\right).$$

In non-aligned grid tetrode valves $\frac{I_s}{I_a}$ is about 0·25 and in aligned grid valves it is about 0·1.

2.8. The Grid Input Admittance of a Valve.[8]

2.8.1. Introduction. The grid input circuit of a valve can usually be represented by a resistance and capacitance in parallel. The magnitude of both components depends on the construction of the valve, the potentials applied to the electrodes, and the impedances in the external circuits connected to the electrodes.

The input resistance can be due to four causes:

(1) Leakage current between the grid and other electrodes. This can be measured when the valve heater or filament is cold, but it may change when it is hot.

(2) Electronic and positive ion current. The first may be produced by grid collection or emission of electrons. Collection of electrons by the grid may generally be prevented by applying sufficient negative bias. Emission from the grid is usually due to the volatilization of active material from the cathode and its condensation on the cooler grid. It can be reduced by preventing cathode temperatures in excess of that required for satisfactory operation, and by maintaining the temperature of grid and anode at as low a value as possible. A high temperature for the grid

electrode encourages emission from any material which has condensed upon it. Cooling fins are often connected to the top of the grid support wires of h gh current valves, and the anode is carbonized. Grid emission causes current in the opposite direction to that from electron collection, and produces across a grid leak resistance a positive bias on the grid. Positive ion current has this same effect and is due to the collection of positive ions—positively charged atoms of residual gas, which have been robbed of electrons by collisions with the primary electrons—by the negatively biased grid. Grid emission and positive ion current are more likely to occur in high current output valves, and the effect may be cumulative if a high grid leak resistance is employed. The positive bias caused by this grid current increases the anode current, thus raising the temperature of the anode, causing it to release absorbed gases; at the same time the grid temperature increases because of greater radiation from the anode, and grid emission current increases. A high grid leak may therefore quickly produce destruction of the vacuum and "softening" of the valve, and it is the probability of "softness" which limits the grid leak in output valves and most types of voltage amplifier valves to maxima of 0·5 and 2 MΩ respectively.

(3) Coupling between the grid and any other electrode containing an impedance to the input frequency. Anode-grid capacitance is the more usual source of this defect, but capacitance between the grid and cathode or screen can also contribute. The equivalent input resistance from this cause may be positive or negative, resulting in degeneration or regeneration of the input voltage, the sign depending on the sign of the particular electrode circuit reactance. The magnitude of this resistance is inversely proportional to the gain of the valve and directly proportional to the ratio of resistance to reactance in the given electrode circuit, zero reactance giving infinite resistance.

(4) Transit time of the electrons between the cathode and grid. When the time of flight of the electrons between grid and cathode becomes comparable with the reciprocal of the input frequency, electrons may be accelerated away from the grid in the direction of the cathode as well as in the direction of the screen of the R.F. amplifier valve. The negatively charged cloud of electrons moving away from both sides of the grid constitutes an A.C. current having a resistive and a capacitive component. The effect is dependent on the total electron flow and is greatest for high space currents, i.e., low grid biases.

The input capacitance is also dependent on four factors :

(1) The grid size and proximity to other earthed electrodes such as the cathode and screen. This capacitance can be measured when the valve is cold.

(2) Space charge. The negative charge due to the electrons clustered round the cathode is equivalent to a reduced distance between the cathode and grid, and the capacitance between the two electrodes is therefore increased upon that of the cold value. The increase is dependent on the grid bias and has a maximum of about 2 $\mu\mu$F at the minimum negative grid bias. It causes mistuning of the signal circuit when the gain of the valve is varied by changing the signal grid bias.

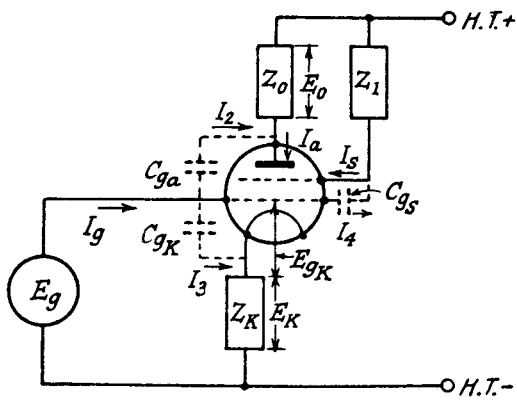


FIG. 2.13.—A Circuit showing Coupling to the Grid by Interelectrode Capacitance.

(3) Coupling between the grid and any other electrode, having an impedance to the operating frequency. This reflected reactance is always capacitive irrespective of the sign of the reactance in the other electrode, and the equivalent capacitance is a maximum for zero external circuit reactance and minimum grid bias, i.e., maximum valve gain.

(4) Transit time of the electrons as set out above.

We will now consider the resistance and capacitance reflected into the grid circuit due to coupling between it and other electrodes. A circuit showing coupling between the grid and other electrodes is given in Fig. 2.13.

2.8.2. Grid Input Admittance and Anode-Grid Capacitance Coupling. Assuming that $Z_k = Z_l = C_{g_k} = C_{g_a} = 0$), the equivalent circuit for a valve having anode-grid capacitance coupling is shown in Fig. 2.14. Using the constant current generator

circuit and replacing the resistances and reactances by their equivalent conductances and susceptances, viz.

$$\frac{1}{R_a} \text{ by } G_a, \quad \frac{1}{Z_0} \text{ by } Y_0 = G_0 + jB_0 \text{ and } \frac{1}{X_{C_{g_a}}} \text{ by } B_{g_a},$$

the voltage and current relationships may be written as follows:

$$I_a = g_m E_g = E_0 \cdot [G_a + G_0 + jB_0] + [E_0 + E_g] \cdot jB_{g_a} \quad . \quad 2.6.$$

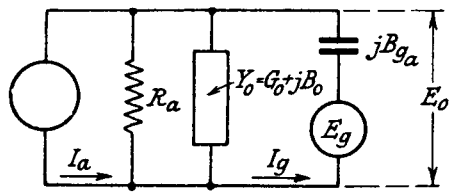


FIG. 2.14.—The Equivalent Constant Current ($I_a = g_m E_g$) Generator Circuit for Anode-Grid Capacitance Coupling.

Note.—E_0 and E_g are in the same direction as regards C_{g_a} and therefore additive.

$$I_g = (E_0 + E_g) jB_{g_a} \quad . \quad . \quad . \quad . \quad 2.7.$$

From 2.7 the admittance of the grid circuit is

$$Y_g = \frac{I_g}{E_g} = \left[1 + \frac{E_0}{E_g}\right] \cdot jB_{g_a}$$

and from 2.6

$$\frac{E_0}{E_g} = \frac{g_m - jB_{g_a}}{(G_a + G_0) + j(B_0 + B_{g_a})} \quad . \quad . \quad . \quad 2.8$$

$$\therefore \; Y_g = \frac{jB_{g_a}[G_a + G_0 + g_m + jB_0]}{(G_a + G_0) + j(B_0 + B_{g_a})}.$$

Rationalizing

$$Y_g = \frac{jB_{g_a}[G_a + G_0 + g_m + jB_0][G_a + G_0 - j(B_0 + B_{g_a})]}{(G_a + G_0)^2 + (B_0 + B_{g_a})^2}$$

$$= \frac{B_{g_a} \cdot [[G_a + G_0 + g_m][B_0 + B_{g_a}] - B_0[G_a + G_0]] + jB_{g_a}[(G_a + G_0 + g_m)(G_a + G_0) + B_0(B_0 + B_{g_a})]}{(G_a + G_0)^2 + (B_0 + B_{g_a})^2}$$

$$= G_g + jB_g$$

$$\therefore \; R_g = \frac{1}{G_g} = \frac{(G_a + G_0)^2 + (B_0 + B_{g_a})^2}{B_{g_a}[g_m \cdot B_0 + B_{g_a}(G_a + G_0 + g_m)]} \quad . \quad . \quad 2.9a$$

$$\text{and } C_g = \frac{B_g}{\omega} = \frac{C_{g_a}[(G_a + G_0 + g_m)(G_a + G_0) + B_0(B_0 + B_{g_a})]}{(G_a + G_0)^2 + (B_0 + B_{g_a})^2} \quad . \quad 2.9b.$$

These expressions show that when B_0 is positive or negative and infinite, R_g is infinite and $C_g = C_{g_a}$. This is to be expected since $B_0 = \infty$ means zero anode reactance and the anode is virtually connected to earth, leaving only the anode-grid capacitance across the grid input circuit. R_g is also infinite when $B_0 = -B_{g_a}\frac{(G_a+G_0+g_m)}{g_m}$. For values of $B_0 > -B_{g_a}\frac{(G_a+G_0+g_m)}{g_m}$, R_g is positive and for $B_0 < -B_{g_a}\frac{(G_a+G_0+g_m)}{g_m}$, R_g is negative. Thus anode-grid capacitive coupling with an inductive anode load can, and usually does, provide regeneration, whereas with a capacitive anode load, degeneration always occurs. There are two minima for R_g, one positive and the other negative, and they are found by differentiating 2.9*a*, with respect to B_0 and equating to 0. The resulting equation is

$$(B_0+B_{g_a})^2 g_m + 2B_{g_a}(B_0+B_{g_a})(G_a+G_0) - g_m(G_a+G_0)^2 = 0.$$

$$\text{Hence } B_0' = -B_{g_a} \pm (G_a+G_0)\left[\frac{B_{g_a}}{g_m} \pm \sqrt{1+\left(\frac{B_{g_a}}{g_m}\right)^2}\right] \quad . \quad 2.10a$$

The maximum value of C_g can be found in a similar manner from 2.9*b* and the resulting equation is

$$B_{g_a}(B_0+B_{g_a})^2 - 2(B_0+B_{g_a})(G_a+G_0)g_m - B_{g_a}(G_a+G_0)^2 = 0$$

$$B_0'' = -B_{g_a} \pm \left[\frac{g_m \pm \sqrt{{g_m}^2+{B_{g_a}}^2}}{B_{g_a}}\right](G_a+G_0)$$

the maximum condition is given from

$$B_0'' = -B_{g_a} + \left[\frac{g_m - \sqrt{{g_m}^2+{B_{g_a}}^2}}{B_{g_a}}\right](G_a+G_0) \quad . \quad 2.11a.$$

Generally, $B_{g_a} \ll g_m$ and (G_a+G_0) so that 2.10*a* and 2.11*a* reduce to

$$B_0' = \pm (G_a+G_0) \quad . \quad . \quad . \quad . \quad 2.10b$$

and

$$B_0'' = -B_{g_a}. \quad . \quad . \quad . \quad . \quad 2.11b.$$

Combining 2.10*b* and 2.9*a* and neglecting B_{g_a} in comparison with the other factors

$$R_g(\text{min.}) = \pm\frac{2(G_a+G_0)}{g_m B_{g_a}} \quad . \quad . \quad . \quad 2.12.$$

Combining 2.11*b* and 2.9*b* and again neglecting B_{g_a}

$$C_g(\text{max.}) = C_{g_a}\left[\frac{G_a+G_0+g_m}{G_a+G_0}\right]$$

$$= C_{g_a}\left[1+\frac{g_m}{G_a+G_0}\right]$$

$$= C_{g_a}\left[1+\frac{g_m R_a R_0}{R_a+R_0}\right]. \quad . \quad . \quad 2.13a$$

$$= C_{g_a}\left[1+\frac{\mu R_0}{R_a+R_0}\right]. \quad . \quad . \quad 2.13b.$$

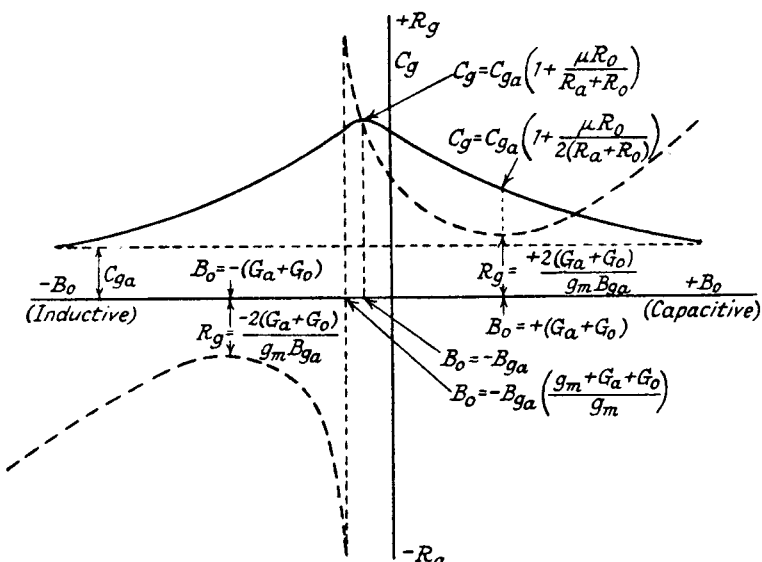


FIG. 2.15.—The Effect on Input Capacitance (C_g) and Input Resistance (R_g) of Varying Anode Circuit Susceptance (B_0).

Expression 2.13*b* is in the well-known form illustrating the Miller effect.

The changes of R_g and C_g for values of B_0 varying from $+\infty$ to $-\infty$ are shown in Fig. 2.15. These effects have a very important bearing on the frequency response of an R.F. amplifier and they are fully discussed in relation to this problem in Section 7.8.

In many practical examples $G_a \ll G_0$ and $B_{g_a} \ll B_0$ so that the expressions for R_g and C_g can be simplified to

$$R_g = \frac{G_0^2+B_0^2}{g_m B_{g_a} B_0} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.14a$$

$$= \frac{B_0}{g_m B_{g_a}} \text{ when } G_0 \ll B_0$$

$$= -\frac{1}{g_m \omega^2 L_0 C_{g_a}} \text{ when } B_0 \text{ is inductive.}$$

$$= \frac{C_0}{g_m C_{g_a}} \text{ when } B_0 \text{ is capacitive.}$$

$$C_g = C_{g_a}\left[\frac{(G_0+g_m)G_0+B_0^2}{G_0^2+B_0^2}\right] = C_{g_a}\left[1+\frac{g_m G_0}{G_0^2+B_0^2}\right] \quad . \quad 2.14b.$$

A valve may have an impedance in its cathode circuit (intentionally as in a negative feedback amplifier or unintentionally due to inductance of the lead from cathode to valve-pin) and the current-voltage equations are modified to (Fig. 2.13 with $C_{g_k} = 0$, $Z_1 = 0$, $I_2 = I_g$)

$$I_g = (E_g+E_0)jB_{g_a}$$

$$Y_g = \left[1+\frac{E_0}{E_g}\right]jB_{g_a} \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.15.$$

$$I_a = \frac{\mu E_{g_k} - (E_0+E_k)}{R_a} = g_m(E_g - E_k) - (E_0+E_k)G_a$$

$$I_a = g_m E_g - E_k(g_m+G_a) - E_0 G_a$$
$$= E_k(G_k+jB_k)$$

$$\therefore\ E_k(G_a+G_k+g_m+jB_k)+E_0G_a = g_m E_g \quad . \quad 2.16a$$

$$(I_a - I_g) = E_0(G_0+jB_0) = E_g(g_m - jB_{g_a}) - E_k(g_m+G_a) - E_0(G_a+jB_{g_a})$$

$$\therefore\ E_k(g_m+G_a)+E_0.[G_a+G_0+j(B_{g_a}+B_0)] = E_g(g_m - jB_{g_a}) \quad 2.16b.$$

Combining 2.16a and 2.16b to eliminate E_k

$$\frac{E_0}{E_g} = \frac{(g_m - jB_{g_a})(G_a+G_k+g_m+jB_k) - g_m(G_a+g_m)}{[G_a+G_0+j(B_{g_a}+B_0)][(G_a+G_k+g_m+jB_k)] - (G_a+g_m)G_a} \quad 2.17a.$$

It will be noted that when G_k and B_k are infinite, i.e., the cathode impedance is zero, 2.17a reduces to 2.8. Normally B_{g_a} and G_a can be neglected in comparison with the other components and so

$$\frac{E_0}{E_g} = \frac{g_m(G_k+jB_k)}{(G_0+jB_0)(G_k+g_m+jB_k)}. \quad . \quad . \quad 2.17b.$$

From 2.15 and 2.17*b*

$$Y_g = jB_{g_a}\left[\frac{(G_0+jB_0)(G_k+g_m+jB_k)+g_m(G_k+jB_k)}{(G_0+jB_0)(G_k+g_m+jB_k)}\right].$$

Rationalizing

$$Y_g = \frac{jB_{g_a}[(G_0+jB_0)(G_k+g_m+jB_k)+g_m(G_k+jB_k)]}{(G_0^2+B_0^2)((G_k+g_m)^2+B_k^2)}\cdot\frac{[(G_0-jB_0)(G_k+g_m-jB_k)]}{}$$

$$R_g = \frac{1}{G_g} = \frac{(G_0^2+B_0^2)((G_k+g_m)^2+B_k^2)}{g_m B_{g_a}[B_0(G_k^2+B_k^2)+g_m(B_0G_k-B_kG_0)]}$$

$$= \frac{G_0^2+B_0^2}{g_m B_{g_a} B_0}\left[1+\frac{g_m(g_m+2G_k)}{G_k^2+B_k^2}\right] \quad . \quad . \quad . \quad . \quad . \quad 2.18a$$

when $B_0G_k - B_kG_0$ is very small.

Comparing this with 2.14*a* we see that the reflected resistance is increased, i.e., the anode damping is reduced. The result was to be expected because the voltage across Z_k is a negative feedback voltage reducing the gain of the valve.

$$C_g = \frac{B_g}{\omega} = C_{g_a}\left[\frac{[(G_0+g_m)G_0+B_0^2][(G_k+g_m)^2+B_k^2] + g_m^2(B_kB_0 - G_kG_0) - g_m^3G^0}{(G_0^2+B_0^2)[(G_k+g_m)^2+B_k^2]}\right]$$

$$= C_{g_a}\left[\frac{(G_0+g_m)G_0+B_0^2}{G_0^2+B_0^2} - \frac{g_m[g_m^2G_0+g_m(G_kG_0-B_kB_0)}{(G_0^2+B_0^2)[(G_k+g_m)^2+B_k^2]}\right]. \quad 2.18b.$$

This is identical with 2.14*b* when G_k and B_k are infinite. The effect of Z_k is therefore to reduce C_g and this also agrees with the reduced valve amplification due to the negative feedback voltage across Z_k.

The above expressions assume that the screen voltage is decoupled by a capacitance to the cathode, and when it is decoupled to earth the following formulae are valid for R_g and C_g.

$$R_g = \frac{G_0^2+B_0^2}{g_m B_{g_a} B_0}\left[1+\frac{g_k(g_k+2G_k)}{G_k^2+B_k^2}\right] \quad . \quad . \quad . \quad . \quad . \quad 2.18c$$

$$C_g = C_{g_a}\left[\frac{(G_0+g_m)G_0+B_0^2}{G_0^2+B_0^2} - \frac{g_m[g_k^2G_0+g_k(G_kG_0-B_kB_0)}{(G_0^2+B_0^2)[(G_k+g_k)^2+B_k^2]}\right]. \quad 2.18d$$

where g_k = mutual conductance with reference to the cathode voltage

$$= g_m\left[\frac{I_a+I_s}{I_a}\right]$$

and I_a and I_s are the anode and screen currents respectively.

2.8.3. Grid Input Admittance and Grid-Cathode Capacitance Coupling. The grid-cathode capacitance of a valve is often comparatively large, about 3 $\mu\mu$F, but it acts only as a capacitance in parallel with the grid input circuit if there is no impedance in the cathode circuit of the valve. At high frequencies (from about 20 Mc/s) the inductance of the cathode lead can provide sufficient impedance to modify to a very great extent the grid input admittance. Fig. 2.13 shows the circuit (C_{g_a}, C_{g_s} and Z_1 are assumed to be zero, and $I_3 = I_g$) and the equations are given below

$$I_g = (E_g - E_k)jB_{g_k}$$

$$Y_g = \frac{I_g}{E_g} = \left[1 - \frac{E_k}{E_g}\right].jB_{g_k} \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.19$$

$$I_a = \frac{\mu(E_g - E_k) - (E_0+E_k)}{R_a} = g_m(E_g - E_k) - (E_0+E_k)G_a$$

$$= E_0(G_0+jB_0)$$

$$I_a+I_g = E_k(G_k+jB_k)$$

from which $E_0(G_a+G_0+jB_0)+E_k(G_a+g_m) = g_mE_g$

and

$$E_0G_a+E_k[G_a+G_k+g_m+j(B_{g_k}+B_k)] = E_g(g_m+jB_{g_k}).$$

Combining the above equations to eliminate E_0

$$\frac{E_k}{E_g} = \frac{(g_m+jB_{g_k})(G_a+G_0+jB_0) - g_mG_a}{[G_a+G_k+g_m+j(B_{g_k}+B_k)][G_a+G_0+jB_0] - (G_a+g_m)G_a} \quad 2.20a.$$

If G_a can be neglected

$$\frac{E_k}{E_g} = \frac{g_m+jB_{g_k}}{G_k+g_m+j(B_{g_k}+B_k)} \quad . \quad . \quad . \quad 2.20b$$

and we see that the anode impedance (Z_0) has practically no effect on the result. Combining 2.19 and 2.20*b*

$$Y_g = jB_{g_k}\frac{(G_k+jB_k)(G_k+g_m - j(B_{g_k}+B_k))}{(G_k+g_m)^2+(B_{g_k}+B_k)^2}.$$

$$R_g = \frac{1}{G_g} = \frac{(G_k+g_m)^2+(B_{g_k}+B_k)^2}{B_{g_k}(G_kB_{g_k} - g_mB_k)} \quad . \quad . \quad 2.21a$$

$$C_g = \frac{B_g}{\omega} = \frac{C_{g_k}[G_k(G_k+g_m)+B_k(B_{g_k}+B_k)]}{(G_k+g_m)^2+(B_{g_k}+B_k)^2} \quad . \quad 2.21b.$$

These two expressions are similar in form to those for anode grid capacitance coupling.

Differentiating R_g with respect to B_k gives the following condition for minimum R_g.

$$B_k' = -B_{g_k}+(G_k+g_m)\left[\frac{B_{g_k}}{g_m}\pm\sqrt{1+\left(\frac{B_{g_k}}{g_m}\right)^2}\right] \quad . \quad 2.22a.$$

C_g has a minimum value at

$$B_k'' = -B_{g_k}-\left[\frac{g_m-\sqrt{g_m^2+B_{g_k}^2}}{B_{g_k}}\right](G_k+g_m) \quad . \quad 2.22b$$

and its maximum value is C_{g_k} for $B_k = \infty$.

This is the reverse of anode-grid coupling, but is to be expected since the voltage developed across a given impedance in the cathode circuit with regard to earth is 180° out-of-phase with that developed across the same impedance in the anode circuit. In the same way it is found that the sign of R_g is reversed, i.e., a positive or capacitive B_k gives a negative R_g. This is clearly shown below in the simplified expressions for R_g and C_g obtained by assuming $B_{g_k} \ll g_m$. Thus 2.22*a* and 22*b* become

$$B_k' = \pm (G_k+g_m)$$
$$B_k'' = -B_{g_k}$$

and
$$R_g(\text{min.}) = \mp\frac{2(G_k+g_m)}{g_m B_{g_k}} \quad . \quad . \quad . \quad 2.23$$

and
$$C_g(\text{min.}) = C_{g_k}\frac{G_k}{G_k+g_m}$$
$$= C_{g_k}\left[1-\frac{g_m}{G_k+g_m}\right] \quad . \quad 2.24.$$

The general shape of the variations of R_g and C_g for changes of B_k from $-\infty$ to $+\infty$ is shown in Fig. 2.16. R_g is infinite when

$$B_k = +B_{g_k}\frac{G_k}{g_m}.$$

Many practical cases allow simplification of the general expressions 2.21*a* and 21*b* for R_g and C_g. Thus if $G_k \ll B_k$ and g_m.

$$R_g = -\frac{g_m^2+(B_{g_k}+B_k)^2}{g_m B_{g_k} B_k} \quad . \quad . \quad . \quad 2.21c$$

and
$$C_g = C_{g_k}\frac{B_k(B_{g_k}+B_k)}{g_m^2+(B_{g_k}+B_k)^2} \quad . \quad . \quad . \quad 2.21d.$$

Often B_{g_k} can be neglected in comparison with B_k and then

$$R_g = -\frac{g_m^2+B_k^2}{g_m B_{g_k} B_k} \quad . \quad . \quad . \quad 2.21e$$

and

$$C_g = C_{g_k}\frac{B_k^2}{g_m^2+B_k^2} \quad . \quad . \quad . \quad 2.21f.$$

In the simplest case when $g_m \ll B_k$

$$R_g = -\frac{B_k}{g_m B_{g_k}} = +\frac{1}{g_m \omega^2 L_k C_{g_k}} = -\frac{C_k}{g_m C_{g_k}}$$

when B_k is respectively inductive and capacitive.

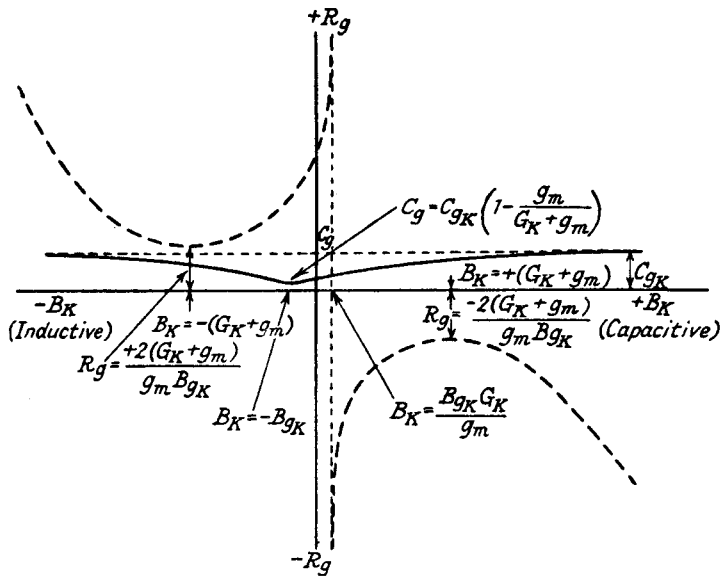


FIG. 2.16.—The Effect on Input Capacitance (C_g) and Input Resistance (R_g) of Varying Cathode Circuit Susceptance (B_k).

(Note that when the screen voltage is not decoupled to cathode, g_m in all the above expressions must be replaced by g_k, the mutual conductance referred to the cathode.)

In all the above expressions we have assumed that the screen is decoupled to cathode, and if it is decoupled to earth g_m must everywhere be replaced by g_k, where $g_k = g_m\left(1+\frac{I_s}{I_a}\right)$. When the effect of cathode lead inductance is considered, g_k must replace g_m. To illustrate the influence of the cathode lead inductance in an R.F. amplifier valve at high frequencies, let us consider a valve operating at 30 Mc/s. and having the following constants

$$g_k = 3 \text{ mA/volt}, \quad C_{g_k} = 3 \ \mu\mu\text{F}, \quad L_k = 0{\cdot}2 \ \mu\text{H}.$$

Using formula 2.21c

$$R_g = \frac{(3 \times 10^{-3})^2 + (0{\cdot}0565 \times 10^{-2} - 2{\cdot}65 \times 10^{-2})^2}{3 \times 10^{-3} \times 5{\cdot}65 \times 10^{-4} \times 2{\cdot}65 \times 10^{-2}}$$

$$= 15{,}200\Omega$$

$$C_g \simeq C_{g_k} \text{ as } g_k \ll B_k.$$

This result shows how serious an effect even a small cathode inductance can have when amplification at comparatively high frequencies is desired. Since R_g is inversely proportional to g_k, the amplifying properties of a high g_m valve may be completely nullified unless L_k is made very small. To reduce additions to L_k from external wiring, leads from the cathode decoupling capacitor to earth should be as short as possible and the capacitor itself must be non-inductive. The part contributed to L_k by the valve internal wiring may be reduced by bringing out by separate leads more than one connection from the cathode.

If the frequency is raised to 60 Mc/s, the resistance R_g falls to the very low value of 3,500Ω, i.e., almost in inverse proportion to the square of the frequency ratio change, 4 to 1.

It is possible to increase the input resistance of a valve by inserting a resistance R_k in the cathode lead, but this entails loss of amplification in the valve owing to the degenerative effect of R.F. voltages set up across R_k. Under certain circumstances the increase in input resistance may be greater than the decrease in valve amplification, and a net increase in overall amplification may result. An example in Section 4.10.3 illustrates this.

The formula for input grid resistance when R_k is included is that given in expression 2.21a.

$$R_g = \frac{(G_k + g_k)^2 + (B_k + B_{g_k})^2}{B_{g_k}(G_k B_{g_k} - g_k B_k)}.$$

To have any appreciable effect R_k must be comparable with the reciprocal of mutual conductance, so let us assume that $R_k = 300\Omega$, other components remaining as in the example quoted above. Converting the cathode impedance to an admittance

$$Y_k = \frac{1}{Z_k} = \frac{1}{300 + 37{\cdot}7j} = \frac{300}{300^2 + 37{\cdot}7^2} - \frac{37{\cdot}7j}{300^2 + 37{\cdot}7^2}$$

$$\therefore G_k = 3{\cdot}28 \times 10^{-3} \text{ mhos and } B_k = -0{\cdot}412 \times 10^{-3} \text{ mhos.}$$

$$R_g = \frac{(6{\cdot}28 \times 10^{-3})^2 + (1{\cdot}53 \times 10^{-4})^2}{5{\cdot}65 \times 10^{-4}(18{\cdot}5 \times 10^{-7} + 12{\cdot}36 \times 10^{-7})}$$

$$= 22{,}700\Omega.$$

The inclusion of R_k has, however, reduced the equivalent mutual conductance of the valve to (expression 2.5*e*).

$$g_m' = \frac{g_m}{1+g_k Z_k} \simeq \frac{g_m}{1+3 \times 10^{-3} \times 300} = 0{\cdot}526\ g_m.$$

This is a serious reduction and the increase in R_g could not be expected to offset it. However, in this example we have not taken into account the probable value of stray capacitance across R_k, and we shall see in the next example that even a small capacitance value can have a profound influence on the result. Let us consider R_k as paralleled by stray capacitance of 5 $\mu\mu$F. The cathode impedance Z_k now becomes

$$\begin{aligned} Z_k &= \frac{R_k}{1+j\omega C_k R_k} + j\omega L_k \\ &= \frac{R_k}{1+(\omega C_k R_k)^2} + j\left[\omega L_k - \frac{\omega C_k R_k^2}{1+(\omega C_k R_k)^2}\right] \\ &= 278 + j(37{\cdot}7 - 78{\cdot}5) \\ &= 278 - j40{\cdot}8 \end{aligned}$$

and $$\begin{aligned} Y_k &= 3{\cdot}52 \times 10^{-3} + j5{\cdot}17 \times 10^{-4} \\ &= G_k + jB_k. \end{aligned}$$

Hence $$\begin{aligned} R_g &= \frac{(6{\cdot}52 \times 10^{-3})^2 + (1{\cdot}082 \times 10^{-3})^2}{5{\cdot}65 \times 10^{-4}(19{\cdot}9 \times 10^{-7} - 15{\cdot}51 \times 10^{-7})} \\ &= 176{,}000\Omega \left(G_g = \frac{1}{R_g} = 5{\cdot}68\mu \text{ mhos}\right). \end{aligned}$$

This represents a very big improvement in input resistance, though the actual value of R_g is now much more dependent on frequency, and the application of this principle will clearly be more satisfactory in receivers operating at a fixed frequency (e.g., for television reception) or over a restricted tuning range.

One result of including R_k must not be overlooked; R_g is no longer infinite when $g_k = 0$. In the above example it has the comparatively low value of 12,050 ohms, steadily increasing as cathode mutual conductance is increased, to 176,000Ω at $g_k = 3$ mA/volt. This particular property may be used to reduce input resistance variations under A.G.C. conditions, for by a suitable choice of R_k it is possible to make R_g almost independent of changes of g_k. This can be illustrated by taking $R_k = 120\Omega$, all other components being as before:

$Z_k = 118{\cdot}3 + j24{\cdot}32.$

$G_k = 8{\cdot}11 \times 10^{-3}; \quad B_k = -1{\cdot}665 \times 10^{-3}$

$$R_g = \frac{(11{\cdot}11 \times 10^{-3})^2 + (1{\cdot}1 \times 10^{-3})^2}{5{\cdot}65 \times 10^{-4}[5{\cdot}65 \times 10^{-4} \times 8{\cdot}11 \times 10^{-3} + 3 \times 10^{-3} \times 1{\cdot}665 \times 10^{-3}]}$$

$= 23{,}050\Omega$ ($G_g = 43{\cdot}4$ micromhos)

for $g_k = 2$ mA/volt, $R_g = 23{,}100\Omega$, $G_g = 43{\cdot}2\mu$ mhos
" $= 1$ " " $= 23{,}800\Omega$, " $= 42{\cdot}0$ "
" $= 0$ " " $= 25{,}900\Omega$, " $= 38{\cdot}5$ "

The use of a cathode resistance may increase or decrease the variation of input capacitance when g_k is varied. From expression 2.21*b*.

$$C_g = AC_{g_k}$$

where

$$A = \frac{G_k(G_k + g_k) + B_k(B_{g_k} + B_k)}{(G_k + g_k)^2 + (B_{g_k} + B_k)^2}.$$

Least variation of A is obtained when $(B_{g_k} + B_k)$ is large compared with g_k, and under these conditions increase of G_k increases the variation of C_g. In the examples given above for $G_k = 0$, A varies from 1·01 ($g_k = 3$ mA/volt) to 1·02 ($g_k = 0$), whereas when $R_k = 120\Omega$, $C_k = 5$ $\mu\mu$F, it changes from 0·74 ($g_k = 3$ mA/volt) to 1·01 ($g_k = 0$). The inclusion of R_k therefore increases the variation of C_g as g_k is varied. If $(B_{g_k} + B_k)$ is comparable with g_k, increase of G_k generally reduces the capacitance variation, so that as the operating frequency increases (B_{g_k} and B_k decrease) we find that a cathode resistance can be used to advantage for reducing input capacitance variations.

2.8.4. Grid Input Admittance and Combined Anode-Grid and Grid-Cathode Capacitance Coupling. The circuit for this coupling is shown in Fig. 2.13. [$C_{g_a} = 0$ and $Z_1 = 0$] and the equations are

$$I_a = g_m(E_g - E_k) - G_a(E_k + E_0) \quad . \quad . \quad . \quad 2.25.$$
$$I_3 = (E_g - E_k)jB_{g_k} \quad . \quad . \quad . \quad . \quad . \quad 2.26.$$
$$I_2 = (E_g + E_0)jB_{g_a} \quad . \quad . \quad . \quad . \quad . \quad 2.27.$$
$$I_a + I_3 = E_k(G_k + jB_k) \quad . \quad . \quad . \quad . \quad . \quad 2.28.$$
$$I_a - I_2 = E_0(G_0 + jB_0) \quad . \quad . \quad . \quad . \quad . \quad 2.29.$$

$$Y_g = \frac{I_g}{E_g} = \frac{I_2 + I_3}{E_g} = \left(1 - \frac{E_k}{E_g}\right)jB_{g_k} + \left(1 + \frac{E_0}{E_g}\right)jB_{g_a} \quad 2.30.$$

Combining 2.25, 2.26 and 2.28

$$E_k[G_a+G_k+g_m+j(B_k+B_{g_k})]+E_0G_a = E_g[g_m+jB_{g_k}] \quad 2.31a.$$

From 2.25, 2.27 and 2.29

$$E_0[G_0+G_a+j(B_0+B_{g_a})]+E_k[g_m+G_a] = E_g(g_m - jB_{g_a}) \quad 2.31b.$$

Eliminating E_0 from 2.31a and b

$$\frac{E_k}{E_g} = \frac{[g_m+jB_{g_k}][G_a+G_0+j(B_0+B_{g_a})] - [g_m - jB_{g_a}]G_a}{[G_a+G_k+g_m+j(B_{g_k}+B_k)]\,[G_a+G_0+j(B_0+B_{g_a})] - [g_m+G_a]G_a} \quad 2.32.$$

If B_{g_a} can be neglected in comparison with its associated factors, 2.32 reduces to 2.20a, so that the part contributed to Y_g by the grid-cathode capacitance is the same as that given in Section 2.8.3.

Combining 2.31a and b to eliminate E_k

$$\frac{E_0}{E_g} = \frac{[g_m - jB_{g_a}][G_a+G_k+g_m+j(B_{g_k}+B_k)] - [g_m+jB_{g_k}][G_a+g_m]}{[G_a+G_0+j(B_0+B_{g_a})]\,[G_a+G_k+g_m+j(B_{g_k}+B_k)] - [g_m+G_a]G_a} \quad 2.33.$$

Again, if B_{g_k} can be neglected in comparison with the other factors, 2.33 reduces to 2.17a, and its contribution to Y_g is as calculated in the second part of 2.8.2. We may therefore take the values of G_g and C_g calculated in Sections 2.8.2. and 2.8.3, and obtain the combined effects of anode-grid and grid-cathode coupling by adding them.

Thus, from the modified expression 2.18a and from 2.21a (neglecting B_{g_k} in comparison with other factors)

$$G_g = \frac{g_m B_{g_a} B_0 (G_k^2+B_k^2)}{(G_0^2+B_0^2)[(G_k+g_m)^2+B_k^2]} - \frac{g_m B_{g_k} B_k}{(G_k+g_m)^2+B_k^2}$$

$$G_g = \frac{g_m}{[(G_k+g_m)^2+B_k^2]}\left[B_{g_a}B_0.\left[\frac{G_k^2+B_k^2}{G_0^2+B_0^2}\right] - B_{g_k}B_k\right]$$

$$R_g = \frac{[(G_k+g_m)^2+B_k^2][G_0^2+B_0^2]}{g_m[B_{g_a}B_0(G_k^2+B_k^2) - B_{g_k}B_k(G_0^2+B_0^2)]} \quad 2.34a.$$

Neglecting B_{g_a} and the second part in 2.18b, and also B_{g_k} in 2.21b.

$$C_g \simeq C_{g_a}.\left[\frac{(G_0+g_m)G_0+B_0^2}{G_0^2+B_0^2}\right]+C_{g_k}\left[\frac{(G_k+g_m)G_k+B_k^2}{(G_k+g_m)^2+B_k^2}\right] \quad 2.34b.$$

Examining 2.34a, we see that it is possible to make R_g infinite [7] and therefore independent of the mutual conductance of the valve if the denominator is zero, i.e.,

$$B_{g_a}B_0.(G_k^2+B_k^2) = B_{g_k}B_k(G_0^2+B_0^2)$$

or

$$\frac{B_{g_a}}{B_{g_k}} = \frac{\dfrac{B_k}{G_k^2+B_k^2}}{\dfrac{B_0}{G_0^2+B_0^2}} \quad . \quad . \quad . \quad 2.35a.$$

But $\dfrac{B_k}{G_k^2+B_k^2}$* is the reactance of the cathode circuit viewed as a series circuit consisting of R_k and L_k (or C_k, though this is less likely to be realized at ultra high frequencies), and similarly $\dfrac{B_0}{G_0^2+B_0^2}$ is the equivalent series reactance of the anode. Hence 2.35*a* becomes

$$\frac{B_{g_a}}{B_{g_k}} = \frac{C_{g_a}}{C_{g_k}} = \frac{L_k}{L_a} \quad . \quad . \quad . \quad 2.35b.$$

It will be noticed that B_0 and B_k must have the same sign (B_{g_a} and B_{g_k} are always capacitive and positive), i.e., they must both be inductive or capacitive. Thus in satisfying expression 2.35*b* by including an inductance $L_a = L_k\dfrac{C_{g_k}}{C_{g_a}}$ between the anode load impedance and the anode, it is possible to raise the input resistance to a very high value. Just as for the added cathode resistance method, there is loss of valve amplification due to the added anode inductance, which may need to be as much as fifty times that of the stray cathode inductance. Input resistance neutralization is also possible by means of an inductance in the screen electrode of the valve.

Expression 2.34*b* for input capacitance may be rewritten as

$$C_g = C_{g_a}.\left[1+\frac{g_mG_0}{G_0^2+B_0^2}\right]+C_{g_k}\left[1-\frac{(G_k+g_m)g_m}{(G_k+g_m)^2+B_k^2}\right]$$

$$= C_{g_a}+C_{g_k}+g_m\left[\frac{C_{g_a}G_0}{G_0^2+B_0^2}-\frac{C_{g_k}(G_k+g_m)}{(G_k+g_m)^2+B_k^2}\right] \quad . \quad 2.34c.$$

If $g_m \ll G_k$ it is also possible to prevent change of input capacitance when g_m is varied (under A.G.C. conditions); this is achieved by making

$$\frac{C_{g_a}}{C_{g_k}} = \frac{\dfrac{G_k}{G_k^2+B_k^2}}{\dfrac{G_0}{G_0^2+B_0^2}} = \frac{R_k}{R_0} \quad . \quad . \quad . \quad 2.36$$

* See Appendix 1A.

where R_k and R_0 are the equivalent series resistance components of the cathode and anode circuits.

It will similarly be found for capacitance coupling between the grid and other electrodes that each coupling can be treated independently of any other, and the resultant input conductances and capacitances can be added to obtain the combined effect.

2.8.5. Grid Input Admittance and Grid-Screen Capacitance Coupling. The equations for grid-screen capacitance coupling are identical with those for anode-grid coupling if G_a is replaced by G_s, the screen conductance $\frac{\Delta I_s}{\Delta E_s}$ (this is the slope of the I_sE_s curve), g_m by g_s, the screen mutual conductance, $\frac{\Delta I_s}{\Delta E_g}$, G_0, B_0 and B_{g_s} by G_1, B_1, the conductance and susceptance of the external impedance of the screen circuit, and B_{g_s}, the screen-to-grid capacitance.

Thus
$$R_g = \frac{(G_s+G_1)^2+(B_1+B_{g_s})^2}{B_{g_s}[g_sB_1+B_{g_s}(G_s+G_1+g_s)]} \quad . \quad . \quad . \quad 2.37a$$

$$C_g = C_{g_s}\left[\frac{(G_s+G_1+g_s)(G_s+G_1)+B_1(B_1+B_{g_s})}{(G_s+G_1)^2+(B_1+B_{g_s})^2}\right] \quad . \quad 2.37b.$$

In an R.F. amplifier, having undesired inductance in the screen lead, B_{g_s} and (G_s+G_1) can usually be neglected in comparison with B_1, so that

$$R_g \simeq \frac{B_1}{g_sB_{g_s}} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 2.38$$

$$\simeq -\frac{1}{g_s\omega^2L_1C_{g_s}} \text{ when } B_1 \text{ is inductive}$$

$$\simeq \frac{C_1}{g_sC_{g_s}} \text{ when } B_1 \text{ is capacitive.}$$

As for combined anode-grid and grid-cathode capacitance coupling, so for combined grid-screen and grid-cathode coupling it is possible to raise the input resistance to an infinite value; in this instance the condition is that

$$\frac{C_{g_s}}{C_{g_k}} = \frac{L_k}{L_s} \quad . \quad . \quad . \quad . \quad 2.39.$$

Since C_{g_s} and C_{g_k} are comparable in value it means that L_s is approximately equal to L_k.

2.8.6. Grid Input Admittance and Electron Transit Time.[3] Over the medium and long wave ranges the time taken for a given electron to travel from the cathode to the anode is a very small fraction of one period of the input voltage wave. Consequently the A.C. current flow in the grid circuit, due to the passage of electrons through the grid wires from cathode to anode, is almost entirely capacitive and leads the grid voltage producing it by 90°. Hence there is no absorption of power in the grid circuit and grid input resistance is very high. At the high-frequency end of the short wave range (from about 15 Mc/s upwards), the transit time of an electron between cathode and grid becomes an appreciable fraction of the grid voltage cycle. This causes the current action in the grid circuit to be delayed and grid current leads upon the

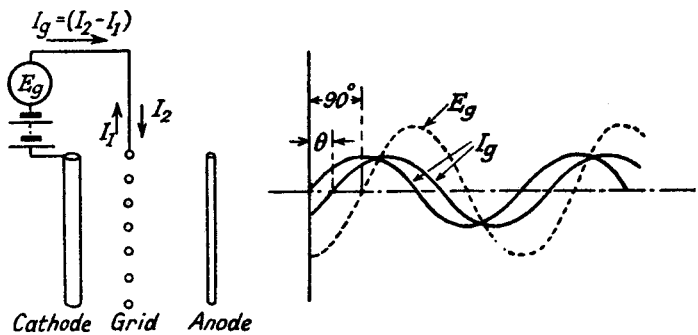


FIG. 2.17.—Grid Current due to Electrons Passing from the Cathode to Anode. (θ is the angle of lag due to the transit time of the electrons.)

grid voltage by an angle less than 90°; i.e., a conductance component is introduced into the grid input admittance. This conductance component increases as the frequency of the grid voltage input increases and in general-purpose valves is quite marked at 20 Mc/s.

Ferris [2] has explained the action as follows. In Fig. 2.17 is shown a cathode, grid and third electrode with electrons passing through the grid mesh. The approach of the negatively charged electrons, to the grid from the cathode, induces a charge on the grid which is equivalent to a current I_1 flowing into the grid circuit. Similarly as the electrons progress from the grid towards the third electrode they induce a charge on the grid mesh, which is equivalent to a current I_2 flowing out of the grid circuit. When there is no A.C. voltage on the grid and sufficient negative bias to prevent the collection of electrons, the grid current due to the approaching

electrons is completely neutralized by that due to the departing electrons. If an A.C. voltage is now applied to the grid, the distribution of the electrons is modulated by this voltage so that as E_g is increasing negatively the density of the electron stream is greatest on the cathode side, and there is a net current into the grid, i.e., I_g is negative. The reverse is true when E_g is decreasing and the density is greatest on the side opposite to the cathode, giving a net current out of the grid. The magnitude of the net current is dependent on the rate at which the electrons are being accelerated or decelerated, and is a maximum when this is maximum (provided there is no delay between the voltage E_g and its effect on the electrons), i.e., when E_g is changing at its greatest rate. This occurs when the A.C. component of E_g is passing through zero. The net current I_g (see Fig. 2.17) therefore leads E_g by 90°, being zero when E_g is maximum or minimum, increasing positively as E_g increases from its minimum value, and reaching a maximum when the A.C. component of E_g is zero. Thus the input admittance has no conductance component. If the time taken for electrons to travel from the cathode region to the grid is an appreciable fraction of the input voltage cycle, the effect of E_g on the electron stream distribution is delayed and the net current leads E_g by less than 90°. The new phase relationship is $(90° - \theta)$, where θ is the angular lag caused by the electron transit time from cathode to grid. The input grid admittance Y_g is

$$Y_g = \frac{I_g}{E_g} = |\, Y_g \,| \; \underline{/90 - \theta} \quad . \quad . \quad . \quad 2.40.$$

and it has a conductance component

$$G_g = |\, Y_g \,| \cos (90 - \theta) = |\, Y_g \,| \sin \theta \quad . \quad . \quad 2.41$$

Ferris suggests that the value of G_g is given by

$$G_g = K g_m f^2 \tau^2$$

where g_m = mutual conductance of the valve

K = a factor dependent on the ratio of the electron transit times in the cathode-grid and grid-third electrode spaces

f = frequency

τ = electron transit time to an arbitrary point in the system.

Electron transit time is reduced by using smaller electrode spacings as in the acorn type valve.

In the above description of electron transit time effects it is assumed that the valve is operating under space-charge current

limitations, i.e., the total current is far below its saturation value. If the current approaches its saturation value (due to reduced cathode temperature, or to a positively biased screen between the signal electrode and cathode as in the heptode frequency changer), current due to the flow of electrons past the grid wires may lag behind the voltage controlling the electrons, i.e., the input susceptance of the signal electrode is equivalent to a negative capacitance. Electron transit time causes the current to lag by more than 90°, so introducing a negative conductance component. This feature is discussed more fully in Section 5.8.3., where an alternative explanation of the positive conductance component under space-charge limited conditions is also given.

There are other forms of electron transit time effects in multi-electrode frequency changers, such as negative resistance [6] coupling between the oscillator and signal grids of a heptode (Section 5.8.3.), and the high negative bias start of grid current on the signal grid of a hexode valve (Section 5.8.2.).

BIBLIOGRAPHY

1. A Note on an Alternative Equivalent Circuit for the Thermionic Valve. N. R. Bligh, *Wireless Engineer*, Sept. 1930, p. 480.
2. The Input Resistance of Vacuum Tubes as Ultra High Frequency Amplifiers. W. R. Ferris, *Proc. I.R.E.*, Jan. 1936, p. 82.
3. The Analysis of the Effects of Space Charge on Grid Impedance. D. O. North, *Proc. I.R.E.*, Jan. 1936, p. 108.
4. The Anode to Accelerating Electrode Space in Thermionic Valves. J. H. O. Harries, *Wireless Engineer*, April 1936, p. 190.
5. Modern Receiving Valves. M. Benjamin, C. W. Cosgrove, and G. W. Warren, *Journal I.E.E.*, April 1937, p. 401.
6. Electron Transit Time Effects in Multigrid Valves. M. J. O. Strutt, *Wireless Engineer*, June 1938, p. 315.
7. Input Conductance Neutralization. R. L. Freeman, *Electronics*, Oct. 1939, p. 22.
8. Input Conductance. F. Preisach and I. Zakariás, *Wireless Engineer*, April 1940, p. 147.
9. *Theory of Thermionic Vacuum Tubes.* E. L. Chaffee. Text-book.
10. *Theory and Application of Electron Tubes.* H. J. Reich. Text-book.

CHAPTER 3

AERIALS AND AERIAL COUPLING CIRCUITS

3.1. Introduction. The difficulties presented by the design of the aerial circuit of a receiver are very largely due to the fact that the type of aerial with which the receiver will operate is unknown. For example, the aerial may be erected indoors, where it will have poor pick-up qualities and large capacitive and resistive components in its terminal impedance. (The terminal impedance is defined as the impedance between aerial and earth looking into the aerial from the receiver.) On the other hand, it may be erected outside, as a horizontal wire with perhaps a long lead-in, or as a short vertical wire at the highest point in the building and connected to the receiver by a screened cable. The characteristics of the two outside aerials are widely different, the former may have a terminal impedance with a low resistive and high reactive component varying appreciably over the tuning frequency range, whilst the latter, owing to the predominating effect of the screened cable, presents to the receiver a mainly resistive impedance less dependent on frequency. If the terminal impedance of the aerial is known over the tuning frequency range, the aerial may be replaced by a generator having an internal impedance equal to the terminal impedance of the aerial and an open circuit voltage equal to the effective pick-up voltage in the aerial. The problem then becomes one of matching aerial terminal and receiver input impedance, if maximum voltage transfer is desired. Before dealing with the aerial connection it is essential to set out briefly the chief features of electromagnetic wave propagation through space.

3.2. Propagation of Electromagnetic Waves. An A.C. current flowing through a wire produces a circularly disposed magnetic field about the wire in a plane perpendicular to the wire. The alternate collapse and reversal of the magnetic field due to the A.C. variation of current induces across the ends of the wire a voltage which is equal to the rate of change of the magnetic flux surrounding the wire and is in opposition to the applied voltage. The latter has, therefore, to overcome not only the resistance of the wire but also this induced component from the magnetic field and the voltage equation is represented by

$$E = RI + \frac{d\Phi}{dt} \quad . \quad . \quad . \quad . \quad 3.1a$$

but $$\frac{d\Phi}{dt} = \frac{d\Phi}{dI} \cdot \frac{dI}{dt} = L \cdot \frac{dI}{dt}$$

where L is the inductance of the wire. Expression 3.1*a* thus becomes

$$E = RI + L\frac{dI}{dt} \quad . \quad . \quad . \quad . \quad 3.1b.$$

If $I = \hat{I} \sin \omega t$ and we assume that there is no time lag between the appearance of the magnetic flux and the current producing it, the expression is again modified to

$$E = R\hat{I} \sin \omega t + L\frac{d\hat{I} \sin \omega t}{dt}$$

$$E = R\hat{I} \sin \omega t + \omega L\hat{I} \cos \omega t \quad . \quad . \quad . \quad 3.1c.$$

The power absorbed in the wire $= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} E\hat{I} \sin \omega t \, dt$

$$= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} [R\hat{I}^2 \sin^2 \omega t + \omega L\hat{I}^2 \cos \omega t \sin \omega t]dt.$$

$$= \frac{\omega}{2\pi}\int_0^{\frac{2\pi}{\omega}} \left[R\hat{I}^2\left[\frac{1 - \cos 2\omega t}{2}\right] + \frac{\omega L\hat{I}^2}{2} . \sin 2\omega t\right]dt$$

$$= \frac{\omega}{2\pi}\left[R\hat{I}^2\left(\frac{t}{2} - \frac{\sin 2\omega t}{4}\right) - \omega L\hat{I}^2 . \frac{\cos 2\omega t}{4}\right]_0^{\frac{2\pi}{\omega}}$$

$$= \frac{RI^2}{2}.$$

Hence there is no loss of energy when the magnetic flux is produced instantaneously, because the energy required to produce it is reabsorbed into the wire when the flux collapses again. This conception of instantaneous flux changes is largely true for low frequencies, such as 50 c.p.s. At much higher frequencies we would expect to find some time lag between the establishment of the magnetic field, especially at some distant point, and the current producing it. The expression for the flux is therefore modified to

$$\Phi = L\hat{I} \sin (\omega t - \theta)$$

and equation 3.1*b* to

$$E = R\hat{I} \sin \omega t + \frac{d}{dt}(L\hat{I} \sin (\omega t - \theta))$$

$$E = R\hat{I} \sin \omega t + \omega L\hat{I} \cos (\omega t - \theta)$$

$$= R\hat{I} \sin \omega t + \omega L\hat{I} (\cos \omega t \cos \theta + \sin \omega t \sin \theta)$$

$$= (R\hat{I} + \omega L\hat{I} \sin \theta) \sin \omega t + \omega L\hat{I} \cos \theta \cos \omega t \quad . \quad 3.1d.$$

The time lag in Φ has increased the resistance term by $\omega L\hat{I} \sin \theta$, and it is this energy component which is radiated into space and part of which is picked up by any suitably disposed conductor. It is known as the radiation resistance of the aerial, and it increases as the frequency increases (this explains the high radiating efficiency of the short wave aerial) and as the magnetic field is more spread out. The latter is achieved by the use of a straight open-ended wire. The magnetic field from a vertical open wire is best imagined in the form of concentric circles in a horizontal plane as shown in Fig. 3.1, and these circles grow in diameter, spreading outwards at the speed of light. The motion of the magnetic field automatically produces an electrostatic field at right angles to it in space but in phase as regards time. The electrostatic field is considered as a series of semicircles (if one end of the aerial is connected to earth)

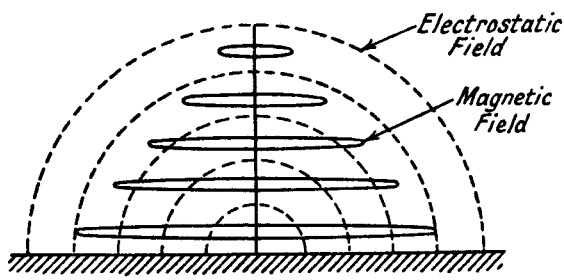


FIG. 3.1.—The Electrostatic and Electromagnetic Fields Developed by an Energised Vertical Aerial.

in a vertical plane (the dotted lines in Fig. 3.1), again moving outwards at the speed of light. The densities of the two fields vary together, i.e., when one is maximum so is the other, and sinusoidally if the current in the vertical wire is sinusoidal. At some distance from the wire (aerial) the magnetic and electrostatic fields are respectively horizontal and vertical, and they may be represented by two vectors at right angles to each other in a plane perpendicular to the direction of travel of the wave front. The horizontal magnetic vector is OM and the vertical electrostatic vector OE in Fig. 3.2*a*. Apart from the simultaneous sinusoidal variation of amplitude with time, the peak value of each vector decreases as the distance from the aerial increases (see $O'M'$ and $O'E'$ in the figure). When the electrostatic vector is vertical as in Fig. 3.2*a*, the wave is said to be vertically polarized, whereas if the electrostatic vector is horizontal (Fig. 3.2*b*), the wave is said to be horizontally polarized. The first condition is realized at some

distance from a vertical aerial if the earth over which the wave passes is a perfect conductor, whilst the second condition occurs by radiation from a horizontal aerial over a perfectly conducting earth. No voltage is induced in a horizontal wire by a vertically polarized wave because a voltage can only be induced in a con-

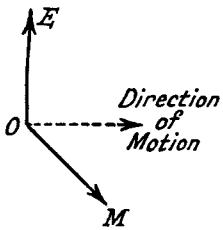
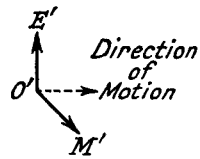


FIG. 3.2*a*.—The Representation of a Vertically Polarized Wave.



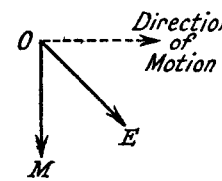


FIG. 3.2*b*.—A Horizontally Polarized Wave.

ductor which is at right angles both to the magnetic field and the direction of motion of the magnetic field.* Similarly a vertical aerial could not be employed for picking up horizontally polarized radiation. Radiation from a vertical aerial does not always result in vertical polarization, for the plane of polarization is tilted forward when the wave passes from a good to a poor conducting earth surface ; the electrostatic vector is tilted forwards and the wave is known as obliquely polarized. Radiation from an aerial at an angle (θ) to the vertical results in an obliquely polarized wave with the two E and M vectors at an angle between the vertical and horizontal polarization positions. Such radiation has vertical and horizontal components of electrostatic and magnetic field, and pick-up is possible with vertical or horizontal aerial. Close to a transmitting aerial, where radiation is mostly by direct ray, the wave is polarized in the plane of the aerial, i.e., is vertically polarized with a vertical aerial, but outside the area of direct ray transmission, where reception is mainly dependent on the indirect ray reflected from the ionosphere (a series of semiconducting layers round the earth), the wave may have any angle of polarization from vertical to horizontal and it may not even be plane polarized, i.e., the vertical and horizontal components of the electric or magnetic field may be out of phase with each other. For example, if the maximum amplitudes of the vertical and horizontal components are equal but have a 90° time phase difference—when the vertical

* This is the basis of Fleming's well-known Right Hand Rule as used in the theory of electrical machines.

is maximum the horizontal is minimum—the wave is said to be circularly polarized. Elliptical polarization is produced by unequal maximum amplitudes with a 90° time phase difference. From the point of view of reception, polarization of the wave front is important in deciding the best orientation of the aerial. For a plane polarized wave, pick-up is maximum when the aerial is parallel to the electrostatic field vector and it is zero at right angles to this. For a circularly or elliptically polarized wave pick-up is possible for any direction of the aerial ; this form of polarization is actually undesirable as it is often produced by conditions (liable to rapid variations) in the earth's upper atmosphere, and distortion of the modulation content of the wave is often severe.

The voltage generated in a receiving aerial may be due to direct or indirect rays from the transmitting aerial, which can be considered as projecting rays in all directions between the horizontal and vertical, though propagation is normally greatest in a horizontal direction or at some shallow angle to this direction. The direct ray travels parallel to the earth's surface, is vertically polarized and is attenuated as it travels away from the transmitting aerial due to the production of eddy currents in the earth. This attenuation increases rapidly as the frequency rises, and for short waves the radius of operation of the direct ray is very limited. An indirect ray is a ray projected at an angle to the earth's surface such that it would ordinarily be lost in space. Owing to the presence of semiconducting layers in the earth's upper atmosphere the ray may be reflected or refracted back to earth. Since the earth itself is a conductor the ray is reflected upwards and is again returned to earth at some distant point. Rays projected at a high angle to the horizontal may make several ricochets between the ionized layers and earth before they reach the receiving aerial.

The semiconducting layers, which return the indirect rays to earth, are caused by ionization of the upper atmosphere from solar radiation. Electrons have been detached from the neutral atoms of the gases in these layers so that besides neutral atoms there are free electrons and positively and negatively charged ions (atoms with a deficiency or excess respectively of electrons in their outer rings).

Subjecting a gas to ultra-violet light or bombardment by a stream of electrons are two methods of producing ionization, and these (due to radiation from the sun) are presumed to be the main causes of the ionized layers surrounding the earth. Pressure plays an important role in ionization and gases at low pressure can

generally be ionized fairly easily and recombine slowly when the cause of ionization is removed. On the other hand, high pressure makes ionization difficult, and recombination is rapid when the ionizing agency disappears. The exact number of ionized layers surrounding the earth is a matter of conjecture, but there are at least two important layers, called the *E* and *F* layers, located at heights of approximately 100 and 250 km. respectively. The lower *E* layer is in a region of comparatively high pressure, so that ionic density is not so high and recombination is rapid when the ionizing agent is removed, i.e., during the night the *E* layer is practically non-existent. The *F* layer, on the other hand, is in a region of lower pressure, ionic density is high and recombination is slow.

The effect produced by these ionized layers depends on their ionic density, the frequency of the transmitted wave, and the angle at which it strikes the layer. The latter may serve as an almost complete reflector, it may allow penetration followed by absorption and/or insufficient refraction for a ray to be returned to earth, or it may allow penetration with little absorption but sufficient refraction to cause the ray to return to earth.

For low frequencies (in the long wave range from approximately 20 to 600 kc/s) the ionosphere acts as an almost perfect reflector and the wave is propagated as if between two concentric reflecting shells formed by the earth and ionized layer. Practically no penetration occurs, and long distance transmission is obtained with steady reception. Direct ray attenuation is not very considerable and this, together with reflected indirect rays, causes the received signal strength to vary in a series of maxima and minima as the distance from transmitter increases. Since these frequencies do not penetrate the ionized layers to any great extent, the indirect ray is less susceptible to variations of ionization, and fading, which is seldom observed, is relatively slow. Owing to the concentric reflector characteristic of the ionosphere and earth, long wave band transmission is almost always vertically polarized at the receiver, though the wave front may be slightly oblique due to earth losses.

Transmission in the medium wave band from about 600 to 1,500 kc/s is characterised by a comparatively limited area of daylight reception; the useful service area, however, is considerably increased at night time. This is due to the fact that the direct ray is fairly rapidly absorbed, due to earth losses, whilst indirect rays penetrate and are absorbed by the lower *E* layer. Indirect rays striking the *E* layer at a shallow angle may be refracted back before penetrating to any depth, and the signal may therefore be

received at some considerable distance from the transmitter. At night when the *E* layer has disappeared, the *F* layer, of much greater ionic density, acts as a reflector for the indirect rays, which are thus returned to earth. High-angle radiation from the transmitting aerial is small, so that the signal from the indirect ray is negligible close to the transmitter ; as the distance from the transmitter increases, lower angle radiation contributes to the indirect ray, the field strength of which increases and remains comparatively high over a considerable distance. The increased power in the low-angle radiation more than compensates for the attenuation in the transmission path. At some point the direct and indirect rays will have comparable field strengths, and in this region night reception is generally unsatisfactory with rapid fading and distortion. This is due to variations in the ionized layer causing variations in the amplitude and phase of the indirect ray so that at one moment it may add and at another subtract from the direct ray. The refractive properties of the ionosphere vary according to frequency, and the phase and amplitude of the modulation sideband frequencies may be independently varied during refraction so that some sidebands may add to those of direct-ray sidebands whilst others subtract. The balance of the modulation frequencies may therefore be completely changed, resulting in severe distortion. The term selective fading is applied when the modulation sideband and carrier frequencies fade independently of each other. Beyond the area of equal direct and indirect ray reception, the signal may vary in strength, and distortion may be produced by the arrival of two indirect rays by different paths, but the resulting signal is generally very much more constant than that obtained in the intermediate area.

Over the short wave band, 2 Mc/s to 25 Mc/s, the direct ray is very rapidly attenuated as it travels over the earth's surface, and long-distance reception is entirely dependent on the indirect ray. The indirect rays penetrate the *E* layer but are refracted by the *F* layer of higher ionic density. If the angle of incidence of the ray is less than a certain critical angle, which falls as the frequency rises, the ray is completely refracted and returned to earth. Above a frequency of about 30 Mc/s, it is almost impossible to obtain a shallow enough critical angle for complete refraction but special freak conditions may produce occasional refraction to earth. Owing to this critical angle effect there are areas, beyond the range of the direct ray, over which no signal can be received except from scattered radiation very erratic in character. The

distance over which reception is almost impossible is known as the skip distance. This distance increases as the frequency increases and varies in length according to the time of day and year. For example, the skip distance at a given frequency is less over a daylight area in summer than over a night area in winter, i.e., refraction is greatest for greatest ionic density. The attenuation of the indirect ray by absorption in the ionized layers is inversely proportional to the square of the frequency (high frequencies in the short wave range are less attenuated than low) and is proportional to the ionic density (high ionic density causes high absorption). Attenuation is therefore greatest over a daylight path in the northern or southern hemisphere during summer.

Short wave reception may suffer from slow or rapid fading due to the arrival of several indirect rays by different paths, some after several ricochets between ionosphere and earth. Selective fading with distortion occurs, and occasionally reception may completely disappear for some hours, especially if the ray path runs close to the earth's magnetic poles. The latter effect normally only occurs during periods of intense solar activity. Another form of distortion sometimes met in short wave reception is that known as the echo effect which may result in either blurred or echoed reception. It is due to the time displacement between the arrival of indirect rays by different routes. The echo effect is often due to the combination of a ray by a more or less direct route and a ray which has made a complete circuit of the earth; a delay of about $\frac{1}{7}$ second occurs between the time of arrival of the same modulation cycle on each wave. It is more commonly noted close to a local short wave station rather than at normal distance operation.

3.3. Types of Aerials.

3.3.1. Introduction. Receiver aerials may conveniently be divided into open and closed (frame) aerials. A wire with an open circuited end is a good example of the former, whilst a frame aerial consisting of closed loops is an example of the latter. The open aerial is the more efficient collector and has maximum pick-up when it is perpendicular to the direction of travel of the wave and parallel to the electric field vector, e.g., with vertically polarized transmission (the electric field component is vertical) in a direction parallel to the earth's surface, an aerial perpendicular to the latter gives maximum pick-up. The voltage generated in the wire is proportional to the height of the wire, but is also considerably affected by the proximity of earthed objects, which tend to distort

the transmitted electric field. A frame aerial is a comparatively inefficient collector and maximum pick-up is obtained when the plane of the frame is that of the electric field and the direction of motion of the wave. The voltage generated is proportional to the area of the loop and the number of turns comprising it.

The open aerial can itself be further subdivided into the vertical, inverted L, the T, and the dipole aerial. The T and inverted L aerials are used principally for long and medium wave reception. They are less satisfactory (especially if they are long) for short wave reception, because the aerial becomes a resonant circuit when its length is approximately one-quarter of the wavelength of the desired transmission. This has a serious mistuning and damping effect on the first tuned circuit unless coupling is very loose. The dipole aerial, consisting of two symmetrical open aerials connected to the receiver by a feeder, is mainly employed for short wave operation, and its particular advantages are that it has directional pick-up and can be conveniently connected to the receiver by a balanced feeder, so reducing interference pick-up. For long and medium wave operation the inverted L aerial having a long horizontal top and vertical lead-in has been extensively used. Its horizontal top [13] enables the voltage picked up in the vertical limb to be used more efficiently, and the equivalent open circuit voltage generated in a given length of vertical is increased by the addition of a horizontal top. It has, however, the disadvantage of collecting vertically and horizontally polarized components of a transmitted wave. These components, having a random phase relationship, tend to cause fading and distortion. The vertical aerial is slowly superseding the inverted L aerial because it responds only to vertically polarized transmission, thus minimizing fading and distortion due to obliquely, circularly or elliptically polarized transmission. It is also simpler to construct and erect.

3.3.2. The Vertical Aerial. As far as reception is concerned there are two important aspects of the aerial connection, its voltage pick-up and its terminal impedance. The vertical aerial may be regarded as a network of series inductive and shunt capacitive arms distributed along its length, with a generator in series with each inductive arm as shown in Fig. 3.3*a*. Owing to the open circuited top the voltages generated in each section cannot be equally effective in driving a current through the receiver impedance. It is clear that the voltage generated at the top has no complete circuit and can therefore contribute nothing to the output, whilst the current sent to the base of the aerial by the second

generator is small because the return path provided by C_0 has a high reactance. The lowest generator is, however, working under most efficient conditions.

This means that standing waves of voltage and current are produced along the aerial as shown in Fig. 3.3*b*. Maximum voltage occurs at the top of the aerial and minimum at the base, whilst

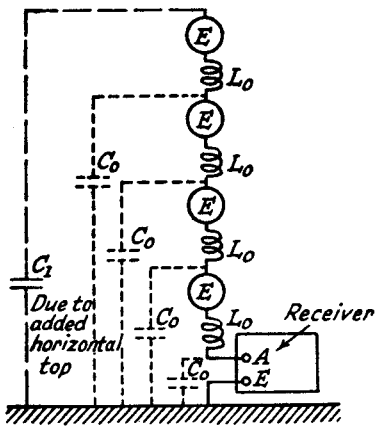


FIG. 3.3*a*.—The Equivalent Generator Circuit for a Vertical Aerial.

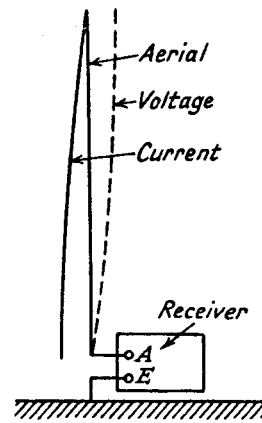


FIG. 3.3*b*.—The Distribution of Current and Voltage in a Vertical Aerial.

current is maximum at the base and minimum (in the special case of the vertical aerial it is zero) at the top. The shape of the standing waves is sinusoidal, but if the height of the aerial is much less than one-quarter of the wavelength the shape may be taken as triangular. Hence the equivalent generated voltage is the average voltage of the whole length of aerial, i.e., is $\frac{Eh}{2}$, where E is the voltage pick-up per unit length and h is the total length of the aerial. It is more usual to associate the $\frac{1}{2}$ with h than with E and $\frac{h}{2}$ is designated as the effective height of the aerial. Thus if we have a vertical aerial 3 metres (9·8 feet) high in a field of 20 μV per metre, the generated voltage is $20 \times \frac{3}{2} = 30$ μV. In this calculation we have assumed that the transmitted electromagnetic field is uniformly distributed from the earth upwards. Changes in earth conductivity and the presence of earthed conductors near the aerial distort and weaken the field. The effective height of low aerials is therefore usually less than $\frac{h}{2}$.

The analysis of aerial terminal impedance has been made by Howe,[1] who treats the aerial as an open-circuited transmission line. Neglecting end effects, he shows that the inductance and capacitance per unit length of a vertical wire close to earth are given by

$$\left.\begin{aligned} L_0 &= 2\left(\log_e \frac{h}{r} - 1\right) \times 10^{-3}\ \mu\text{H per cm.} \\ C_0 &= \frac{1}{1{\cdot}8\left(\log_e \frac{h}{r} - 1\right)}\ \mu\mu\text{F per cm.} \end{aligned}\right\} \quad . \quad 3.2$$

where h = length of aerial in cms.
and r = radius of wire in cms.
The characteristic or surge impedance of the aerial acting as a transmission line is

$$Z_{0} = \sqrt{\frac{R_0+j\omega L_0}{G_0+j\omega C_0}} \quad . \quad . \quad . \quad . \quad 3.3a$$

where R_0 and L_0 are the resistance and inductance per unit length, and G_0 and C_0 are the conductance and capacitance per unit length. Generally $\omega L_0 \gg R_0$ and $\omega C_0 \gg G_0$ so that

$$Z_0 = \sqrt{\frac{L_0}{C_0}} = 60\left(\log_e \frac{h}{r} - 1\right) = 138\left(\log_{10} \frac{h}{r} - 0{\cdot}435\right) . \quad 3.3b$$

(note that C_0 must be in μF per centimetre when L_0 is in μH per centimetre in the above formula).

Applying normal transmission line procedure, we have for the terminal impedance Z_{a0} of the aerial

$$\begin{aligned} Z_{a0} &= Z_0 \coth \sqrt{(R_0+j\omega L_0)(G_0+j\omega C_0)} \times h \\ &= Z_0 \coth (\alpha+j\beta)h \quad . \quad . \quad . \quad . \quad . \quad . \quad 3.4a \end{aligned}$$

where α = attenuation constant of the aerial

$$= \sqrt{\tfrac{1}{2}[\sqrt{(R_0^2+\omega^2L_0^2)\,(G_0^2+\omega^2C_0^2)}+(G_0R_0 - \omega^2L_0C_0)]}$$

and β = phase constant of the aerial

$$= \sqrt{\tfrac{1}{2}[\sqrt{(R_0^2+\omega^2L_0^2)\,(G_0^2+\omega^2C_0^2)} - (G_0R_0 - \omega^2L_0C_0)]}$$

when $\omega L_0 \gg R_0$ and $\omega C_0 \gg G_0$

$$\alpha = \sqrt{\tfrac{1}{2}(G_0R_0)} \quad . \quad . \quad . \quad . \quad 3.5a$$

and

$$\beta = \omega\sqrt{L_0C_0} \quad . \quad . \quad . \quad . \quad 3.5b.$$

Since the voltages and currents induced in the aerial by the transmitted wave themselves produce electrostatic and electromagnetic waves in space, the resistance term in the above formulae must

include the effect of the radiation resistance noted in expression 3.1*c*. In a well-constructed aerial R_0 and G_0 of the aerial and its insulators are negligible in comparison with the term due to radiation. The method of incorporating radiation resistance is illustrated in a later example. If we neglect the resistance and conductance components, the aerial terminal impedance has only a reactive component which is determined by combining 3.4*a* and 3.5*b*.

Thus
$$Z_{a0} = jX_{a0} = Z_0 \coth j\omega\sqrt{L_0C_0} \times h$$
$$= -jZ_0 \cot \omega\sqrt{(L_0C_0)} \times h$$
$$\therefore X_{a0} = -Z_0 \cot \omega h\sqrt{L_0C_0} \quad . \quad . \quad . \quad . \quad 3.6a.$$

If we assume that there is no retardation of the wave along the wire, the velocity of propagation (the speed of light) v cms. per second is related to $\sqrt{L_0C_0}$ as follows

$$v = \frac{1}{\sqrt{L_0C_0}} = \text{wavelength} \times \text{frequency} = \lambda f = \frac{\lambda\omega}{2\pi} \quad . \quad 3.7$$

$$\therefore X_{a0} = -Z_0 \cot \frac{2\pi h}{\lambda} \quad . \quad . \quad . \quad . \quad . \quad . \quad 3.6b.$$

From 3.6*b* we note that X_{a0} is 0, i.e., the aerial is resonant, when

$$\cot \frac{2\pi h}{\lambda} = 0 \text{ and } \frac{2\pi h}{\lambda} = \frac{\pi}{2}, \frac{3\pi}{2}, \text{ etc.},$$

or
$$h = \frac{\lambda}{4}, \frac{3}{4}\lambda, \text{ etc.} = \frac{n\lambda}{4}$$

where n is an odd positive integer.

Similarly X_{a0} is $-\infty$ when $h = 0, \lambda, 2\lambda$, etc., and is $+\infty$ at $h = \frac{\lambda}{2}, \frac{3\lambda}{2}$, etc. Generally we shall not be interested in the variation of X_{a0} for varying heights of aerial but for a fixed height h_0 and varying frequency. Expression 3.6*b* therefore becomes

$$X_{a0} = -Z_0 \cot \frac{2\pi h_0}{\lambda}$$

where h_0 = actual height of the aerial,

but
$$\frac{2\pi h_0}{\lambda} = \frac{2\pi}{\lambda}\cdot\frac{\lambda_0}{4} = \frac{\pi}{2}\frac{\lambda_0}{\lambda} = \frac{\pi}{2}\cdot\frac{f}{f_0}$$

where λ_0 = natural wavelength of the aerial
and f_0 = natural frequency of the aerial.

$$\therefore X_{a0} = -Z_0 \cot \frac{\pi}{2}\cdot\frac{f}{f_0} \quad . \quad . \quad . \quad . \quad . \quad 3.6c$$

which is negative, i.e., capacitive, when $\frac{f}{f_0} < 1$.

The equivalent terminal capacitance C_{a0} is

$$C_{a0} = \frac{1}{\omega Z_0 \cot(\pi/2.f/f_0)} = \frac{\tan(\pi/2.f/f_0)}{\omega Z_0} \quad . \quad 3.8a$$

but from 3.3*b* and 3.7

$$Z_0 = \sqrt{\frac{L_0}{C_0}} = \frac{1}{vC_0} = \frac{h_0}{vC_1} = \frac{\frac{1}{4}\lambda_0}{f_0\lambda_0 C_1} = \frac{1}{4f_0C_1} \quad . \quad 3.3c$$

where $C_1 = C_0h_0$ = total electrostatic capacitance of the aerial, measured when f is very much less than f_0.

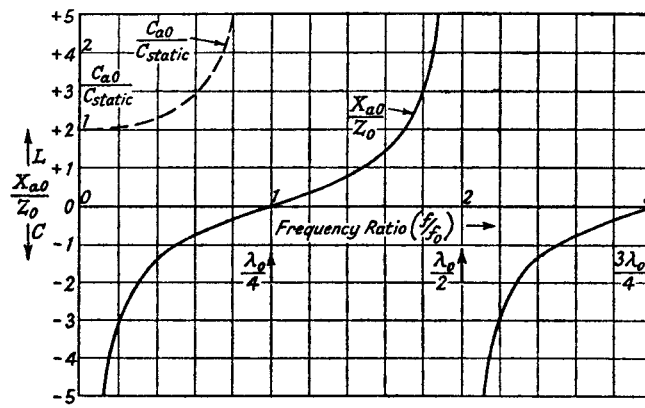


FIG. 3.4*a*.—The Variation of Aerial Terminal Reactance and Capacitance of a Vertical Aerial with Frequency.

Replacing Z_0 in 3.8*a* by 3.3*c*

$$C_{a0} = \frac{4f_0C_1 \tan(\pi/2.f/f_0)}{\omega} = \frac{4f_0C_1}{2\pi f}\tan\left(\frac{\pi}{2}.\frac{f}{f_0}\right)$$

$$= C_1\frac{2}{\pi}.\frac{f_0}{f}\tan\left(\frac{\pi}{2}.\frac{f}{f_0}\right) \quad . \quad . \quad . \quad . \quad . \quad 3.8b$$

when $\frac{f}{f_0}$ is small, $\tan\left(\frac{\pi}{2}.\frac{f}{f_0}\right) \simeq \sin\left(\frac{\pi}{2}.\frac{f}{f_0}\right) \simeq \left(\frac{\pi}{2}.\frac{f}{f_0}\right)$ and

$C_{a0} = C_1$, the electrostatic capacitance of the aerial.

The variation of the ratio of $\frac{C_{a0}}{C_1}$ against $\frac{f}{f_0}$ is shown by the dotted curve in Fig. 3.4*a* over the range $\frac{f}{f_0} = 0$ to 0·8, and the variation of $\frac{X_{a0}}{Z_0}$ (calculated from 3.6*c*) is the full line curve in the

same figure over the range $\frac{f}{f_0} = 0$ to 3. We note that for aerials short compared with the wavelength of the received signal $\left(\frac{f}{f_0} < 0{\cdot}5, \text{ i.e., } h_0 < \frac{\lambda_0}{8}\right)$ the terminal capacitance of the aerial is practically its electrostatic capacitance. Thus over the long and medium wave ranges most aerials can be replaced by a generator having an internal impedance consisting of a resistance (as yet unspecified) in series with the electrostatic capacitance of the aerial.

For $\frac{f}{f_0} > 1$ and < 2, $\cot \frac{\pi}{2} \cdot \frac{f}{f_0}$ is negative so that X_{a0} is positive and therefore inductive. It can be shown by a method similar to that used for calculating C_{a0} that.

$$L_{a0} = -L_1 \frac{\pi}{2} \cdot \frac{f}{f_0} \cot\left(\frac{\pi}{2} \cdot \frac{f}{f_0}\right)$$

$$= L_1 \text{ when } \frac{f}{f_0} \text{ approaches 2.}$$

where $L_1 = L_0 h$ = total inductance of the aerial.

From $\frac{f}{f_0} = 2$ to 3, the aerial is again capacitive and the curve of $\frac{C_{a0}}{C_1}$ is a repetition of that from $\frac{f}{f_0} = 0$ to 1. It is inductive from $\frac{f}{f_0} = 3$ to 4 and the process is repeated as f is increased. The variation of X_{a0} with frequency is that of a cotangent curve and it is plotted in Fig. 3.4*a* as a ratio of $\frac{X_{a0}}{Z_0}$ against $\frac{f}{f_0}$. Reactance is infinite at $\frac{f}{f_0} = 0$ and 2, corresponding to $h_0 = 0$ and $\frac{\lambda}{2}$, between $\frac{f}{f_0} = 0$ and 1 $\left(h_0 = 0 \text{ and } h_0 = \frac{\lambda}{4}\right)$ it is capacitive, being zero at $\frac{f}{f_0} = 1\left(h_0 = \frac{\lambda}{4}\right)$, whilst from $\frac{f}{f_0} = 1$ to 2 $\left(h_0 = \frac{\lambda}{4} \text{ to } \frac{\lambda}{2}\right)$ it is inductive. Above $\frac{f}{f_0} = 2$ it repeats itself periodically between $\frac{f}{f_0} = 2n$ and $2n+2$. If we neglect end effects the primary fundamental frequency, at which an aerial acts as a series resonant circuit

($X_{a0} = 0$), occurs when its height h_0 is equal to a quarter of the wavelength of the exciting frequency, and under these conditions standing quarter waves of E and I are produced on the aerial with maximum current at the base and maximum voltage at the top as shown in Fig. 3.3*b*. In the practical case the electric and magnetic fields do not cease abruptly at the end of the aerial but are projected for a short distance beyond, so that the equivalent electrical height as far as the resonant condition is concerned is rather greater than the actual. A more correct formula for the resonant wavelength and frequency is

$$\lambda_0 = 4{\cdot}2h_0\,;\quad f_0 = \frac{v}{4{\cdot}2h_0} \qquad . \qquad . \qquad . \qquad 3.9$$

and in determining the resonant frequency of any aerial the above formula should be used in preference to $f_0 = \dfrac{v}{4h_0}$.

So far in our discussion we have neglected the resistance component in the aerial terminal impedance. It is not easy to assess because of its dependence on a number of factors, the resistance and conductance characteristic of the aerial, the radiation resistance, and the earth loss due to circulating currents in an imperfectly conducting earth at the base of the aerial. The former, except in badly-erected and indoor aerials running close to earthed conductors, or semiconductors are not often very large and may be neglected, the earth losses are difficult to estimate and depend on site conditions, but the radiation resistance (assuming a perfect earth) has been calculated.

Thus a quarter-wave resonant aerial above a perfectly conducting earth has a terminal impedance which is non-reactive and equal to the radiation resistance, i.e., $36{\cdot}6\Omega$. The radiation resistance [21] for a given height of aerial varies approximately as the square of the signal frequency up to $h = 0{\cdot}4\lambda$, reaches a maximum of 108Ω at $h = 0{\cdot}45\lambda$ and then falls to about 46Ω at $h = 0{\cdot}675\lambda$, passing through 100Ω at $h = 0{\cdot}5\lambda$. To simplify the analysis let us assume that the radiation resistance is 40Ω at $h_0 = \dfrac{\lambda_0}{4{\cdot}2}$ (where h_0 is the physical height) and that radiation resistance is directly proportional to frequency up to $h_0 = \dfrac{\lambda}{2}$. The errors introduced by this simplification are not excessive. We can now illustrate the method of calculating terminal impedance in respect of an aerial of No. 12 S.W.G. copper wire 10 metres (30 ft. approx.) height.

From 3.9 the fundamental wavelength $= 4{\cdot}2 \times 10 = 42$ metres.
,, frequency $= 7{\cdot}14$ Mc/s.
Radius of No. 12 S.W.G. wire $= 0{\cdot}132$ cms.

From 3.2 the capacitance per unit length is

$$C_0 = \frac{1}{1{\cdot}8\left[\log_e \frac{1000}{0{\cdot}132} - 1\right]} \mu\mu\text{F/cm.}$$

$$= 0{\cdot}0705 \ \mu\mu\text{F/cm.}$$

The electrostatic capacitance is $C_0 l = 70{\cdot}5$ $\mu\mu$F.
From 3.3*b*

$$Z_0 = 138 \ (3{\cdot}445) = 475\Omega.$$

Combining expressions 3.4*a*, 3.5*b* and 3.9

$$Z_{a0} = Z_0 \coth \left(\alpha + j\frac{\pi}{2}\right)\frac{f}{f_0} \quad . \quad . \quad . \quad 3.4b.$$

For the quarter wave resonant condition $\frac{f}{f_0} = 1$ and Z_{a0} = radiation resistance $= 40\Omega$.

$$\therefore \ 40 = 475 \coth \left(\alpha + j\frac{\pi}{2}\right)$$

$$= 475 \tanh \alpha$$

or $$\alpha = 0{\cdot}0843.$$

$$\therefore \ Z_{a0} = 475 \coth \left(0{\cdot}0843 + j\frac{\pi}{2}\right)\frac{f}{7{\cdot}14}.$$

Note that

$$\coth(\alpha + j\beta) = \frac{\cosh(\alpha+j\beta)}{\sinh(\alpha+j\beta)} = \frac{\cosh(\alpha+j\beta)\sinh(\alpha-j\beta)}{\sinh(\alpha+j\beta)\sinh(\alpha-j\beta)}$$

$$= \frac{\sinh 2\alpha - \sinh j2\beta}{\cosh 2\alpha - \cosh j2\beta} = \frac{\sinh 2\alpha - j\sin 2\beta.}{\cosh 2\alpha - \cos 2\beta}$$

$$\therefore \ Z_{a0} = R_{a0} + jX_{a0} = Z_0\left[\frac{\sinh 2\alpha}{\cosh 2\alpha - \cos 2\beta} - \frac{j\sin 2\beta}{\cosh 2\alpha - \cos 2\beta}\right] \quad 3.10.$$

A graph showing the variation of R_{a0} and X_{a0} against frequency ratio from 0 to 3 is shown in Fig. 3.4*b*. The curve from 2 to 3 is calculated on the assumption that the radiation resistance falls back again to 40Ω at $\frac{3}{4}\lambda$, so that the value of R_{a0} is slightly lower than would occur in practice. X_{a0} is not appreciably altered by the radiation resistance except in the region of $\frac{f}{f_0} = 2$, at which value it is zero instead of infinite as for Fig. 3.4*a*, and it has a finite

maximum above and below $\frac{f}{f_0} = 2$. A vertical aerial greater than $\frac{\lambda}{4}$ long is seldom used for reception so that the curve from $\frac{f}{f_0} = 0$ to 1 is the important part.

We see that this particular aerial would function satisfactorily over the long and medium wave ranges because the highest received frequency is 1,500 kc/s giving $\frac{f}{f_0} = 0{\cdot}21$, and up to this frequency there is little difference between the terminal and static capacitance. It would not be satisfactory on the short wave range as

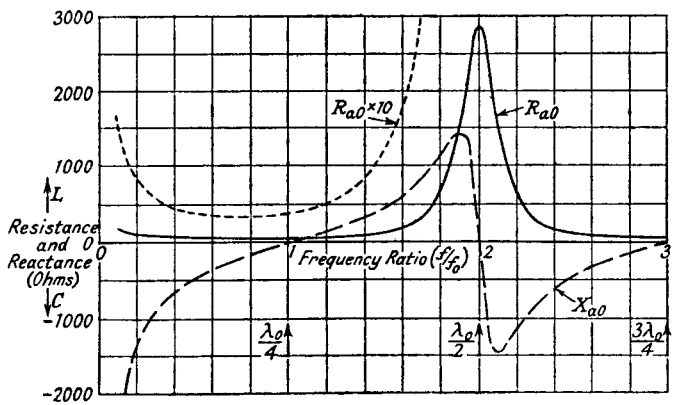


FIG. 3.4*b*.—The Resistance and Reactance Components of the Terminal Impedance of a Vertical Aerial having a Quarter Wavelength Radiation Resistance of 40 Ohms.

it passes through resonance, and would therefore affect considerably the performance of the first tuned circuit, causing large changes of the resistance and reactance components reflected from the aerial into this circuit. The need for keeping aerial resonance outside the desired frequency range is made clear in Section 3.4.

3.3.3. The Inverted L Aerial. The inverted L aerial is formed by adding a horizontal top to a vertical aerial. The addition of the horizontal top affects both generated voltage and terminal impedance : it is equivalent to adding a capacitance from the top of the aerial to earth, see C_1 in Fig. 3.3*a*. With vertically polarized transmission no voltage is induced in the horizontal top, but its capacitance to earth makes the voltage induced in each vertical section more effective, e.g., the top section generator is now operative because the horizontal top capacitance completes its return

path to earth. Similarly, succeeding generators operate more efficiently and the average effective voltage becomes greater than $\frac{Eh}{2}$, i.e., the effective height of the aerial is increased. An inverted L aerial having a ratio of horizontal to vertical length of unity has an effective height of the order of $0{\cdot}6h$.

The terminal impedance of an L aerial is calculated by considering vertical and horizontal sections separately. The horizontal section is treated as an open circuited transmission line parallel to earth, whilst the vertical part is treated as a line terminated by the terminal impedance of the horizontal part. The characteristic impedance Z_{0h} of this part is

$$Z_{0h} = 60\left(\log_e \frac{2h}{r} - 1\right) = 60\left(\log_e \frac{h}{r} + \log_e 2 - 1\right) \quad . \quad 3.11$$

and if $h \gg r$ we see that Z_{0h} is very nearly equal to Z_{0v} (expression 3.3*b*) when the radii of the horizontal and vertical sections are equal. The terminal or input impedance of the horizontal section is

$$Z_h = Z_{0h} \coth (\alpha_h + j\beta_h)l = Z_{0h} \coth \gamma_h l$$

where l = length of horizontal section

α_h and β_h = attenuation and phase constant of the section

and $\gamma_h = \alpha_h + j\beta_h$.

The vertical section is terminated by Z_h and applying the above nomenclature with the suffix "v" replacing "h" we have, by normal transmission line theory, for the terminal impedance of the vertical part

$$Z_{a0} = Z_{0v}\left[\frac{Z_h \cosh \gamma_v h + Z_{0v} \sinh \gamma_v h}{Z_h \sinh \gamma_v h + Z_{0v} \cosh \gamma_v h}\right] \quad . \quad 3.12a$$

where h = height of vertical section.

If $\alpha_v \ll \beta_v$ and $\alpha_h \ll \beta_h$.

$$Z_h = -jZ_{0h} \cot \beta_h l.$$

and $$jX_{a0} = Z_{0v}\left[\frac{-jZ_{0h} \cot \beta_h l. \cos \beta_v h + jZ_{0v} \sin \beta_v h}{Z_{0h} \cot \beta_h l. \sin \beta_v h + Z_{0v} \cos \beta_v h}\right] \quad 3.12b.$$

For aerial resonance $X_{a0} = 0$

i.e., $$\tan \beta_h l. \tan \beta_v h = \frac{Z_{0h}}{Z_{0v}} \quad . \quad . \quad 3.13.$$

If we assume $Z_{0h} = Z_{0v}$, we have that $L_{0h} = L_{0v}$, $C_{0h} = C_{0v}$ and $\beta_h = \beta_v$.

Hence expression 3.13 becomes

$$\tan \beta_v l. \tan \beta_v h = 1$$

$$\sin \beta_v l. \sin \beta_v h = \cos \beta_v l. \cos \beta_v h.$$

or $$\cos \beta_v (h+l) = 0$$

giving $$(h+l)\beta_v = \frac{\pi}{2}$$

so that from 3.5*b* and 3.7. $$(h+l) = \frac{\lambda_0}{4} \quad . \quad . \quad . \quad . \quad 3.14.$$

Similarly, expression 3.12*a* reduces to

$$X_{a0} = -Z_{0v}\left[\frac{\cot \beta_v l. \cos \beta_v h - \sin \beta_v h.}{\cot \beta_v l. \sin \beta_v h + \cos \beta_v h}\right].$$

$$= -Z_{0v}\left[\frac{\cos \beta_v l. \cos \beta_v h - \sin \beta_v h. \sin \beta_v l.}{\cos \beta_v l. \sin \beta_v h + \cos \beta_v h. \sin \beta_v l.}\right].$$

$$= -Z_{0v} \cot \beta_v (h+l) \quad . \quad . \quad . \quad . \quad . \quad 3.15.$$

We therefore see that an inverted L aerial erected well clear of earth has practically the same terminal impedance characteristics as a vertical aerial of height equal to the overall length of the L aerial.

3.3.4. The T Aerial. There is little essential difference in operation between the T and inverted L aerial. The split horizontal section acts as a capacitance to earth, increasing the effective height of the aerial. Terminal impedance is calculated in the same manner as for the inverted L aerial, but we must note that the impedance at the top of the vertical part consists of the impedances of the two horizontal sections in parallel, i.e., $Z_h = \frac{Z_{h1}.Z_{h2}}{Z_{h1}+Z_{h2}}$. Generally $Z_{h1} = Z_{h2}$, and if the T top has the same wire diameter as the vertical section and is not very close to earth

$$Z_h = \frac{Z_{h1}}{2} \simeq \frac{Z_{0v}}{2} \coth (\alpha_v + j\beta_v)\frac{l}{2}$$

where l is the total horizontal length of the T top.

3.3.5. The Dipole Aerial. The dipole (sometimes called the doublet) aerial consists of two equal lengths of open-ended wires connected at their near ends to a feeder balanced to earth. The two wires are in the same plane and at 180° to each other. Such an aerial is very suitable for short wave reception and it has the advantage that if the aerial itself is balanced to earth, i.e., erected horizontally, local interference effects tend to cancel out. Owing to the use of the reflected ray for short wave transmission, the wave

arriving at the aerial has an appreciable horizontally polarized component and this is picked up by the horizontally disposed dipole. The horizontal disposition gives the aerial a directive figure-of-eight diagram in the horizontal plane with maximum pick-up in the two directions perpendicular to the dipole. A vertically erected dipole has no directional effect in the horizontal plane and is unbalanced to earth so that local interference voltages are not cancelled to the same degree.

The terminal impedance of a dipole aerial can be calculated by the method employed for the vertical aerial. Let us assume that we have a horizontal dipole aerial, split at the centre to form two wires, each of 12 S.W.G. copper 7 metres long (Fig. 3.5 inset).

The fundamental wavelength is $4{\cdot}2 \times 7 = 29{\cdot}4$ metres

" " frequency is 10·2 Mc/s

which is in the centre of the short wave band (6 to 15 Mc/s).

The capacitance per unit length between each wire of the dipole when the latter is at least $\frac{\lambda}{2}$ (14 metres) from earth (the image effect can then be neglected) is [1]

$$C = \frac{1}{3{\cdot}6\left(\log_e \frac{l}{r} - 1\right)} \ \mu\mu\text{F/cm.}$$

where l = length of each wire

r = radius of wire = 0·132 cms.

$$C = \frac{1}{3{\cdot}6\left(\log_e \frac{700}{0{\cdot}132} - 1\right)} = 0{\cdot}0366 \ \mu\mu\text{F/cm.}$$

$$Z_0 = 120\left(\log_e \frac{700}{0{\cdot}132} - 1\right) = 910\Omega.$$

The radiation resistance of a $\frac{\lambda}{2}$ dipole split at the centre is twice that of the vertical $\frac{\lambda}{4}$ aerial, i.e., $72{\cdot}2\Omega$, and we will assume that the terminal impedance under these conditions is 80Ω.

Thus

$$80 = Z_0 \coth\left(\alpha + j\frac{\pi}{2}\right)$$

$$\tanh \alpha = \frac{80}{910} = 0{\cdot}0878$$

$$\alpha = 0{\cdot}0880$$

from 3.4*b*
$$Z_{a0} = Z_0 \coth\left(\alpha + j\frac{\pi}{2}\right)\frac{f}{10{\cdot}2}$$
$$= 910 \coth\left(0{\cdot}0880 + j\frac{\pi}{2}\right)\frac{f}{10{\cdot}2}.$$

The values of R_{a0} and X_{a0} are plotted against $\frac{f}{f_0} = 0$ to 3 in Fig. 3.5

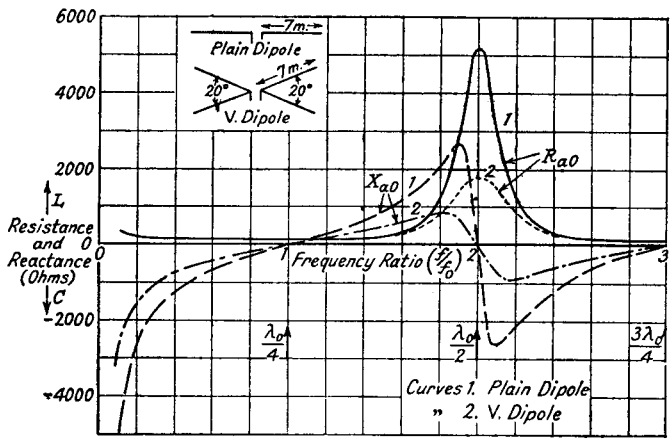


FIG. 3.5.—The Variation of Terminal Resistance and Reactance of a Plain and V Dipole having a Quarter Wavelength Radiation Resistance of 80 Ohms.
(The wavelength scale is referred to the length of one-half of the dipole.)

(curves 1). The variation of terminal impedance over the normal short wave range from 6 to 15 Mc/s (approximately from $\frac{f}{f_0} = 0{\cdot}6$ to 1·5) is considerable, though its effect on the receiver is generally reduced by the feeder connection, which is usually employed. Mismatching between feeder and aerial occurs and there is loss of efficiency in the power transfer at the junction. The loss of power, or transition loss as it is called, is discussed more fully in Section 3.5.3. The variation in terminal impedance can be reduced if the characteristic impedance [8] of the aerial is reduced. A decrease in Z_0 has no effect at $\frac{f}{f_0} = 1$, since the terminal impedance is determined by radiation resistance alone, but it has a large effect at $\frac{f}{f_0} = 2$. Then each half acts as a $\frac{\lambda}{2}$ aerial and

$$Z_{a0}\left(\frac{\lambda}{2}\right) = Z_0 \coth 2\alpha \simeq \frac{Z_0}{2\alpha} \text{ when } \alpha \text{ is small}$$

i.e.,
$$Z_{a0}\left(\frac{\lambda}{2}\right) = \frac{{Z_0}^2}{2R_r\left(\frac{\lambda}{4}\right)}.$$

Hence a small reduction of characteristic impedance reduces considerably Z_{a0}, e.g., halving Z_0 reduces Z_{a0} to 25% of its original value. This reduction of Z_0 can be achieved by making each half of the dipole into a V aerial [5] as shown by the inset in Fig. 3.5.

The capacitance $\mu\mu$F per cm. of this type of aerial is given by [2]

$$C = \frac{1}{1{\cdot}8\left[\log_e \frac{l}{r} - 1 + \log_e\left(1+\sqrt{1+(\operatorname{cosec}\theta+\cot\theta)^2}\right) - \log_e\left(1+\sqrt{1+[\operatorname{cosec}(180-\theta)+\cot(180-\theta)]^2}\right)\right]}$$

where θ = angle between the V.

If this angle is 20° and all other details the same as for the single dipole, the V dipole has a capacitance per centimetre of

$$C_0 = 0{\cdot}063 \ \mu\mu\text{F}.$$

and

$$Z_0 = 526\Omega.$$

Hence

$$\tanh\alpha = \frac{80}{526} = 0{\cdot}152$$

and

$$Z_{a0} = 526 \coth\left(0{\cdot}153 + j\frac{\pi}{2}\right)\frac{f}{10{\cdot}2}.$$

R_{a0} and X_{a0} are plotted as curves 2 in Fig. 3.5, and the reduction in terminal impedance variation is quite marked.

Since the curves in Fig. 3.5 are calculated on the assumption that R_r is directly proportional to frequency up to $\frac{f}{f_0} = 2$, the values of R_{a0} are not strictly correct, though they form a useful basis for the design of the aerial-to-receiver connection. The tendency is for the curves of R_{a0} to be high for values of $\frac{f}{f_0} < 1$ and low for values of $\frac{f}{f_0}$ from 1 to 1·8. A more accurate method of calculating aerial terminal impedance for many types of aerial is to be found in Bibliography 18.

The balanced horizontal dipole is an inefficient collector of signals in the long and medium wave ranges, for which transmission is mainly vertically polarized. If it is to be used for "all wave" reception, provision must be made to change its method of operation at lower frequencies. A method [5] of realizing this is shown in Fig. 3.6. The dipole is connected via a parallel wire feeder, crossed over at intervals, to a transformer at the receiver. The centre tap of the transformer primary is connected to earth by a series *LCR* circuit. One end of the secondary of the transformer is connected

to the aerial terminal of the receiver and the other to the capacitance C of the primary centre tap earth circuit. Over the short wave range the aerial acts as a dipole, the voltage being developed across the primary. The filtering action of the LCR circuit prevents voltages at short wave frequencies (desired or local interference) from appearing across C. An electrostatic screen is included between the primary and secondary to prevent interference currents passing to the unbalanced secondary. Local interference currents developed in each side of the aerial and feeder system cancel in the transformer primary and flow to earth through the LCR circuit. They tend to produce a voltage across C, but it is reduced to small proportions by the addition of the inductance L. As the desired frequency is decreased the reactance of C increases and the system begins to function as a T aerial, the two halves of aerial and feeder being considered as in parallel and the pick-up voltage appearing across C. With this method of operation there is practically no local interference protection at the lower frequencies, and some improvement may be achieved at the expense of signal strength by using a twin feeder enclosed in an earthed shield.

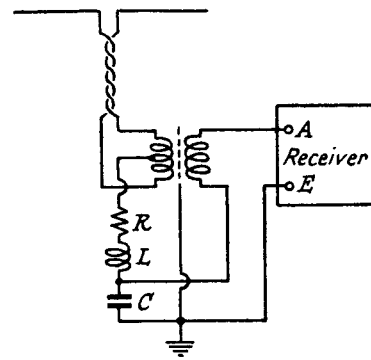


FIG. 3.6.—A Balanced Horizontal Dipole for All Wave Reception.

More complicated methods of making the dipole an all-wave collector have been employed, but the principle involved is the same, viz., the aerial acts as a dipole for short waves and then changes to a T or L aerial for long and medium waves. When a dipole aerial is intended for operation over a very limited range of frequencies, e.g., for television reception, one arm of the dipole may be connected through a parallel resonant wave trap circuit [19] (tuned to the centre of the short or ultra short wave band accepted by the dipole) to a long single wire, which acts as the aerial for the longer short wave ranges and the long and medium wave bands.

If increased signal is desired from a given direction, a reflector, often a similar dipole with the two halves either short circuited or connected together with an inductance or capacitance, is used. The inductance lowers and the capacitance raises the $\frac{\lambda}{2}$ resonant

frequency. The reflector is placed parallel to the first dipole at a distance of $\frac{\lambda}{8}$ to $\frac{\lambda}{4}$ behind it, looking from the direction of the desired transmission source. The effect of the reflector is also to suppress reception from sources behind the main aerial, the directional diagram being heart shaped with maximum pick up at right angles to, and in front of the first dipole. If the reflector is of different physical or electrical length (this is effected by inductance or capacitance at its centre) compared with the main aerial, a more constant response [20] can often be realized over a given frequency range.

3.3.6. The Frame Aerial. The frame or closed loop aerial is an inefficient collector as compared with the normal open aerial, and its chief advantage is its directional property. It operates for vertically polarized transmission only when there is a phase difference between the voltages induced in the two vertical sides. These voltages are equal and cancel each other when the plane of the loop is parallel to the electrostatic component of the transmitted field and perpendicular to its direction of travel. When the frame is not perpendicular to the motion of the field, there is a phase difference between the voltages induced in the two limbs, that induced in the limb nearest the source of the wave leading on the voltage in the other limb. This means that at any given instant the voltages are unequal and there is a net voltage to drive current round the loop. The phase angle between the two voltages is the distance in radians between the projections of the two vertical limbs on to a plane parallel to the motion of the wave. If λ is the wavelength of the desired signal, b the breadth of the frame and α the angle between the plane of the frame and the motion of the wave (Fig. 3.7*a*), the phase angle is $\frac{2\pi}{\lambda} b \cos \alpha$. The phase angle is clearly a maximum when α is 0, i.e., when the frame is parallel to the motion of the wave, and in this position the pick-up is maximum. The maximum effective voltage induced in the frame is $nEh \sin \frac{2\pi b}{\lambda}$ where n is the number of turns in the frame, E is the field strength of the signal at the frame and h is the length of a vertical side. When $\frac{2\pi b}{\lambda}$ is small the induced voltage is

$$\frac{nEh2\pi b}{\lambda} = \frac{2\pi nE}{\lambda} \times \text{area of frame} \quad . \quad . \quad 3.16.$$

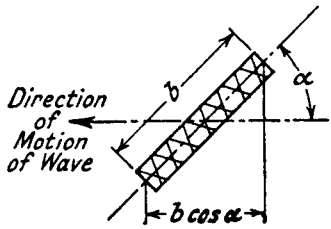


FIG. 3.7.*a*—Plan View of a Frame Aerial.

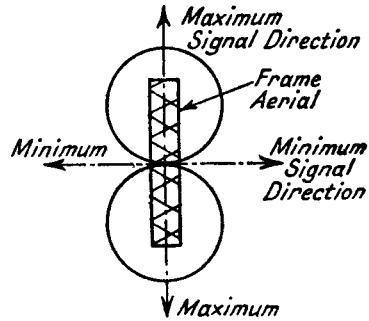


FIG. 3.7*b*.—The Directional Diagram of a Frame Aerial.

The directional diagram of the frame is a figure-of-eight with minimum from directions at right angles on either side of the frame and maximum from the two end-on positions (Fig. 3.7*b*).

3.4. The Coupling between the Aerial and Receiver.[3, 17]

3.4.1. Introduction. An aerial may be coupled to the first tuned circuit of a receiver by inductance and capacitance, separately or combined. Owing to this coupling, a resistance and reactance component is reflected from the aerial into the tuned circuit; this reduces its selectivity and also requires the tuning capacitor setting to be changed if resonance is to be maintained, a disadvantage when the circuit is ganged with succeeding tuned circuits. The object of the coupling is therefore to obtain maximum voltage transfer with minimum effect on the tuned circuit. For any given aerial and tuned circuit conditions, there is always an optimum coupling giving greatest voltage transfer, and if coupling is increased beyond this point voltage transfer falls and the reflected impedance effect from the aerial increases. Hence it is most undesirable to exceed optimum coupling, and indeed it is preferable to use couplings much less than critical since voltage transfer falls at a much slower rate than the reflected aerial impedance. For couplings less than optimum, maximum voltage transfer is realised by adjusting the tuning capacitance for resonance with the tuning coil and added reactance from the aerial.

In the analysis, which follows, the aerial circuit is assumed to consist of a generator of voltage E_1 with an internal impedance equal to the terminal impedance Z_{a0} of the aerial. This impedance is considered as a resistance R_{a0} in series with a reactance jX_{a0}; for most normal aerials the reactance is capacitive on the long and medium wave bands. The fundamental wavelength of an aerial is

given by 4·2 × height, and at this frequency it is resonant and its terminal impedance is resistive only. At frequencies greater than this the reactance becomes inductive. For resonance in the medium wave band (maximum frequency = 1,500 kc/s) the height or length of the aerial would have to exceed 47 metres, a condition hardly likely to be realized in practice. On the short wave range the resonant point may be reached and passed.

The first form of coupling to be considered is by mutual inductance, as it is possible to develop from this generalized formulae applicable to all types of coupling.

3.4.2. Mutual Inductance Coupling. Coupling between the aerial and first tuned circuit of a receiver is quite commonly effected by mutual inductance between a primary coil, to which the aerial is connected, and a secondary coil, which is the inductance element

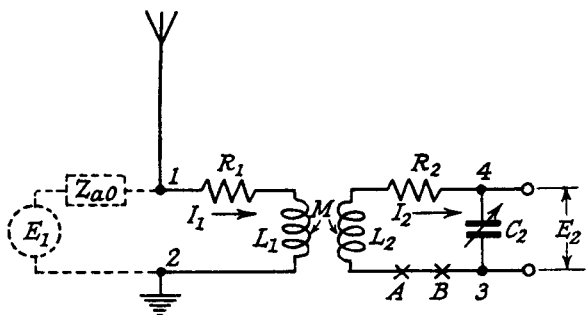


FIG. 3.8*a*.—An Aerial Circuit with Mutual Inductance Coupling.

L_2, of the first tuned circuit (Fig. 3.8*a*). Certain conventions and terms are used and these will first be stated.

The aerial is considered as a generator of voltage E_1 of internal impedance $Z_{a0} = R_{a0} + jX_{a0}$ equal to its terminal impedance.

The series impedance of the aerial terminal impedance and the coupling coil is designated as

$$Z_{a1} = R_{a1} + jX_{a1} = R_{a0} + R_1 + j(X_{a0} + \omega L_1)$$

where R_1 and L_1 are the resistance and inductance of the primary coil. Hence Z_{a1} is the impedance looking from the aerial input voltage E_1 with the tuning coil on open circuit.

The series impedance of the tuned secondary circuit with the aerial disconnected is designated as

$$Z_2 = R_2 + jX_2 = R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right)$$

Transfer Voltage Ratio T_R is the ratio of the voltage E_2 developed across the tuning capacitance C_2 to the aerial generated voltage E_1.

To express selectivity, we will use a term conveniently called the Selectivity Ratio defined as

$$S_R = \frac{R_2}{R_2+R_c+R_{ar}} = \frac{1}{1+\dfrac{R_c+R_{ar}}{R_2}}$$

where R_2 = resistance of the secondary coil
R_c = resistance (if any) of the coupling element
R_{ar} = resistance reflected into the secondary circuit from the aerial.

It is very nearly the ratio of the Q value of the secondary circuit, with the aerial connected, to that of the secondary coil alone. The coupling circuit resistance R_c (it is zero for mutual inductance but not necessarily for other forms of coupling) is excluded from the numerator since it can safely be assumed that the coupling element would not be included unless the aerial connection were required. The maximum value of selectivity ratio is 1, when the selectivity is that of the tuned circuit alone, and for all couplings it is less than 1.

Mistuning is definable in two forms, both of which are useful; the first states it as the capacitance correction, ΔC_2, required to maintain resonance of the secondary circuit. It is the difference in capacitance between C_2, the final tuning capacitance setting, and C_{20}, the initial tuning capacitance satisfying $C_{20} = \dfrac{1}{\omega^2 L_2}$. A knowledge of ΔC_2 indicates how far the ganging error can be corrected, e.g., if ΔC_2 is constant over a tuning range and within the range of the tuning trimmer capacitor, the ganging error can be reduced to zero. The second defines mistuning in the ratio form of $\dfrac{\Delta C_2}{C_{20}}$, and from this the frequency mistune ratio, the ratio of the frequency difference between the desired signal frequency and the actual resonant frequency of the secondary circuit (coupled to the aerial without correcting for reflected aerial and coupling reactance) to the desired signal frequency can be calculated. If $\Delta C_2 \ll C_{20}$, the frequency mistune ratio is half the capacitance mistune ratio, i.e.,

$$\frac{f_1 - f_2}{f_1} = \frac{\Delta C_2}{2C_{20}} = \frac{M_R}{2}$$

where f_1 = desired signal frequency, i.e., the resonant frequency of the uncoupled secondary circuit
f_2 = resonant frequency of the secondary coupled to the aerial
and M_R = mistune ratio.

The sign of mutual inductance is given with reference to the common limb of the equivalent T network in Fig. 3.8*b*. Thus positive M gives $(L_1 - M)$ and $(L_2 - M)$, whilst negative M gives (L_1+M) and (L_2+M) as the series arms. Referring to Fig. 3.8*a*, it means that M is positive if by joining 2 and 3 a measurement of

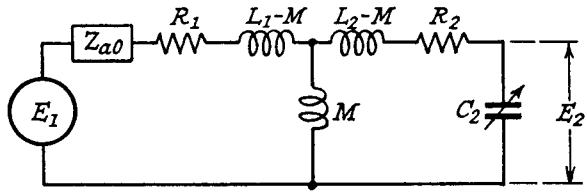


FIG. 3.8*b*.—The Equivalent Circuit for Mutual Inductance Aerial Coupling.

total inductance across 1 and 4 gives $L_T = L_1+L_2 - 2M$. Actually the sign of M is only important when additional coupling, e.g., by capacitance, is employed.

The current and voltage relationships from Fig. 3.8*a*, the actual circuit, are

$$E_1 = I_1[Z_{a0}+R_1+j\omega L_1]+I_2 j\omega M$$

$$0 = I_1 j\omega M+I_2\left(R_2+j\omega L_2+\frac{1}{j\omega C_2}\right)$$

$$E_2 = \frac{I_2}{j\omega C_2}.$$

Solving for the transfer voltage ratio

$$T_R = \frac{E_2}{E_1} = \frac{\dfrac{M}{C_2}}{Z_{a1}Z_2+\omega^2M^2}$$

$$= \frac{\dfrac{M}{C_2}}{(R_{a1}R_2 - X_{a1}X_2+\omega^2M^2)+j(X_2R_{a1}+X_{a1}R_2)} \quad . \quad 3.17.$$

Assuming L_1 to be fixed, there are two possible variables in 3.17, C_2 and M. In most practical circuits M is preset, but C_2 is variable over a wide range. In the absence of ganging requirements C_2 would be adjusted for maximum T_R, and this occurs when the total series reactance of the secondary circuit, including that reflected from the aerial, is zero, i.e., when the secondary circuit is resonant. If M also is variable, a maximum value of T_R, which we will call the optimum, is found.

To determine the resonant condition, the secondary circuit must be opened at points such as AB and E_1 must be short circuited.

The ratio of an applied voltage E (of any value) across these points to the current I it produces gives Z_{2T}, the secondary circuit equivalent series impedance.

Thus
$$Z_{2T} = \frac{E}{I} = R_2 + j\omega L_2 + \frac{1}{j\omega C_2} + \frac{\omega^2 M^2}{Z_{a1}}$$
$$= R_2 + \frac{\omega^2 M^2 R_{a1}}{|Z_{a1}|^2} + j\left[X_2 - \frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2}\right] \quad . \quad 3.18.$$

The resistance and reactance reflected from the aerial are

$$\frac{\omega^2 M^2 R_{a1}}{|Z_{a1}|^2} \text{ and } -\frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2} \text{ respectively.}$$

For secondary circuit resonance

$$X_2 = \frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2} \quad . \quad . \quad . \quad . \quad 3.19.$$

Replacing X_2 in 3.17 by this value

$$\frac{E_2}{E_1}(\text{max.}) = \frac{\frac{M}{C_2}}{\left[R_{a1}R_2 - \frac{\omega^2 M^2 X_{a1}{}^2}{|Z_{a1}|^2} + \omega^2 M^2\right] + j\left[\frac{\omega^2 M^2 X_{a1} R_{a1}}{|Z_{a1}|^2} + X_{a1}R_2\right]}$$
$$= \frac{\frac{M}{C_2}}{Z_{a1}\left[R_2 + \frac{\omega^2 M^2 R_{a1}}{|Z_{a1}|^2}\right]} \quad . \quad . \quad . \quad . \quad . \quad . \quad 3.20a.$$

For optimum $\frac{E_2}{E_1}$ we must differentiate this expression with respect to M and equate to zero. This gives

$$M = \frac{Z_{a1}}{\omega}\sqrt{\frac{R_2}{R_{a1}}} \quad . \quad . \quad 3.21$$

and
$$\frac{E_2}{E_1}(\text{opt.}) = \frac{M}{2C_2 Z_{a1} R_2} \quad . \quad . \quad 3.20b$$
$$= \frac{1}{2\omega C_2 \sqrt{R_2 R_{a1}}} \quad . \quad . \quad 3.20c.$$

$$\text{Selectivity Ratio} \left(\text{max.}\frac{E_2}{E_1}\right) = \frac{R_2}{R_2 + \frac{\omega^2 M^2 R_{a1}}{|Z_{a1}|^2}} \quad . \quad . \quad 3.22a$$

$$S_R\left(\text{opt.}\frac{E_2}{E_1}\right) = \frac{1}{2} \quad . \quad . \quad . \quad . \quad 3.22b.$$

Mistuning is calculated from 3.19 thus

$$X_2 = \omega L_2 - \frac{1}{\omega C_2} = \frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2}$$

$$C_2 = \frac{1}{\omega^2 L_2 \left[1 - \frac{\omega^2 M^2 X_{a1}}{\omega L_2 |Z_{a1}|^2}\right]} \quad . \quad . \quad 3.23a$$

$$= \frac{C_{20}}{1 - \frac{\omega^2 M^2 X_{a1}}{\omega L_2 |Z_{a1}|^2}} \quad . \quad . \quad . \quad 3.23b.$$

Capacitance Correction $\Delta C_2 = C_2 - C_{20}$

$$= \frac{C_{20} \cdot \frac{\omega^2 M^2 X_{a1}}{\omega L_2 |Z_{a1}|^2}}{1 - \frac{\omega^2 M^2 X_{a1}}{\omega L_2 |Z_{a1}|^2}} \quad . \quad . \quad 3.24a$$

$$= \frac{C_{20}}{\frac{\omega L_2 |Z_{a1}|^2}{\omega^2 M^2 X_{a1}} - 1} \quad . \quad . \quad 3.24b.$$

Mistune Ratio $M_R = \frac{\Delta C_2}{C_{20}} = \frac{1}{\frac{\omega L_2 |Z_{a1}|^2}{\omega^2 M^2 X_{a1}} - 1} \quad . \quad . \quad 3.25.$

From expression 3.23*b* it is clear that when X_{a1} is positive, i.e., inductive, C_2 is greater than C_{20}. The reactance reflected into the secondary circuit from the aerial is equivalent to a negative inductance of $\frac{-\omega M^2 X_{a1}}{|Z_{a1}|^2}$ in series with the secondary coil. The reverse is true for capacitive X_{a1}, C_2 is less than C_{20}, and the reflected reactance is equivalent to positive inductance. Over the long and medium wave ranges, the aerial terminal reactance, X_{a0}, is almost certain to be capacitive, so that for X_{a1} to be inductive ωL_1 must be greater than X_{a0}, i.e., a large primary coil is required.

For optimum coupling

$$\Delta C_2 = \frac{C_{20}}{\frac{\omega L_2 R_{a1}}{X_{a1} R_2} - 1} \quad . \quad . \quad . \quad 3.24c.$$

Optimum coupling is rarely employed because M requires to be changed as the tuning frequency is varied $\left(M_{opt} = \frac{Z_{a1}}{\omega}\sqrt{\frac{R_2}{R_{a1}}}\right)$

unless Z_{a1} is proportional to ω (that is preponderatingly inductive) and the ratio $\frac{R_2}{R_{a1}}$ is constant. An added disadvantage is that selectivity ratio is halved and mistuning is large. The most important case is for a fixed value of M, and the expressions 3.20*a*, 22*a*, 24*b* and 25 are more useful.

By combining 3.20*a* and 23*a* we have

$$T_R = \frac{E_2}{E_1} = \frac{\omega M\left(\omega L_2 - \frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2}\right)}{Z_{a1}\left[R_2 + \frac{\omega^2 M^2 R_{a1}}{|Z_{a1}|^2}\right]} \quad . \quad . \quad 3.26a$$

$$= \frac{\omega M}{Z_{a1}} Q_{21} \quad . \quad . \quad . \quad . \quad . \quad 3.26b$$

where Q_{21} = magnification of the secondary circuit when the aerial is coupled to it

$$= \frac{\text{coil reactance} + \text{reflected aerial reactance}}{\text{coil resistance} + \text{reflected aerial resistance}}.$$

A useful approximate formula is applicable when the coupling is loose, i.e., $M \ll \frac{Z_{a1}}{\omega}\sqrt{\frac{R_2}{R_{a1}}}$ for then $Q_{21} \simeq Q_2$, the magnification of the uncoupled secondary coil and

$$T_r = \frac{\omega M}{Z_{a1}} Q_2 = \frac{\text{coupling impedance}}{\text{aerial circuit impedance}} \times Q_2 \quad . \quad 3.26c.$$

In a similar way 3.24*a* may be written as

$$\Delta C_2 = \frac{C_{20}.\ \text{reflected aerial reactance}}{\text{coil reactance} + \text{reflected aerial reactance}} \quad . \quad 3.27.$$

Let us now consider combined mutual inductance and resistance coupling. Such a circuit (Fig. 3.9) is practically never used, but it is valuable in studying the effects of a more general form of coupling.

3.4.3. Combined Mutual Inductance and Resistance Coupling. The coupling impedance is $R + j\omega M$ and

$$T_R = \frac{\frac{R + j\omega M}{j\omega C_2}}{Z_{a1}Z_2 - (R + j\omega M)^2}$$

where $Z_{a1} = R_{a0} + R_1 + R + j(X_{a0} + \omega L_1)$

and $Z_2 = R_2 + R + j\left(\omega L_2 - \frac{1}{\omega C_2}\right)$.

The impedance reflected from the aerial into the tuned circuit is

$$\frac{-(R+j\omega M)^2}{Z_{a1}} = \frac{-(R+j\omega M)^2(R_{a1}-jX_{a1})}{|Z_{a1}|^2}$$

$$= \frac{(\omega^2M^2-R^2)R_{a1}-2R\omega MX_{a1}}{|Z_{a1}|^2} - \frac{j[X_{a1}(\omega^2M^2-R^2)+2\omega MRR_{a1}]}{|Z_{a1}|^2}.$$

Generally $\omega M \gg R$ and $X_{a1}\omega M \gg 2RR_{a1}$ so that the reflected aerial impedance is

$$Z_{a1} = \frac{\omega^2M^2R_{a1}-2R\omega MX_{a1}}{|Z_{a1}|^2} - \frac{j\omega^2M^2X_{a1}}{|Z_{a1}|^2}.$$

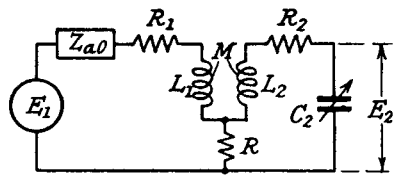


FIG. 3.9.—An Aerial Circuit with Combined Mutual Inductance and Resistance Coupling.

The resistance component is appreciably affected by the addition of R, but the reactance is almost unchanged. This means that T_R and S_R, but not ΔC_2 or M_R, are changed. The expression for T_R is

$$T_R = \frac{R+j\omega M}{j\omega C_2 Z_{a1}\left[R_2+R+\frac{\omega^2M^2R_{a1}}{|Z_{a1}|^2}-\frac{2R\omega MX_{a1}}{|Z_{a1}|^2}\right]} \quad . \quad 3.28$$

and

$$S_R = \frac{R_2}{R_2+R+\frac{\omega^2M^2R_{a1}}{|Z_{a1}|^2}-\frac{2R\omega MX_{a1}}{|Z_{a1}|^2}} \quad . \quad . \quad . \quad 3.29.$$

This analysis may be used as a basis for developing generalized formulæ applicable to any type of coupling which exists as, or is convertible into, a T section network.

3.4.4. Generalized Formulæ for T_R, S_R, ΔC_2 and M_R. The generalized form of T network for coupling common to aerial and secondary circuits is that of Fig. 3.10, and it is identical with Fig. 3.9, when

$$Z_\alpha+Z_\beta = Z_{a1} = R_{a0}+R_1+R+j(X_{a0}+\omega L_1)$$
$$Z_\beta = R+j\omega M$$
$$Z_\Delta = R_2+j\omega(L_2-M)$$
$$Z_\lambda = \frac{-j}{\omega C_2}.$$

By assuming that $Z_\alpha = R_\alpha + jX_\alpha$, etc., the generalized formula for transfer voltage ratio is from 3.28.

$$T_R = \frac{Z_\beta Z_\lambda}{(Z_\alpha+Z_\beta)\left[R_\beta+R_\Delta+R_\lambda+\frac{X_\beta^2(R_\alpha+R_\beta)}{|(Z_\alpha+Z_\beta)|^2}-\frac{2R_\beta X_\beta(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}\right]} \quad 3.30a$$

where $|(Z_\alpha+Z_\beta)|^2 = (R_\alpha+R_\beta)^2+(X_\alpha+X_\beta)^2$

and the formula for selectivity ratio is

$$S_R = \frac{R_\beta}{R_\beta+R_\Delta+R_\lambda+\frac{X_\beta^2(R_\alpha+R_\beta)}{|(Z_\alpha+Z_\beta)|^2}-\frac{2R_\beta X_\beta(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}} \quad 3.31a.$$

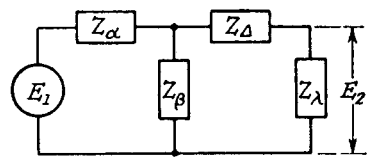


FIG. 3.10.—A Generalized T Network Coupling Circuit.

To obtain the generalized expressions for capacitance correction and mistune ratio we must return to 3.23*a*.

$$C_2 = \frac{1}{\omega\left(\omega L_2 - \frac{\omega^2 M^2 X_{a1}}{|Z_{a1}|^2}\right)} = \frac{1}{\omega\left(X_\beta+X_\Delta-\frac{X_\beta^2(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}\right)}$$

but $\frac{1}{\omega} = \omega L_2 C_{20}$.

$$\therefore C_2 = C_{20}\frac{\omega L_2}{X_\beta+X_\Delta-\frac{X_\beta^2(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}}$$

$$\Delta C_2 = C_2 - C_{20} = C_{20}\left[\frac{\omega L_2}{X_\beta+X_\Delta-\frac{X_\beta^2(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}}-1\right] \quad 3.32a$$

$$M_R = \frac{\Delta C_2}{C_{20}} = \left[\frac{\omega L_2}{X_\beta+X_\Delta-\frac{X_\beta^2(X_\alpha+X_\beta)}{|(Z_\alpha+Z_\beta)|^2}}-1\right] \quad 3.33a.$$

Simplification of the above expressions is possible by combining some of the terms into a single term as follows:

Let $R_{a1} = R_\alpha + R_\beta$ the series resistance of the aerial and coupling circuit with the secondary coil disconnected.

$X_{a1} = X_\alpha + X_\beta$ the series reactance of the same circuit.

$R_{21} = R_\beta + R_\Delta + R_\lambda$ the series resistance of the secondary and coupling circuit with the aerial disconnected.

$X_{21} + X_\lambda = X_\beta + X_\Delta + X_\lambda$ the series reactance of the same circuit.

and
$$x = \frac{X_\beta}{|(Z_\alpha + Z_\beta)|} = \frac{X_\beta}{|Z_{a1}|}.$$

Then
$$S_R = \frac{1}{\frac{R_{21}}{R_2} + \frac{x^2}{R_2}\left(R_{a1} - \frac{2R_\beta X_{a1}}{X_\beta}\right)} \quad . \quad 3.31b.$$

$$\Delta C_2 = C_{20}\left[\frac{\omega L_2}{X_{21} - x^2 X_{a1}} - 1\right] \quad . \quad . \quad 3.32b.$$

$$M_R = \frac{\omega L_2}{X_{21} - x^2 X_{a1}} - 1$$

$$= \frac{1}{\frac{\omega L_2}{\omega L_2 - X_{21} + x^2 X_{a1}} - 1} \quad . \quad . \quad 3.33b.$$

If $R_{a1} \ll X_{a1}$, a further simplification is possible for $x = \frac{X_\beta}{X_{a1}}$ and

$$S_R = \frac{1}{\frac{R_{21}}{R_2} + \frac{x^2 R_{a1}}{R_2} - \frac{2xR_\beta}{R_2}} \quad . \quad . \quad 3.31c.$$

$$M_R = \frac{1}{\frac{\omega L_2}{\omega L_2 - X_{21} + xX_\beta} - 1} \quad . \quad . \quad 3.33c.$$

If coupling is by mutual inductance $X_{21} = \omega L_2$ and

$$M_R = \frac{1}{\frac{\omega L_2}{xX_\beta} - 1} \quad . \quad . \quad 3.33d$$

and often for other forms of coupling $X_{21} = \omega L_2 + X_\beta$, hence

$$M_R = \frac{1}{-\frac{\omega L_2}{X_\beta(1-x)} - 1} \quad . \quad . \quad 3.33e.$$

Generally $R_\beta \ll X_\beta$ and transfer voltage ratio may be written as

$$T_R = \frac{xZ_\lambda S_R}{R_2}.$$

R_λ is usually very small so that

$$Z_\lambda = jX_\lambda = -\frac{j}{\omega C_2} = -\frac{j}{\omega C_{20}\left(1+\frac{\Delta C_2}{C_{20}}\right)}$$

$$= -\frac{j}{\omega C_{20}(1+M_R)}$$

$$\therefore\ T_R = \frac{xS_R}{j\omega C_{20}R_2(1+M_R)} = \frac{xQ_2S_R}{(1+M_R)} \quad . \quad . \quad . \quad 3.30b.$$

When x is much less than its optimum value $\sqrt{\frac{R_2}{R_{a1}}}$ (see expression 3.21) T_R is very nearly xQ_2; for $x = 0{\cdot}4$ and $0{\cdot}7.\sqrt{\frac{R_2}{R_{a1}}}$ the approximate value of T_R is about 10% and 25% high respectively.

Let us now consider the values of R_{a1}, X_{a1}, x, etc., for the probable forms of aerial coupling circuit The couplings will be designated as shunt, if the coupling impedance is common to aerial and tuned circuit (this is the procedure adopted in network analysis), and series when the impedance joins the aerial to the top or a tapping point on the secondary coil. The formulae for the resistance and reactance elements, together with selectivity and mistune ratio, are given. The formulae for capacitance correction and transfer voltage ratio are not listed since expressions 3.33a and 3.30b show their direct connection with M_R and S_R.

3.4.5. Combined Mutual Inductance and Shunt Capacitance Coupling. The circuit for this type is shown in Fig. 3.11.

$$R_{a1} = R_{a0}+R_1,\ R_{21} = R_2,$$

$$X_{a1} = X_{a0}+\omega L_1 - \frac{1}{\omega C_3},\ X_{21} = \omega L_2 - \frac{1}{\omega C_3}$$

$$x = (a)\ \frac{\omega M - \frac{1}{\omega C_3}}{Z_{a1}} \text{ and } (b)\ \frac{-\omega M - \frac{1}{\omega C_3}}{Z_{a1}}$$

where (*a*) refers to positive M

and (*b*) ,, ,, negative M.

$$S_R = \frac{1}{1+\dfrac{x^2 R_{a1}}{R_2}}$$

$$M_R = \frac{1}{\dfrac{\omega L_2}{x^2 X_{a1}+\dfrac{1}{\omega C_3}} - 1}.$$

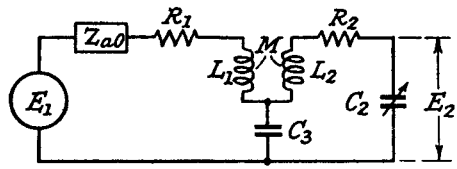


FIG. 3.11.—An Aerial Circuit with Combined Mutual Inductance and Shunt Capacitance Coupling.

Approximate expressions for M_R, obtained by assuming $x = \dfrac{X_\beta}{X_{a1}}$, are

$$M_R = (a)\ \frac{1}{\dfrac{L_2}{Mx+\dfrac{1}{\omega^2 C_3}(1-x)} - 1}$$

and

$$M_R = (b)\ \frac{1}{\dfrac{L_2}{-Mx+\dfrac{1}{\omega^2 C_3}(1-x)} - 1}.$$

3.4.6. Shunt Capacitance Coupling. Shunt Capacitance coupling is illustrated in Fig. 3.12.

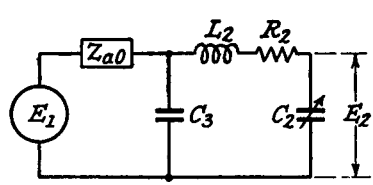


FIG. 3.12.—An Aerial Circuit with Shunt Capacitance Coupling.

$$R_{a1} = R_{a0},\ R_{21} = R_2$$

$$X_{a1} = X_{a0} - \frac{1}{\omega C_3},\ X_{21} = \omega L_2 - \frac{1}{\omega C_3}$$

$$x = \frac{-\dfrac{1}{\omega C_3}}{Z_{a1}}$$

$$S_R = \frac{1}{1+\dfrac{x^2 R_{a0}}{R_2}}$$

$$M_R = \frac{1}{\dfrac{\omega L_2}{x^2 X_{a1} + \dfrac{1}{\omega C_3}} - 1}$$

$$\simeq \frac{1}{\dfrac{\omega^2 L_2 C_3}{1-x} - 1} \qquad \left(x = \frac{X_\beta}{X_{a1}}\right).$$

3.4.7. The Tapped Tuned Circuit. Coupling to a tapping point on the tuned secondary coil is similar in form to mutual inductance coupling, and the equivalent T section can be deduced by analogy from Figs. 3.8*a* and 8*b*. The development is illustrated in Figs. 3.13*a* and *b*. The mutual inductance between the coils is such as to increase the total inductance between 1 and 4 and so by definition in Section 3.4.2 is given a negative sign

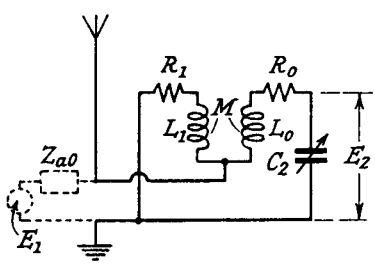


FIG. 3.13*a*.—Tapped Tuned Circuit Aerial Coupling.

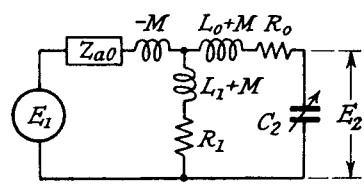


FIG. 3.13*b*.—The Equivalent Circuit of Tapped Tuned Circuit Aerial Coupling.

$$R_{a1} = R_{a0} + R_1, \quad R_{21} = R_1 + R_0,$$

$$X_{a1} = X_{a0} + \omega L_1, \quad X_{21} = \omega(L_1 + L_0 + 2M)$$

$$x = \frac{\omega(L_1+M)}{Z_{a1}}$$

$$S_R = \frac{1}{1+\dfrac{x^2 R_{a1}}{R_1+R_0} - \dfrac{R_1}{(R_1+R_0)}\dfrac{2x^2 X_{a1}}{X_\beta}}$$

$$\simeq \frac{1}{1+\dfrac{x^2R_{a1}}{R_1+R_0}-\dfrac{2xR_1}{R_1+R_0}} \left(x=\frac{X_\beta}{X_{a1}}\right)$$

$$M_R = \frac{1}{\dfrac{\omega(L_1+L_0+2M)}{x^2X_{a1}}-1}$$

$$\simeq \frac{1}{\dfrac{L_1+L_0+2M}{(L_1+M)x}-1}.$$

3.4.8. Series Capacitance Coupling. Fig. 3.14 gives the circuit diagram for series capacitance coupling.

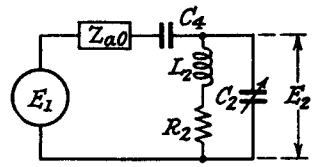


FIG. 3.14.—An Aerial Circuit with Series Capacitance Coupling.

$$R_{a1} = R_{a0}+R_2,\ R_{21} = R_2,$$

$$X_{a1} = X_{a0}+\omega L_2-\frac{1}{\omega C_4},\ X_{21} = \omega L_2$$

$$x = \frac{\omega L_2}{Z_{a1}}$$

$$S_R = \frac{1}{1+\dfrac{x^2R_{a1}}{R_2}-\dfrac{2x^2X_{a1}}{X_\beta}}$$

$$\simeq \frac{1}{1+\dfrac{x^2R_{a1}}{R_2}-2x}$$

$$M_R = \frac{1}{\dfrac{\omega L_2}{x^2X_{a1}}-1}$$

$$\simeq \frac{1}{\dfrac{1}{x}-1}.$$

3.4.9. Combined Series Capacitance and Shunt Inductance Coupling. In order to apply the generalized formulae the π section enclosed in the dotted lines (Fig. 3.15*a*) must be converted into the equivalent T section. The necessary transformation is

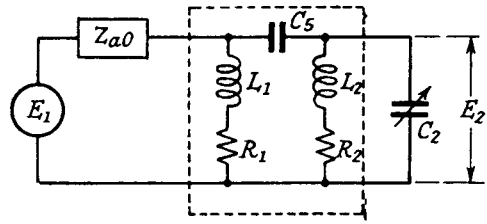


FIG. 3.15*a*.—An Aerial Circuit with Combined Series Capacitance and Shunt Inductance Coupling.

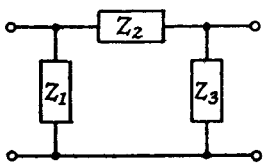


FIG. 3.15*b*.—An Unsymmetrical π Section.

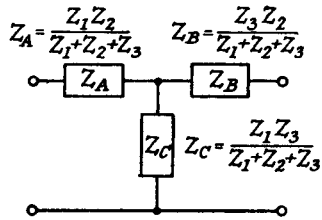


FIG. 3.15*c*.—The T Network Equivalent of Fig. 3.15*b*.

illustrated in Figs. 3.15*b* and 3.15*c*, and the impedances of the series and shunt arms of the T section are

$$Z_A = \frac{(R_1+j\omega L_1)\dfrac{1}{j\omega C_5}}{(R_1+R_2) - j\left(\dfrac{1}{\omega C_5} - \omega(L_1+L_2)\right)}$$

Rationalizing

$$Z_A = \frac{\left[(R_1+R_2)\omega L_1 + R_1\left(\dfrac{1}{\omega C_5} - \omega(L_1+L_2)\right)\right]\dfrac{1}{\omega C_5}}{(R_1+R_2)^2+\left[\dfrac{1}{\omega C_5} - \omega(L_1+L_2)\right]^2}$$

$$+\frac{j\left[\omega L_1\left(\dfrac{1}{\omega C_5} - \omega(L_1+L_2)\right) - R_1(R_1+R_2)\right]\dfrac{1}{\omega C_5}}{(R_1+R_2)^2+\left[\dfrac{1}{\omega C_5} - \omega(L_1+L_2)\right]^2}.$$

Similarly $Z_B = \dfrac{\left[(R_1+R_2)\omega L_2+R_2\left(\dfrac{1}{\omega C_5}-\omega(L_1+L_2)\right)\right]\dfrac{1}{\omega C_5}}{(R_1+R_2)^2+\left[\dfrac{1}{\omega C_5}-\omega(L_1+L_2)\right]^2}$

$$+\frac{j\left[\omega L_2\left(\frac{1}{\omega C_5}-\omega(L_1+L_2)\right)-R_1(R_1+R_2)\right]\frac{1}{\omega C_5}}{(R_1+R_2)^2+\left[\frac{1}{\omega C_5}-\omega(L_1+L_2)\right]^2}$$

and $Z_C = -\dfrac{\left[(\omega^2L_1L_2-R_1R_2)(R_1+R_2)+(R_1\omega L_2+R_2\omega L_1)\left(\dfrac{1}{\omega C_5}-\omega(L_1+L_2)\right)\right]}{(R_1+R_2)^2+\left[\dfrac{1}{\omega C_5}-\omega(L_1+L_2)\right]^2}$

$$-\frac{j\left[(\omega^2L_1L_2-R_1R_2)\left(\frac{1}{\omega C_5}-\omega(L_1+L_2)\right)+(R_1+R_2)(R_1\omega L_2+R_2\omega L_1)\right]}{(R_1+R_2)^2+\left[\frac{1}{\omega C_5}-\omega(L_1+L_2)\right]^2}$$

In practice $\dfrac{1}{\omega C_5}-\omega(L_1+L_2) \gg R_1+R_2$, $\omega L_1 \gg R_1$ and $\omega L_2 \gg R_2$, so that Z_A, Z_B and Z_C can be simplified to

$$Z_A = A\left[R_1+\frac{(R_1+R_2)\omega L_1}{B}\right]+jA\omega L_1$$

$$Z_B = A\left[R_2+\frac{(R_1+R_2)\omega L_2}{B}\right]+jA\omega L_2$$

$$Z_C = -\left[\frac{R_2\omega L_1+R_1\omega L_2}{B}+\frac{\omega^2L_1L_2(R_1+R_2)}{B^2}\right]-j\frac{\omega^2L_1L_2}{B}$$

where $A = \dfrac{1}{B\omega C_5}$, and $B = \dfrac{1}{\omega C_5}-\omega(L_1+L_2)$.

Thus

$$R_{a1} = R_{a0}+A\left[R_1+\frac{(R_1+R_2)\omega L_1}{B}\right]-\frac{R_2\omega L_1}{B}-\frac{R_1\omega L_2}{B}-\frac{\omega^2L_1L_2(R_1+R_2)}{B^2}$$

$$= R_{a0}+AR_1+\frac{A\omega L_1}{B}(R_1+R_2)\left(1-\frac{\omega L_2}{AB}\right)-\left[\frac{R_2\omega L_1+R_1\omega L_2}{B}\right]$$

$$R_{a1} = AR_2+\frac{A\omega L_2}{B}(R_1+R_2)\left(1-\frac{\omega L_1}{AB}\right)-\left[\frac{R_2\omega L_1+R_1\omega L_2}{B}\right]$$

$$X_{a1} = X_{a0} + A\omega L_1 - \frac{\omega^2 L_1 L_2}{B}$$

$$X_{21} = A\omega L_2 - \frac{\omega^2 L_1 L_2}{B}$$

$$x = -\frac{\omega^2 L_1 L_2}{B Z_{a1}}$$

$$S_R = \frac{1}{\frac{R_{21}}{R_2} + \frac{x^2 R_{a1}}{R_2} + \left[\frac{R_2\omega L_1 + R_1\omega L_2}{B} + \frac{\omega^2 L_1 L_2 (R_1+R_2)}{B^2}\right]\left(\frac{2x^2 X_{a1}}{X_\beta}\right)}$$

when $x \simeq \frac{X_{a1}}{X_\beta}$, the last factor in the denominator of S_R becomes $(2x)$

$$M_R = \frac{1}{\frac{\omega L_2}{(1-A)\omega L_2 + \frac{\omega^2 L_1 L_2}{B} + x^2 X_{a1}} - 1}$$

$$\simeq \frac{1}{\frac{1}{(1-A) + \frac{\omega L_1}{B}(1-x)} - 1} \left(x = \frac{X_\beta}{X_{a1}}\right).$$

3.4.10. Combined Mutual Inductance and Series Capacitance Coupling. The circuit is shown in Fig. 3.16*a*; the equivalent circuit is the same as that of Fig. 3.8*b* except that the capacitance C_5 joins the opposite ends of the resistances R_1 and R_2. The

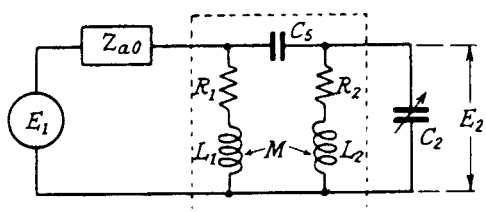


FIG. 3.16*a*.—An Aerial Circuit with Combined Mutual Inductance and Series Capacitance Coupling.

part enclosed between the dotted lines can therefore be transformed to the unsymmetrical bridged T network of Fig. 3.16*b*.

$$Z_1 = R_1 + j\omega(L_1 \mp M),\ Z_2 = R_2 + j\omega(L_2 \mp M);\ Z_3 = j\omega(\pm M),\ Z_4 = \frac{1}{j\omega C_5}$$

and this can in turn be changed to the T section of Fig. 3.16c. After making the same simplifications as in 3.4.9 we have

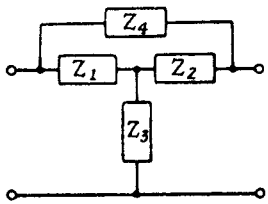


FIG. 3.16b.—An Unsymmetrical Bridged T Section.

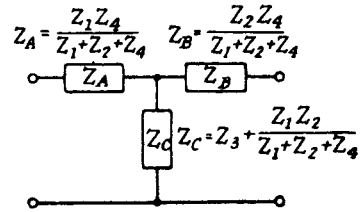


FIG. 3.16c.—The T Network Equivalent of Fig. 3.16b.

$$Z_A = A_1\left[R_1 + \frac{(R_1+R_2)\omega(L_1 \mp M)}{B_1}\right] + j\omega A(L_1 \mp M)$$

$$Z_B = A_1\left[R_2 + \frac{(R_1+R_2)\omega(L_2 \mp M)}{B_1}\right] + j\omega A(L_2 \mp M)$$

$$Z_C = -\left[\frac{[R_2\omega(L_1 \mp M)+R_1\omega(L_2 \mp M)]}{B_1} + \frac{\omega^2(L_1 \mp M)(L_2 \mp M)(R_1+R_2)}{B_1^2}\right] + j\left[\omega M - \frac{\omega^2(L_1 \mp M)(L_2 \mp M)}{B_1}\right]$$

where $A_1 = \dfrac{1}{\omega C_5 B_1}$ and $B_1 = \dfrac{1}{\omega C_5} - \omega(L_1+L_2 \mp 2M)$.

The sign taken by M when combined with L_1 or L_2 is negative for positive M and positive for negative M. It may be noted that when $M = 0$ the expressions for Z_A, Z_B and Z_C are identical with those in Section 3.4.9.

$$R_{a1} = R_{a0} + A_1\left[R_1 + \frac{(R_1+R_2)\omega(L_1 \mp M)}{B_1}\right] - \left[\frac{R_2\omega(L_1 \mp M)+R_1\omega(L_2 \mp M)}{B_1} + \frac{\omega^2(L_1 \mp M)(L_2 \mp M)(R_1+R_2)}{B_1^2}\right]$$

$$R_{21} = A_1\left[R_2 + \frac{(R_1+R_2)\omega(L_2 \mp M)}{B_1}\right] - \left[\frac{R_2\omega(L_1 \mp M)+R_1\omega(L_2 \mp M)}{B_1} + \frac{\omega^2(L_1 \mp M)(L_2 \mp M)(R_1+R_2)}{B_1^2}\right]$$

$$X_{a1} = X_{a0} + A_1\omega(L_1 \mp M) \pm \omega M - \frac{\omega^2(L_1 \mp M)(L_2 \mp M)}{B_1}$$

$$X_{21} = A_1\omega(L_2 \mp M) \pm \omega M - \frac{\omega^2(L_1 \mp M)(L_2 \mp M)}{B_1}$$

$$x = \frac{\pm\omega M - \dfrac{\omega^2(L_1 \mp M)(L_2 \mp M)}{B_1}}{Z_{a1}}$$

$$S_R = \frac{1}{\dfrac{R_{s1}}{R} + \dfrac{x^2 R_{a1}}{R_s} + \left[\dfrac{R_s\omega(L_1 \mp M) + R_1\omega(L_2 \mp M)}{B_1} + \dfrac{\omega^2(L_1 \mp M)(L_2 \mp M)(R_1 + R_2)}{B_1^2}\right]\left(\dfrac{2x^2 X_{a1}}{X_\beta}\right)}.$$

The last factor in the denominator becomes $2x$ when $x \simeq \dfrac{X_\beta}{X_{a1}}$.

$$M_R = \frac{1}{\dfrac{\omega L_2}{(1 - A_1)\omega(L_2 \mp M) + \dfrac{\omega^2(L_1 \mp M)(L_2 \mp M)}{B_1} + x^2 X_{a1}} - 1}$$

$$\simeq \frac{1}{\dfrac{1}{(1 - A_1)\dfrac{(L_2 \mp M)}{L_2} + \dfrac{\omega(L_1 \mp M)(L_2 \mp M)(1 - x)}{B_1 L_2} \pm \dfrac{Mx}{L_2}} - 1}$$

when $x = \dfrac{X_\beta}{X_{a1}}$

3.4.11. Selectivity Ratio Variation over a Tuning Range. From the approximate expression 3.31*c*, we can estimate the trend of selectivity ratio over a given tuning range. Neglecting the variation of the resistance ratios (these are less affected by frequency because numerator and denominator tend to vary together in the same direction), we see that S_R is dependent on x. When x is independent of frequency so is S_R, and hence if X_β and X_{a1} are the same type of reactance, both inductive or both capacitive, S_R is constant over the tuning range (see Curve 4 in Fig. 3.17*a*).

For X_β inductive and X_{a1} capacitive, x is proportional to $-f^2$, hence S_R decreases as the tuning frequency rises (curve 1*a* in Fig. 3.17*a*). The reverse is true when X_β is capacitive and X_{a1} inductive and S_R increases as the tuning frequency increases.

3.4.12. Mistune Ratio and Capacitance Correction Variation over a Tuning Range. An examination of Mistune ratio is possible, from expressions 3.33*d* and 3.33*e*. For mutual inductance coupling (3.33*d*) and X_{a1} capacitive, x is proportional to $-f^2$ and

$$M_R = \frac{1}{\dfrac{K}{-f^2} - 1}.$$

Hence M_R is negative and it increases as f is increased. Since $\Delta C_2 = C_{20}.M_R = \frac{K_1}{f^2}M_R$, ΔC_2 is negative and tends to fall as f is increased. In practice ΔC_2 tends to remain almost constant (curve 1*a* in Fig. 3.17*c*). When X_{a1} is inductive (large primary coil)

$$M_R = \frac{1}{\frac{L_2}{Mx} - 1} \text{ is positive and constant,}$$

because $\frac{L_2}{Mx} > 1$ and x is constant. ΔC_2 is also positive and it decreases as the frequency increases (curve 1*b* in Fig. 3.17*c*).

For couplings other than mutual inductance, expression 3.33*e* indicates that if X_β and X_{a1} are both inductive (x and $\frac{\omega L_2}{X_\beta}$ are positive constants) M_R is negative and independent of frequency. ΔC_2 is negative, decreasing for increasing frequency. For capacitive X_β and X_{a1}, x is a positive constant, but $\frac{\omega L_2}{X_\beta}$ is proportional to $-f^2$. Since X_β is generally much less than ωL_2 and x than 1, $\frac{\omega L_2}{-X_\beta(1-x)} > 1$ and M_R is positive, decreasing as the frequency increases. ΔC_2 is also positive and decreases (curve 4 in Fig. 3.17*c*).

If X_β is inductive and X_{a1} capacitive, x is proportional to $-f^2$ and $\frac{\omega L}{X_\beta}$ is constant. Hence

$$M_R = \frac{1}{-\left(\frac{K_1}{1+K_2f^2}\right) - 1},$$

i.e., is negative and increasing with increasing frequency. ΔC_2 is negative and increases slightly as the frequency rises.

When X_β is capacitive and X_{a1} inductive

$$\left(x = \frac{-K_1}{f^2},\ \frac{\omega L_2}{X_\beta} = -K_2f^2\right) M_R = \frac{1}{\frac{K_2f^2}{1+K_1f^2} - 1}$$

is positive and decreases as the frequency increases. ΔC_2 is positive and almost constant.

3.4.13. Transfer Voltage Ratio Variation over a Tuning Range. Expression 3.30*b* shows that T_R is almost proportional to x and its variation over a tuning range is largely that of x.

The above conclusions are best summarised in a table.

TABLE 3.1. TUNING FREQUENCY INCREASING

X_β	X_{a1}	x	S_R	M_R	ΔC_2	T_R	Δf
L	L	constant	constant	constant —	decreases —	constant	increases —
C	C	constant	constant	decreases +	decreases +	constant	decreases +
L	C	$-Kf^2$	decreases	increases —	decreases —	increases	increases —
C	L	$-\frac{K}{f^2}$	increases	decreases +	almost constant +	decreases	almost constant +
M	L	constant	constant	constant +	decreases +	constant	almost constant +
M	C	$-Kf^2$	decreases	increases —	slight decrease —	increases	increases —

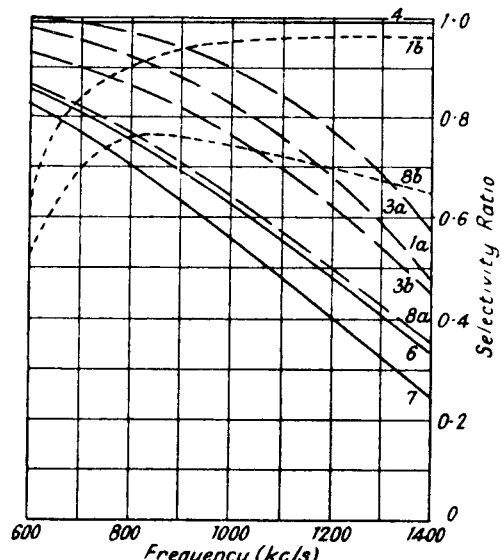


FIG. 3.17*a*.—Typical Selectivity Ratio Variations over the Medium Wave Range.

Typical examples of the variation of S_R, M_R, ΔC_2, T_R and Δf are shown in Figs. 3.17*a*, *b*, *c*, *d* and *e* for the following types of coupling.

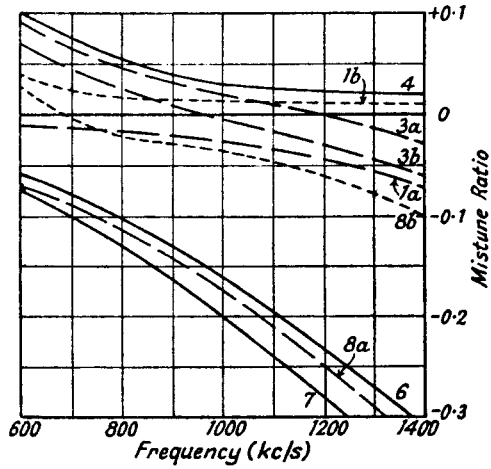


FIG. 3.17b.—Typical Mistune Ratio Variations over the Medium Wave Range.

1a. Mutual inductance, small primary coil, X_{a1} capacitive.
1b. ,, ,, , large ,, ,, , X_{a1} inductive.

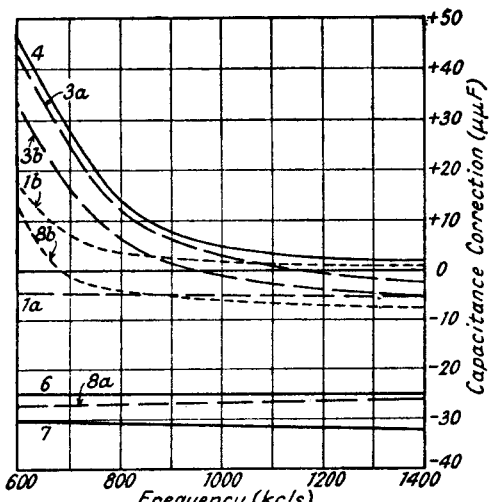


FIG. 3.17c.—Typical Capacitance Correction Variations over the Medium Wave Range.

3a. Positive mutual inductance, small primary coil, and shunt capacitance.
3b. Negative mutual inductance, small primary coil, and shunt capacitance.
4. Shunt capacitance.
6. Series capacitance.
7. Series capacitance and shunt inductance.
8a. Positive mutual inductance, small primary coil, and series capacitance.
8b. Negative mutual inductance, large primary coil, and series capacitance.

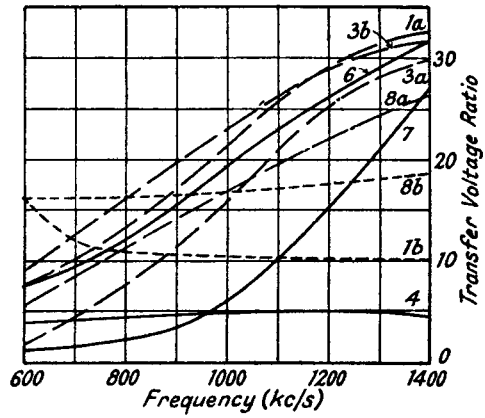


FIG. 3.17*d*.—Typical Transfer Voltage Ratio Variations over the Medium Wave Range.

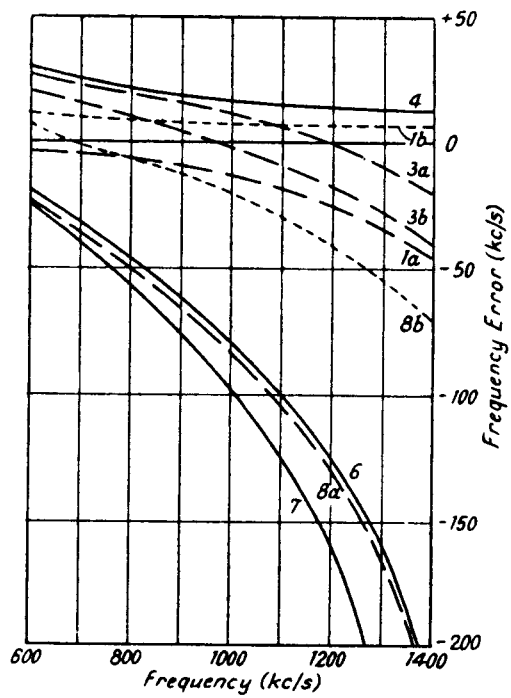


FIG. 3.17*e*.—Typical Frequency Error Variations over the Medium Wave Range.

The curves follow the predictions of Table 3.1, except for slight modifications due to an approach to aerial and coupling circuit resonance. Thus mutual inductance with a large primary coil (X_β and X_{a1} inductive) shows a decrease of S_R and an increase in M_R and T_R as the tuning frequency falls, because X_{a1} is approaching zero at the low frequency end of the range.

Combined couplings mainly affect x and those reducing the variation of x over the frequency range are advantageous in reducing the variation of T_R and S_R. For example, negative mutual inductance and shunt capacitance (curves 3*b*) giving $X_\beta = -\left(\omega M + \frac{1}{\omega C_5}\right)$ and positive mutual inductance with series capacitance (curves 8*a*) giving $X_\beta = +\omega M - \frac{\omega^2(L_1 - M)(L_2 - M)}{B}$ show reduced variation of T_R and S_R as compared with mutual inductance alone. There is a disadvantage; the average value of S_R is reduced and damping of the aerial circuit is therefore increased. Series capacitance coupling may be added to negative mutual inductance with a large primary coil so as to raise T_R at the high frequency end (curve 8*b*), but the average S_R is reduced.

The variation of ΔC_2 over the frequency range follows Table 3.1. For combined couplings it may change sign at some point in the frequency range (curves 3*a*, 3*b* and 8*b*), and this is due to the fact that the reflected aerial reactance $x^2.X_{a1}$ is opposite in sign to that of the coupling reactance. The most desirable curve of ΔC_2 against frequency is that giving a horizontal line, i.e., ΔC_2 is independent of frequency. Provided the constant value of ΔC_2 is within the range of the trimmer across the tuning capacitance, mistuning effects can be cancelled. From this point of view series capacitance (6), mutual inductance with small primary coil (1*a*), positive mutual inductance (small primary) and series capacitance (8*a*), and series capacitance with shunt inductance (7) are satisfactory.

We may note that if $R_{a0} \ll X_{a0}$ and X_{a0} is capacitive, ΔC_2 for three of the couplings may be written

Shunt capacitance (4) $\Delta C_2 = \dfrac{-C_{20}^2}{C_3 + C_{a0} - C_{20}}$

Series capacitance (6) $\Delta C_2 = \dfrac{-C_4 C_{a0}}{C_4 + C_{a0}}$ (independent of f)

Combined series capacitance and inductance (7) $\Delta C_2 = \dfrac{C_5}{1 - \dfrac{\omega^2 L_1 C_5}{1 - \omega^2 L_1 C_{a0}}}$.

Shunt capacitance coupling gives a large variation of ΔC_2 and it is not therefore desirable.

If no correction is made for coupling and reflected aerial reactance, mutual inductance with a large primary coil [9] gives least variation of frequency error (curve 1*b* in Fig. 3.17*e*). Shunt capacitance 4 is much better than series capacitance coupling 6. When frequency error varies over the range it is preferable to select a coupling giving least error at the low frequency end where the selective properties of the secondary circuit are greatest.

It is important next to consider the effect of aerial terminal impedance on S_R, M_R, ΔC_2 and T_R.

3.4.14. Aerial Terminal Impedance and S_R, ΔC_2 and T_R. —(*a*) *Increasing* R_{a0}. With a high value of X_{a1}, increase of R_{a0} increases R_{a1}, but hardly affects Z_{a1} and x. Hence M_R and ΔC_2 are almost unchanged; S_R is decreased because the second term in 3.31*c* is increased. If the aerial and coupling circuit approach resonance in the tuning range, i.e., $X_{a1} \rightarrow 0$, the reflected aerial reactance tends to disappear so that M_R and ΔC_2 fall. The second term in 3.31*c* approaches $\frac{X_\beta^2}{R_{a1}R_2}$ and so S_R tends to increase when R_{a0}, i.e., R_{a1}, increases. The general shape of S_R curve over a tuning range in which $X_{a1} = 0$ is shown in Fig. 3.18; it is a condition rarely occurring in practice, and increase of R_{a0} normally decreases S_R.

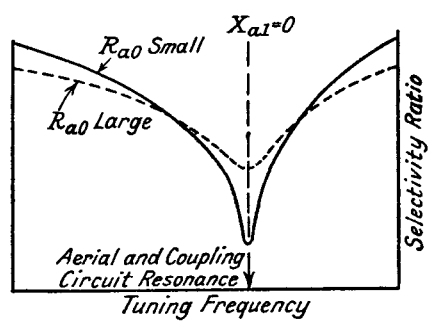


FIG. 3.18.—Selectivity Ratio Variation over a Tuning Range passing through Aerial and Coupling Circuit Resonance for a Large and Small Value of R_{a0}.

(*b*) *Increasing* X_{a0}. Increase of aerial terminal reactance X_{a0}, provided it has the same sign as X_{a1}, increases Z_{a1} and decreases x. Hence T_R is decreased and S_R increased. The statement with regard to T_R needs qualification if the coupling reactance X_β is initially greater than its optimum value, for increase of Z_{a1} then causes optimum coupling to be approached and T_R therefore increases. The normal decrease follows when optimum coupling is passed.

From 3.33*b* we see that M_R is decreased if X_{a1} and X_{a0} are negative, but is increased if X_{a1} is positive and X_{a0} is negative.

The following tables summarize the conclusions.

TABLE 3.2*a*. R_{a0} INCREASING

X_{a1}	x	S_R	M_R	ΔC_2	T_R
Very large	little affected	decreased	little affected	little affected	decreased
Small and comparable with R_{a0}	decreased	increased	tends to decrease	tends to decrease	decreased

TABLE 3.2*b*. X_{a0} INCREASING

X_{a1}	x	S_R	M_R	ΔC_2	T_R
Same sign as X_{a0}	decreased	increased	decreased	decreased	decreased if initial coupling < optimum increased if > optimum
Opposite sign to X_{a0}	increased	decreased	increased	increased	increased if initial coupling < optimum decreased if > optimum

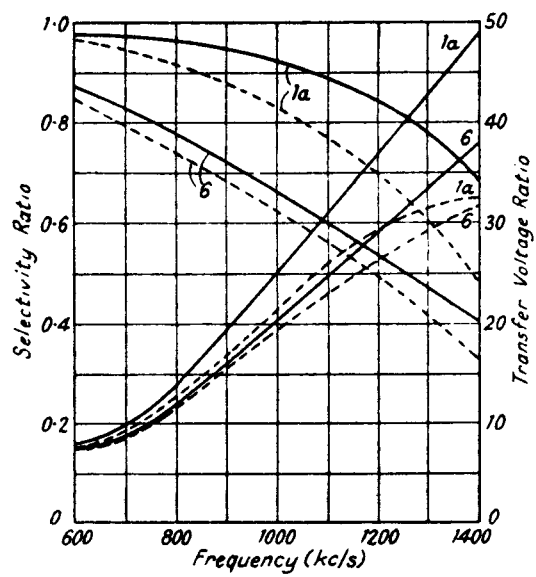


FIG. 3.19*a*.—The Effect on Selectivity and Transfer Voltage Ratios of a Change of R_{a0}.

Full line, R_{a0} = 15 ohms. C_{a0} = 177 $\mu\mu$ F.
Dotted line, R_{a0} = 42 ohms.

The result that a decrease of X_{a0} (X_{a1} of the same sign) tends to increase T_R when coupling is less than optimum shows that the addition of a horizontal top to a vertical aerial can not only increase the effective generated voltage (Section 3.3.3) but also give a greater transfer voltage ratio. The horizontal section reduces X_{a0} and X_{a1}, thus increasing x and T_R.

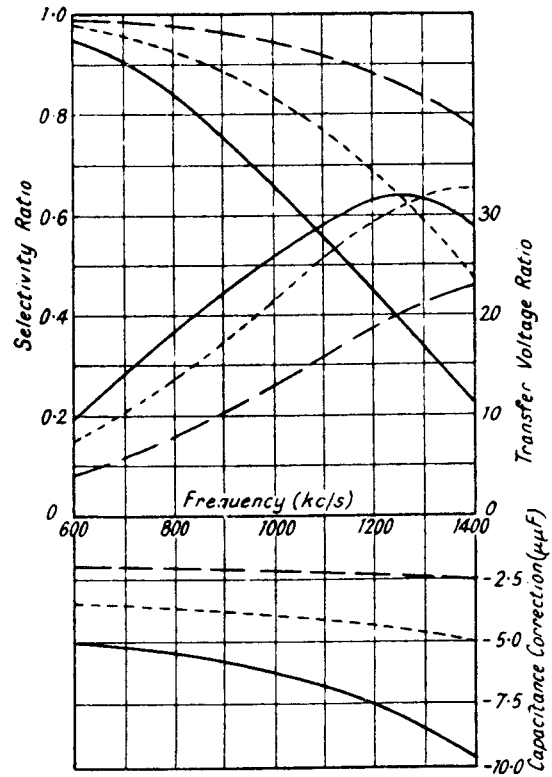


FIG. 3.19*b*.—The Effect on Selectivity and Transfer Voltage Ratios and Capacitance Correction of a Change of X_{a0} with Small Primary Coil.

Broken line, C_{a0} = 104 μμF. R = 42 ohms.
Dotted line, C_{a0} = 177 μμF. R = „ „
Full line, C_{a0} = 240 μμF. R = „ „

Typical examples of varying R_{a0} and X_{a0} are shown in Figs. 3.19*a* and 3.19*b* over the medium wave range for mutual inductance with a small primary coil (1*a*) and series capacitance coupling (6) (X_{a0} and X_{a1} capacitive). It is seen that increase of R_{a0} has greatest effect at the high frequency end of the range due to an approach to aerial circuit resonance (X_{a1} decreasing rapidly), and the variation

of T_R over the tuning range is reduced, though at the expense of a reduction of average S_R. This applies to all forms of inductive coupling with capacitive X_{a1} (e.g., 1*a*, 2, 3*a*, 3*b*, 6, 7 and 8). For inductive X_β and X_{a1} (large primary coil) variation of T_R is also reduced, but increase of R_{a0} has greatest effect at the low frequency end.

Increase of aerial terminal reactance X_{a0}, when X_{a0} and X_{a1} are both capacitive, increases Z_{a1} and decreases x. T_R is decreased and S_R increased. Most reduction of T_R and increase of S_R occurs at the end of the range approaching aerial circuit resonance, in this case the high frequency end. A modification occurs with $C_{a0} = 240$ $\mu\mu$F, T_R is reduced at the high frequency end as compared with $C_a = 177$ $\mu\mu$F, and this is due to optimum coupling having been exceeded. Capacitance correction (ΔC_2) and its variation over the tuning range is decreased because increase of X_{a0} increases X_{a1}.

3.5. Interference Reducing Aerial Systems.

3.5.1. Introduction. In the early days of broadcast reception the receiver was comparatively insensitive, and the aerial had therefore to be as good a collector as possible. As receiver design progressed and sensitivity was increased (mainly by the adoption of the superheterodyne principle), the necessity for a good aerial became less pressing and the tendency was to neglect this aspect of reception, employing inefficient indoor aerials or the mains wiring to supply the pick-up voltage. Attention is once again being focused on the aerial because of the increased use of electrical apparatus producing fortuitously or intentionally R.F. voltage components. For example, any form of sparking (in commutator machines, switching on apparatus, etc.), distortion of wave form by gas discharge illuminated signs, short wave therapy apparatus, etc., generate appreciable R.F. components, which can be transmitted for considerable distances over the supply lines, unless special precautions are taken to suppress or confine them to the apparatus. Since the supply lines are the transmission media the interference field is mainly localized in the building housing the receiver. The disadvantage of the indoor aerial is at once apparent; not only is it an inefficient collector (due to the screening effect of the building), but it is also located in the area of maximum interference field. It is possible to reduce the effect on the receiver by including R.F. filters in the incoming mains leads to the house, and by erecting an aerial well clear of the latter and coupling it to the

receiver by twisted, screened twin or concentric feeder, in which the interference voltage is either cancelled or eliminated. An example of a common type of interference reducing aerial [7] is shown in Fig. 3.20. The vertical aerial (often in the form of a spike attached to the highest point on the building outside the region of intense interference) is self-supporting and rarely longer than about 6 feet (1·83 metres), and is coupled at its base either direct to the screened twin feeder or by a matching transformer, T_1. At the other end of the feeder is a transformer T_2 (with earthed centre tapped primary), matching the low impedance feeder to the high impedance input of the receiver. It is almost impossible to obtain correct matching between the aerial and feeder over a large fre-

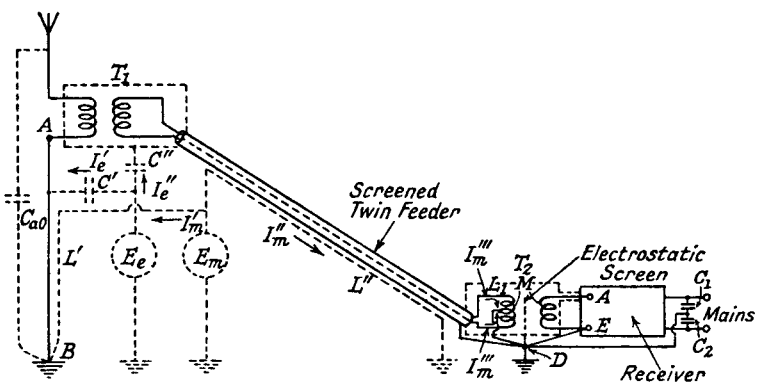


FIG. 3.20.—A Local-Interference Reducing Aerial System.

quency range, and the use of a transformer is only justified under certain conditions. The feeder-to-receiver transformer is generally designed on the basis of a receiver input impedance of 2,000 to 3,000 ohms. The dotted circuits in Fig. 3.20 indicate the possible ways by which interference is induced in the aerial circuit. The electrostatic component of the interference field is represented by the generator E_e and capacitances C' and C'', and the magnetic component by the generator E_m and loops L' and L''. The earthed lead AB from the aerial feeder matching transformer T_1, which is enclosed in a screening box earthed to the feeder shield but not to A, carries interference currents from E_e and E_m and an interference voltage is induced in the lead due to its inductive reactance. This voltage has a return path to earth via the aerial capacitance C_{a0}, but the current it produces in the primary of T_1 is usually negligible. If transformer T_1 is located in the interference field and connected by an unscreened lead to the aerial, considerable

interference is caused because the currents from E_e and E_m then pass through the primary of T_1. The earthed screen round T_1, the feeders and T_2 prevents electrostatic pick-up, but current in the loop L'' induces currents in the shield and feeders. The currents I_m''' in the feeders are equal and opposite, provided the feeders are balanced to earth; to ensure this the centre point of the primary of T_2 is earthed. An electrostatic screen between the primary and secondary also helps to preserve this balance and prevents interference voltage transfer (from the currents I_m''' in the half primaries) to the secondary by interwinding capacitance. The screening for feeders and transformers is earthed at one point only, preferably D, which is likely to be nearer to earth than A. If it is also earthed at A a loop ABD is formed in the interference field and pick-up due to the magnetic component is greatly increased. The earthed point of any filter system reducing interference entering by the mains (see capacitance C_1 and C_2 in Fig. 3.20) should be taken separately to earth rather than via the receiver earth. The improvement to be expected in signal-to-interference ratio by the use of this type of aerial circuit is from 30 to 40 dbs., but it is obtained at the cost of a reduction in desired signal (compared with the unshielded aerial and lead in) of from 3 to 6 dbs.

3.5.2. The Characteristic Impedance of Feeders. The characteristic impedance of any uniform feeder may be determined by measuring the capacitance of a comparatively short length of feeder at any suitable radio frequency, e.g., about 1,000 kc/s, and noting (expression 3.3c) that

$$Z_0 = \frac{l}{vC_1} = \frac{l \times 10^{12}}{3 \times 10^{10} C_1} = \frac{100l}{3C_1}\,\Omega$$

where l = length of feeder (cms.)
and C_1 = total capacitance of feeder in $\mu\mu$F.

The formula for calculating Z_0 of a concentric feeder from its dimensions is

$$Z_0 = \frac{138}{k} \log_{10} \frac{r_2}{r_1} \text{ ohms} \quad . \quad . \quad . \quad 3.34$$

where k = specific inductive capacitance of the dielectric separating the lines.
r_2 = inner radius of outer conductor.
and r_1 = outer radius of inner conductor.
For a pair of parallel unscreened wires

$$Z_0 = \frac{276}{k} \log_{10} \left(\frac{d}{r} + \sqrt{\frac{d^2}{r^2} + 1} \right) \quad . \quad . \quad . \quad 3.35$$

where d = distance between conductor centres.
and r = radius of conductor.

Curves are given in Fig. 3.21 of characteristic impedance against $\frac{r_2}{r_1}$ and $\frac{d}{r}$ for the concentric and parallel wire feeders with $k = 1$.

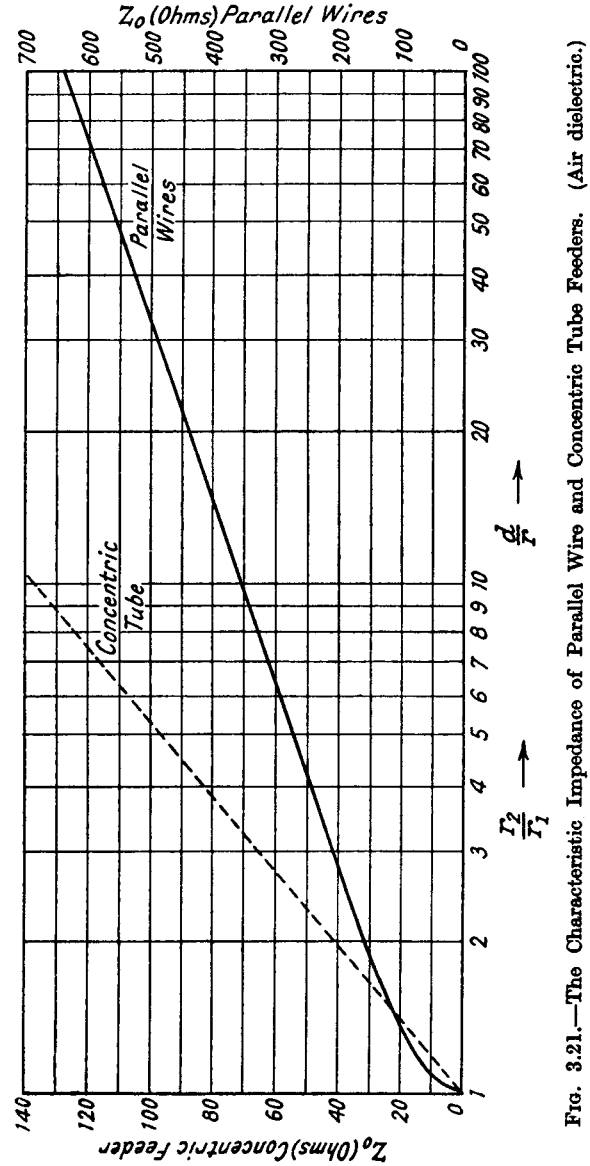


FIG. 3.21.—The Characteristic Impedance of Parallel Wire and Concentric Tube Feeders. (Air dielectric.)

3.5.3. The Aerial to Feeder Connection. We will consider first the case of a feeder correctly terminated at the receiver end in its characteristic impedance Z_0. Owing to the fact that the aerial input impedance varies over the frequency range and is reactive, except at its resonant frequencies, there must be a loss of power transfer at the junction of feeder and aerial. The maximum power transfer occurs when the load resistance (the characteristic impedance of the feeder) is equal to the generator (aerial) terminal resistance and the reactance of the latter is zero. The power transferred to the feeder under these conditions is

$$P_0 = \frac{E_1^2}{4R_{a0}} \quad . \quad . \quad . \quad . \quad 3.36$$

where E_1 = the generator (aerial) voltage.

Expression 3.36 is written as $\frac{E_1^2}{4R_{a0}}$ rather than $\frac{E_1^2}{4Z_0}$ since it is the generator resistive component which is variable and which determines the maximum power transfer.

When the generator (aerial) impedance is equal to $R_{a0} \pm jX_{a0}$, the power delivered to the feeder is

$$P_1 = I_1^2 Z_0 = \left[\frac{E_1}{\sqrt{(R_{a0}+Z_0)^2+X_{a0}^2}}\right]^2 Z_0 \quad . \quad . \quad 3.37.$$

The ratio of maximum to actual power is

$$\frac{P_0}{P_1} = \frac{(R_{a0}+Z_0)^2+X_{a0}^2}{4R_{a0}Z_0} \quad . \quad . \quad . \quad 3.38a$$

$$= \frac{\left(\frac{R_{a0}}{Z_0}+1\right)^2+\left(\frac{X_{a0}}{Z_0}\right)^2}{\frac{4R_{a0}}{Z_0}} \quad . \quad . \quad . \quad 3.38b.$$

Replacing $\frac{P_0}{P_1}$, $\frac{R_{a0}}{Z_0}$ and $\frac{X_{a0}}{Z_0}$ by a, x and y

we have $(x+1)^2+y^2 = 4ax$

or $x^2+2(1-2a)x+1+y^2 = 0.$

This is the equation to a circle with centre on the x axis, and solving for $y = 0$ gives the two values of x corresponding to the intersections of the circle with the x axis. Thus

$$x_2 = \frac{2(2a-1)+\sqrt{4(2a-1)^2-4}}{2}$$

$$= (2a-1)+\sqrt{(2a-1)^2-1}$$

and $$x_1 = (2a-1)-\sqrt{(2a-1)^2-1}.$$

The centre of the circle is at $x_3 = \frac{x_2+x_1}{2}$

or $$x_3 = 2a - 1$$

and the radius is $$r = \frac{x_2 - x_1}{2}$$

$$= \sqrt{(2a-1)^2 - 1}.$$

Taking a particular example, for a loss of 3 dbs., $a = 2$, $x_3 = 3$, $r = 2{\cdot}82$.

A series of curves, due to Wheeler,[4] are shown in Fig. 3.22. The semicircles connect values of $\frac{X_{a0}}{Z_0}$ and $\frac{R_{a0}}{Z_0}$ giving the particular

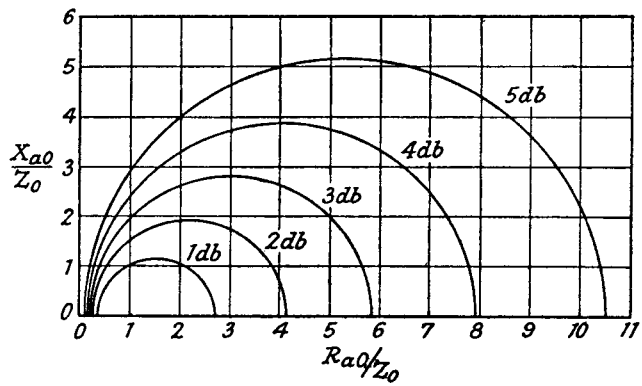


FIG. 3.22.—A Chart for Estimating Transition Loss due to Mismatching between an Aerial and Correctly Terminated Feeder.

transition loss marked. Thus the *V* dipole, whose impedance characteristic is given by curves 2 in Fig. 3.5, has at 13·2 Mc/s $\left(\frac{f}{f_0} = 1{\cdot}3\right)$ an impedance of $R_{a0}+jX_{a0} = 129+j255$, and coupling this to a correctly terminated 600Ω feeder gives $\frac{R_{a0}}{Z_0} = 0{\cdot}215$ and $\frac{X_{a0}}{Z_0} = 0{\cdot}425$, which from Fig. 3.22 corresponds to a loss of approximately 3 dbs. Calculation from $10 \log_{10} \frac{P_0}{P_1}$ (expression 3.38*b*) gives 2·8 dbs.

The loss (over the range 6 to 15 Mc/s) at the junction of the plain and *V* dipole aerials of Fig. 3.5 with a 600Ω feeder is shown in Fig. 3.23.

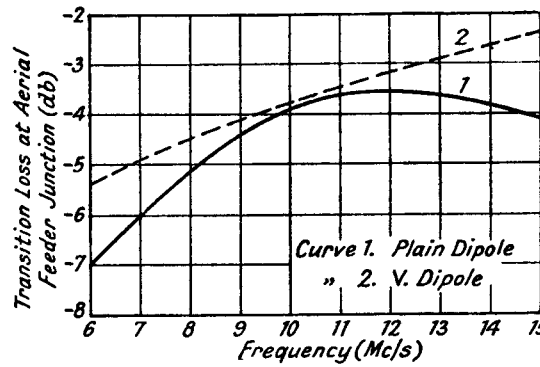


FIG. 3.23.—Curves of Transition Loss at the Junction of a Plain and *V* Dipole Aerial and a 600-ohm Feeder.

When the feeder is not correctly terminated the problem becomes more complicated. Let us suppose that a vertical aerial is connected to a receiver of input impedance Z_R via a feeder of characteristic impedance Z_0, having a propagation constant $Y = \alpha + j\beta$ (see Fig. 3.24). The aerial terminal impedance Z_{a0} (calculated as indicated in Section 3.3.2) is the terminating impedance of the

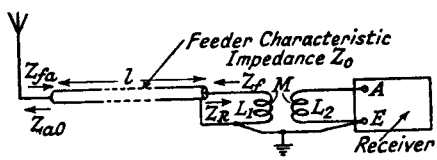


FIG. 3.24.—A Single Concentric Tube Feeder with no Aerial Matching Transformer.

feeder so that the impedance [6] at the receiver end of the feeder (Fig. 3.24) is given by

$$Z_f = Z_0 \frac{Z_{a0} \cosh \gamma l + Z_0 \sinh \gamma l}{Z_{a0} \sinh \gamma l + Z_0 \cosh \gamma l} \quad . \quad . \quad 3.39$$

$$= R_f + jX_f$$

where Z_0 = characteristic impedance of the feeder
and l = length of feeder.

The aerial and feeder may be replaced by a generator having an internal impedance Z_f and an open circuit voltage of $\frac{KE_1Z_R}{Z_{a0}+Z_R}$

where E_1 = aerial generated voltage

$$\text{and } K = \frac{2Z_0(Z_{a0}+Z_R)\varepsilon^{-\gamma l}}{(Z_{a0}+Z_0)(Z_R+Z_0)\left[1 - \frac{(Z_{a0}-Z_0)(Z_R-Z_0)}{(Z_{a0}+Z_0)(Z_R+Z_0)}\varepsilon^{-2\gamma l}\right]} \quad 3.40.$$

This formula is developed in Chapter 4, p. 116, of Bibliography 22, where curves of transition loss due to mismatching will also be found.

A simplification of the problem is possible in calculating over the long and medium wave ranges, because the feeder length is usually short compared with the lowest wavelength, and the attenuation constant α is negligible. In section 3.3.2 it is shown that the impedance of a vertical aerial of height less than $\frac{\lambda}{8}$ is practically equivalent to that of its electrostatic capacitance, and the same rule applies to a feeder line. Thus, if the lowest medium wavelength is 200 metres, feeder lines not exceeding 25 metres (82 feet) in length may be replaced by a shunt capacitance equal to the electrostatic capacitance. If the feeder attenuation constant cannot be neglected the feeder may be replaced by two shunt capacitances of half the feeder electrostatic capacitance and a series resistance. If a matching transformer [14] is employed between aerial and transformer, the equivalent circuit becomes that of Fig. 3.8*b* where C_2 is now the feeder capacitance C_f. Resonance effects between C_{a0}, L_1, L_2 and C_f can be used to produce a more level overall frequency response over a given range.

3.6. Aerials for Automobile Receivers.[11, 16] The design of aerials for automobile receivers is chiefly concerned with finding an aerial which is a reasonably efficient collector. Mechanical rigidity and location of the aerial, which should be as far from the ignition system as possible, are important. There is little difficulty in designing the coupling to the receiver first tuned circuit as it follows the lines set out in Section 3.4. Before the use of all-steel car construction the most efficient type of aerial consisted of gauze strip let into the roof of the car. The under-car aerial, mounted below the running-board, has been used, but it is a much less efficient collector, is liable to damage, and its pick-up properties are seriously impaired in wet weather due to mud adhering to the insulators and offering a low resistance path to earth. Owing to its position its capacitance to earth is comparatively high (about 200 $\mu\mu$F). The roof aerial mounted above and parallel to the roof of the car is a better collector and has a lower earth capacitance (about 50 $\mu\mu$F). A third type of aerial is the " telescopic whip "

vertical aerial located usually at the side of the car behind the bonnet. It may be capable of extension to a length of about 5 feet, and the sliding sections should have clamps making intimate contact between sections so as to prevent interference from intermittent contacts when travelling. Its capacitance to earth may be as low as 25 $\mu\mu$F. A concentric feeder with the outer conductor earthed to the car chassis is used to couple the aerial to receiver ; it should be as short as possible and should have the least possible capacitance value. The equivalent circuit is that of Fig. 3.24, and the feeder capacitance C_f acts with the aerial capacitance as a potential divider for the pick-up voltage ; a large value of C_f reduces the receiver input voltage and also necessitates looser coupling to the receiver for a given ganging error in the first tuned circuit.

3.7. The Connection of Several Receivers to one Aerial System. Difficulties are met if attempts are made to connect several receivers to the same aerial. Unless the coupling between each receiver and the aerial lead-in is loose, interaction may occur between the receivers, causing mistuning of the first tuned circuit of one receiver when the tuning of another is changed. There may also be sufficient oscillator voltage injected into the aerial circuit from the frequency changer to cause whistles in the other receivers as one is tuned over its range. It is possible, where loss of signal strength can be tolerated, to couple several receivers through suitable transformers to the feeder line of the interference reducing aerial described in Section 3.5.1. A more satisfactory arrangement, particularly for a large installation, such as that in a block of flats, is to distribute the signal frequencies via a low impedance feeder line, preceded by an aperiodic amplifier or with an aperiodic buffer amplifier between each receiver and the feeder. The second method has been employed but is obviously not so satisfactory economically as the first. A suitable circuit [15] for the first method is shown in Fig. 3.25. The aerial, a short vertical rod, erected at the highest point in the building, is connected by a short screened lead to two aperiodic amplifiers, covering a total frequency range from 150 kc/s to 15 Mc/s. It is not practicable to cover the whole range with one amplifier without appreciable loss of amplification. Between the aerial lead and each amplifier is a filter selecting the particular frequency range required ; to reduce further the possibilities of cross modulation the valves should have as near linear $I_a E_g$ characteristics as possible. Wave traps may also be necessary if the aerial is in the high field strength area of a local station. The outputs from the two amplifiers are connected in parallel, through step-

down transformers, to a low impedance feeder line of low loss terminated in its characteristic impedance. The advantage of the terminated feeder is that it has a wide frequency response and conveys power with low losses and small noise pick-up. Furthermore, its low impedance makes the reduction of interaction between receivers much simpler, and the comparatively high impedance inputs of the receivers can be connected to the feeder without impairing its transmission characteristics. The line is matched to the amplifier valves by transformers but no attempt is made to match the receiver input to the feeder. A resistance R is connected in series with

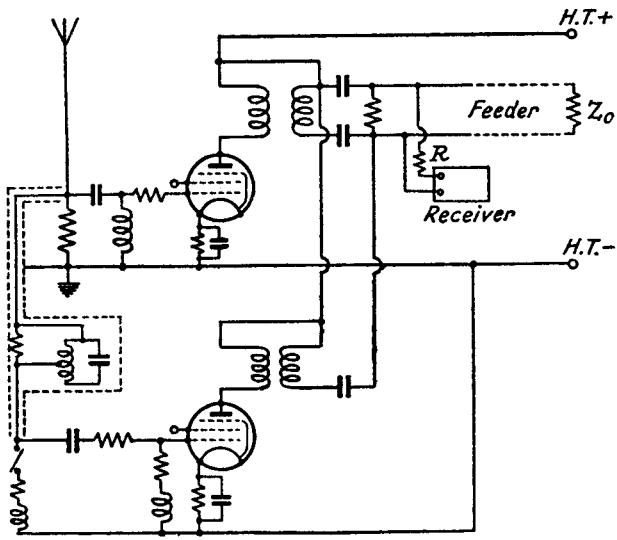


FIG. 3.25.—An Amplifier for a Feeder supplying Several Receivers.

each receiver input to reduce interaction still further. This resistance need only affect the input voltage at the receiver to a small extent, since it can be much lower than the receiver input impedance and yet much higher than the feeder impedance. For example, if the receiver input impedance is 3,000 Ω and the series resistance 1,000Ω, the greatest possible reduction of input voltage is to $\frac{3}{4}$ of the voltage at the feeder, whilst interaction is reduced to $\frac{70}{1070} \simeq 7\%$, if the characteristic impedance of the feeder is 70 ohms. The length of the feeder is determined by the highest received frequency because the feeder losses are greatest at this frequency; a length of 400 metres (about $\frac{1}{4}$ mile) is possible for the frequency range

quoted above, 150 kc/s to 15 Mc/s, but it may be increased to 2,000 metres (about $1\frac{1}{4}$ miles) if reception is restricted to the long and medium wave ranges.

3.8. Diversity Reception. Owing to its dependence on the reflected ray, short wave reception is liable to considerable variations in signal strength. The actual value of the signal at any given time instant depends on the position of the aerial, and there may be large differences in field strength between positions spaced only a few wavelengths apart. By suitably combining the outputs from two or more aerials with at least a wavelength separation, the average signal may be maintained at a more constant level. Diversity reception, as it is called, is quite often employed for commercial reception, and a large number of aerials are located at different positions and connected to the receiver by feeder lines. This is clearly not feasible for broadcast receiver purposes, but some improvement can be obtained by using two aerials, one vertical and the other horizontal, separated by at least a wavelength. One method [12] of combining the two aerial outputs is by means of a gas-filled valve with an interrupter relay in its anode circuit. The arm of the relay drives a sprocket wheel attached to a differential capacitor, the rotor of which is connected to the aerial terminal of the receiver. The stators are connected each to an aerial, and the gas-filled relay is biased by the D.C. component of the detected I.F. output voltage of the receiver. When the output is greater than a certain predetermined value, sufficient negative bias voltage is developed to render the gas-filled valve inoperative. If it falls below this value the bias is reduced, the valve conducts, and the interrupter relay drives the sprocket wheel until the output voltage is large enough to shut down the valve again. The sprocket wheel comes to rest with the capacitor giving more coupling to the aerial having greatest signal. If both aerials give outputs less than the value required to stop the relay the capacitor continues to rotate.

When the same programme is transmitted on two different carrier frequencies, improved reception may be gained by receiving each transmission on a separate receiver and combining the audio frequency outputs.[10]. A common A.G.C. system actuated from the strongest signal is employed for the two receivers, and A.F. combination may be achieved by dual loudspeakers, two output transformers, the secondaries of which are connected in series with the speech coil of a loudspeaker, or by a single centre-tapped-primary output transformer with the output valve of each receiver con-

nected to a half primary. A common aerial system may be used, but it is preferable to have (for example) a horizontal aerial for one receiver and a vertical for the other.

BIBLIOGRAPHY

1. On the Capacity of Radio Telegraphic Antennae. G. W. O. Howe, *Wireless Engineer*, Dec. 1914, p. 546; Jan. 1915, p. 612; Feb. 1915, p. 680.
2. The Capacity of Aerials of the Umbrella Type. G. W. O. Howe, *Wireless Engineer*, Oct. 1915, p. 426.
3. The Balance of Power in Aerial Tuning Circuits. F. M. Colebrook, *Wireless Engineer*, March, 1930, p. 129.
4. Transition Loss Chart. H. A. Wheeler, *Electronics*, Jan. 1936, p. 27.
5. The Design of Doublet Antenna Systems. H. A. Wheeler and V. E. Whitman, *Proc. I.R.E.*, Oct. 1936, p. 1257.
6. The Calculation of Input and Sending End Impedance of Feeders and Cables terminated with a Complex Load. H. Cafferata, *Marconi Review*, Jan.–Feb. and Sept.–Dec. 1937, pp. 12 and 21.
7. Screened Aerials. F. R. W. Strafford, *Wireless World*, Nov. 25th, 1937, p. 516.
8. The Impedance Characteristic of Short Wave Dipoles. T. Walmsley, *Phil. Mag.*, June 1938, p. 381.
9. The Aerial Connection. M. G. Scroggie, *Wireless World*, June 23rd, 1938, p. 548; June 30th, 1938, p. 579.
10. Diversity Reception at Home. R. H. Tanner, *Wireless World*, Sept. 1st, 1938, p. 194.
11. Car Aerials. F. R. W. Strafford, *Wireless World*. Oct. 20th, 1938, p. 344.
12. Dual Diversity Reception. *Wireless World*, March 2nd, 1939, p. 208.
13. Vertical or Inverted L Aerials. F. R. W. Strafford, *Wireless World*, June 15th, 1939, p. 554; June 22nd, 1939, p. 575.
14. Receiving Aerials. J. van Slooten, *Philips Technical Review*, Nov. 1939, p. 320.
15. A Central Antenna System. D. J. Fruin, *Electronics*, Nov. 1939, p. 37.
16. Design Problems in Automobile Radio Receivers. G. C. Hall, *A.W.A. Technical Review*, Vol. 4, No. 3, 1939, p. 105.
17. Receiver Aerial Coupling Circuits. K. R. Sturley, *Wireless Engineer*, April 1941, p. 137; May 1941, p. 190.
18. The Theory of Antennas of Arbitrary Size and Shape. S. A. Schelkunoff, *Proc. I.R.E.*, Sept. 1941, p. 493.
19. Antennas for Frequency Modulated Reception. J. G. Aceves, *Electronics*, Sept. 1941, p. 42.
20. The Design of Television Receiving Apparatus. B. J. Edwards, *Journal I.E.E.*, Wireless Section, Sept. 1941, p. 191.
21. Aerial Characteristics. N. Wells, *Journal I.E.E.*, June 1942, Part III, p. 76.
22. *Transmission Networks and Wave Filters*. T. E. Shea. Text-book.

CHAPTER 4

RADIO FREQUENCY AMPLIFICATION

4.1. Introduction. In this chapter radio frequency amplification is considered in relation to amplifiers which are tunable over a range of frequencies from a minimum of 150 kc/s to a maximum of 50 Mc/s. Special problems are involved due to the need for covering a range of frequencies ; for example, the gain and selectivity of an amplifier may vary widely over the frequency range, leading perhaps to instability at the high frequency end. Again owing to interelectrode capacitance coupling the grid input admittance of an R.F. amplifier valve (see section 2.8) may be comparable with the reciprocal of the dynamic resistance of the tuned circuit to which it is connected. An anode-grid capacitance of 0·005 $\mu\mu$F can produce a high input admittance at frequencies of the order of 1·5 Mc/s, if the anode external load impedance is high. At higher frequencies (50 Mc/s) in the short wave band other effects such as the inductance of the cathode-earth connection and electron transit time tend to give a very high input admittance, so that gain is limited.

The coupling impedances between the amplifier stages generally consist of parallel tuned circuits since high selectivity and gain are required (aperiodic circuits are rarely used). Band-pass filters are often employed, and the design of two coupled tuned circuits producing a band-pass effect is given in detail in Section 7.3 on I.F. amplifiers. We shall therefore only deal with those aspects of design which result from variable tuning. It is recommended that Sections 7.2 to 7.5 should be read before Sections 4.5 and 6, since the I.F. amplifier band-pass filter is a special case of the tuned R.F. filter operating at a fixed frequency.

The circuit analysis, which follows, begins with the simplest tuned filter, the parallel resonant circuit, and progresses to the more complicated band-pass filter. The R.F. valve, generally a multi-electrode valve, [19] is considered as a generator of constant current $I_a = g_m.E_g$ and its resistance R_a, except where stated otherwise, is assumed to be much larger than the resonant impedance of the tuned filter in its anode circuit.

4.2. The Parallel Resonant Circuit.

4.2.1. Magnification. The most important criterion for a coil forming part of a tuned circuit is its magnification. This term, designated by Q, is the ratio at resonance of the voltage across the inductance/or capacitance to the voltage injected in series with the circuit.

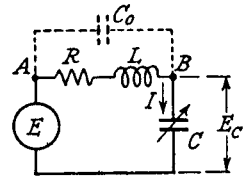


FIG. 4.1.—The Voltage Magnification of a Coil.

Referring to Fig. 4.1.

$$Q = \frac{E_L}{E} = \frac{I\omega_r L}{E}$$

Alternatively $$Q = \frac{Ec}{E} = \frac{I}{\omega_r CE}$$

But $$E = I\left[R + j\left(\omega_r L - \frac{1}{\omega_r C}\right)\right]$$

$$= IR \text{ at resonance, for } \omega_r L = \frac{1}{\omega_r C}$$

$$\therefore Q = \frac{\omega_r L}{R} = \frac{1}{\omega_r CR}. \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.1.$$

4.2.2. The Impedance of a Parallel Resonant Circuit and its Equivalent Series and Parallel Circuits. The impedance of the parallel resonant circuit of Fig. 4.2 is

$$Z = \frac{R + j\omega L}{j\omega C} \Big/ R + j\left(\omega L - \frac{1}{\omega C}\right) \quad . \quad . \quad 4.2.$$

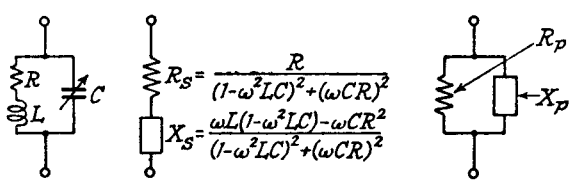


FIG. 4.2.—The Parallel Resonant Circuit and its Equivalent Series and Parallel Circuits.

$$Rp = \frac{R^2 + \omega^2L^2}{R}; \quad Xp = \frac{R^2 + \omega^2L^2}{\omega L(1 - \omega^2LC) - \omega CR^2}$$

At resonance if $R \ll \omega_r L$

$$Z_r = \frac{L}{CR} = \frac{\omega_r^2 L^2}{R} = Q\omega_r L = \frac{Q}{\omega_r C} \quad . \quad . \quad 4.3.$$

The resonant impedance Z_r is thus a resistance, and it is often called the dynamic resistance of the circuit and denoted by R_D.

To determine its frequency discriminating action, the impedance must be calculated for frequencies other than resonance.

Rewriting 4.2

$$Z = \frac{R+j\omega L}{1 - \omega^2 LC + j\omega CR}.$$

Rationalizing

$$Z = \frac{(R+j\omega L)(1 - \omega^2 LC - j\omega CR)}{(1 - \omega^2 LC)^2 + (\omega CR)^2}$$

$$= \frac{R}{(1 - \omega^2 LC)^2 + (\omega CR)^2} + j\frac{\omega L(1 - \omega^2 LC) - \omega CR^2}{(1 - \omega^2 LC)^2 + (\omega CR)^2} \quad . \quad 4.4$$

$$= R_S + jX_S$$

where R_S and X_S are the equivalent series resistance and reactance.

An interesting point to observe is that reactance resonance, i.e., $X_S = 0$, gives

$$\omega_r L(1 - \omega_r^2 LC) - \omega_r CR^2 = 0$$

or

$$\omega_r = \sqrt{\frac{1}{LC}\left(1 - \frac{CR^2}{L}\right)} \quad . \quad . \quad . \quad 4.5.$$

Replacing this in R_S (4.4) gives $Z_r = \frac{L}{CR}$, which is identical with expression 4.3, for which R was assumed to be much less than $\omega_r L$.

The equivalent series circuit is rarely so useful as the equivalent parallel circuit. The latter is determined by considering the admittance Y.

Thus

$$Y = j\omega C + \frac{1}{R+j\omega L}$$

$$= \frac{(1 - \omega^2 LC + j\omega CR)(R - j\omega L)}{R^2 + \omega^2 L^2}$$

$$= \frac{R}{R^2 + \omega^2 L^2} + j\frac{\omega CR^2 - \omega L(1 - \omega^2 LC)}{R^2 + \omega^2 L^2} \quad . \quad . \quad 4.6$$

$$= g + jb$$

$$= \frac{1}{R_p} + \frac{1}{jX_p}$$

where R_p and X_p are the equivalent parallel resistance and reactance. By making the assumption that $R \ll \omega L$ most useful approximate formulae for R_p and X_p are obtained as follows

$$\left.\begin{aligned} R_p &\simeq \frac{\omega^2 L^2}{R} \simeq R_D \\ X_p &\simeq \frac{\omega L}{1-\omega^2 LC} \end{aligned}\right\} \quad . \quad . \quad . \quad 4.7.$$

It is important to note that X_s and X_p are inductive for frequencies below and are capacitive for frequencies above the resonant frequency.

4.2.3. The Selectivity Characteristic.[6] The selectivity characteristic of a parallel tuned circuit may be defined as the ratio of the resonant impedance to the impedance at any given off-tune frequency. It is more conveniently expressed as loss in decibels, $20 \log_{10} \frac{|Z_r|}{|Z|}$, the reference level corresponding to 0 db. being the resonant impedance Z_r. Expression 4.2 simplifies to

$$\begin{aligned} Z &= \frac{\frac{L}{CR}}{1+j\frac{\omega_r L}{R}\left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega}\right)} \\ &= \frac{R_D}{1+jQ\left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega}\right)} \quad . \quad . \quad . \quad 4.8a \end{aligned}$$

when $R \ll \omega L$ and $\omega_r = \frac{1}{\sqrt{LC}}$.

But
$$\begin{aligned} \frac{\omega}{\omega_r} - \frac{\omega_r}{\omega} &= \frac{\omega^2 - \omega_r^2}{\omega_r \omega} = \frac{f^2 - f_r^2}{f_r f} \\ &= \frac{(f-f_r)(f+f_r)}{f_r f} \end{aligned}$$

If $f = f_r + \Delta f$, where Δf is the off-tune frequency

$$\begin{aligned} \frac{(f-f_r)(f+f_r)}{f_r f} &= \frac{\Delta f(2f_r + \Delta f)}{f_r(f_r + \Delta f)} \\ &\simeq \frac{2\Delta f}{f_r} \text{ when } \Delta f \ll f_r. \end{aligned}$$

For convenience we will designate $\frac{2\Delta f}{f_r}$ by F.

Expression 4.8*a* now becomes

$$Z = \frac{R_D}{1+jQF} \quad . \quad . \quad . \quad . \quad 4.8b$$

hence
$$20 \log_{10} \frac{|Z_r|}{|Z|} = 20 \log_{10} \sqrt{1+Q^2F^2}$$
$$= 10 \log_{10} (1+Q^2F^2) \quad . \quad . \quad 4.8c.$$

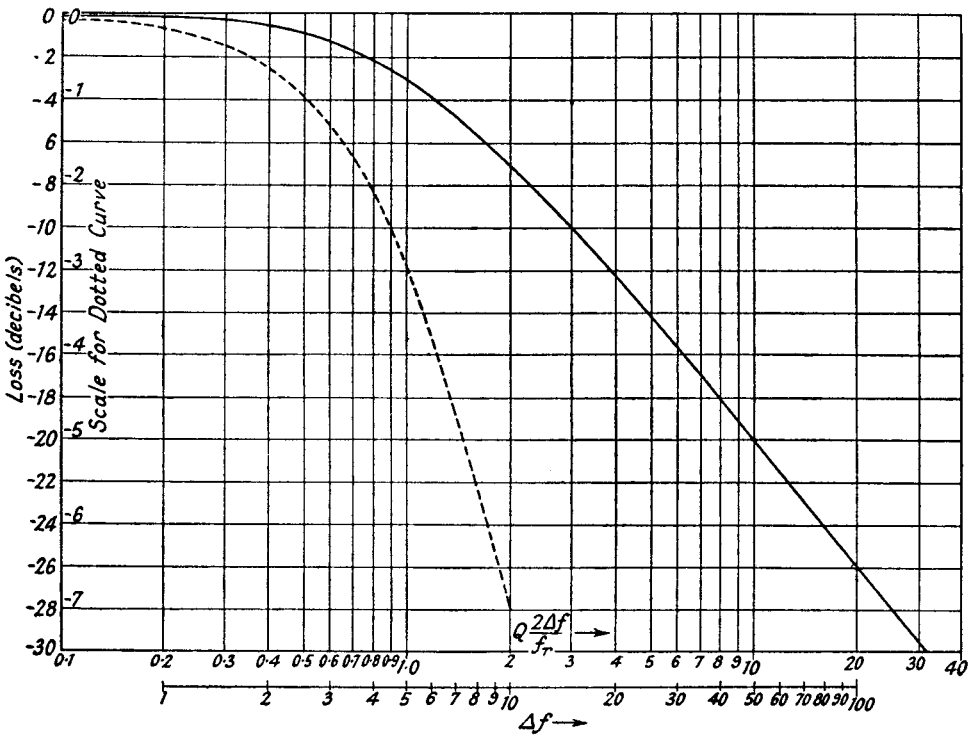


FIG. 4.3.—The Generalized Selectivity Curve for a Single Tuned Circuit.

Expression 4.8*c* is most important because it gives a generalized selectivity curve,[14] which enables the frequency response of any parallel tuned circuit to be obtained immediately. In Fig. 4.3 a curve is plotted of $10 \log_{10} (1+Q^2F^2)$ against QF. The scale of the latter is logarithmic. Only one-half of the curve is plotted since it is symmetrical about f_r, and the loss at equal numerical positive and negative values of Δf is the same.

Since $QF = \frac{2Q\Delta f}{f_r}$, $\Delta f = \frac{QF.f_r}{2Q}$

and $$\log_{10} \Delta f = \log_{10} QF + \log_{10} f_r - \log_{10} 2Q,$$

but for a given circuit f_r and Q are fixed, therefore

$$\log_{10} \Delta f = \log_{10} QF + \text{constant}.$$

Hence by sliding and locating correctly a logarithmic scale (identical with the QF scale) marked in Δf, the off-tune frequency, underneath the QF scale, the selectivity characteristic at different off-tune frequencies may be read directly. The correct position is obtained by locating $\Delta f' = \frac{f_r}{2Q}$ immediately beneath $QF = 1$. As an example, let $f_r = 1{,}000$ kc/s and $Q = 100$ be the constants of a particular tuned circuit, then

$$\Delta f' = \frac{f_r}{2Q} = \frac{1000}{200} = 5 \text{ kc/s}.$$

The Δf scale is therefore adjusted as shown in Fig. 4.3, with $\Delta f' = 5$ kc/s against $QF = 1$. The loss at any value of Δf may now be read directly, e.g., $\Delta f = 5$ and 10 kc/s gives a loss of 3 and 6·9 db. respectively.

The width of the pass band of any filter circuit is important, and in the case of a tuned circuit it is defined as the frequency range over which the loss is not greater than 3 db., i.e., between the frequency points at which the voltage has fallen to 0·707 of its maximum value. Referring to Fig. 4.3, the off-tune frequency corresponding to a loss of 3 db. is defined by

$$QF = \frac{Q2\Delta f}{f_r} = 1 \quad . \quad . \quad . \quad 4.9$$

so that the pass band $2\Delta f = \frac{f_r}{Q}$.

4.2.4. Constant Selectivity over a Range of Tuning Frequencies. In an ideal R.F. amplifier the overall selectivity should be independent of the tuning frequencies. From expression 4.8c we see that this requires a constant loss at a given off-tune frequency, i.e.,

$$1 + (QF)^2 = \text{constant at } \Delta f = \Delta f'$$

or $$QF' = \text{constant} = \frac{2Q\Delta f'}{f_r}$$

$$Q \propto f_r.$$

Hence for a constant selectivity characteristic over a range of resonant frequencies Q must vary linearly with f_r. For some types of coil, notably iron-cored coils, Q decreases as the frequency rises,

and the selectivity at low frequencies is much better than at high frequencies. This decrease of Q with increase of frequency has one advantage; the amplification of an R.F. stage is directly proportional to the dynamic resistance $R_D = Q\omega_r L$, which tends to remain more constant over the tuning range when Q decreases as the frequency increases. Permeability tuning,[16] i.e., tuning by variation of inductance with fixed capacitance, produces less change of selectivity and amplification over the tuning range because L decreases as the frequency increases. Its chief disadvantage is that variation of inductance is usually more difficult and costly to achieve than variation of capacitance over a wide tuning range.

In order to obtain constant selectivity from the normal type of coil it is necessary either to apply controlled reaction having greater effect at high than at low tuning frequencies, or to introduce resistance to reduce the Q of the coil as the tuning frequency is decreased. In the first method it is difficult to design a simple feedback circuit giving the desired Q compensation. The second method,[36] though not so satisfactory since it involves degrading the circuit, can be achieved by means of a simple resistance-reactance network.

It is essential in this method to obtain the highest Q possible at the highest frequency since the selectivity characteristic over the tuning range is to be entirely determined by the selectivity at the highest frequency. Let us assume that the highest possible Q is 150 for a medium wave band tuning inductance of 156 μH and that the Q varies as shown in Fig. 4.4. The variation of coil resistance is also plotted. Now $Q = \dfrac{\omega_r L}{R} = \dfrac{2\pi L}{R} f_r$, so that if Q is to be proportional to f_r the coil resistance R must remain constant over the range of f_r. It is therefore necessary to insert a resistance, in series with the coil, having a resistance variation with frequency shown by the curve $R_{required}$ (obtained from $R_{1,500\ kc/s} - R_{coil}$) in Fig. 4.4.

Let us consider the network shown in Fig. 4.4. It can be resolved into a series circuit of resistance and reactance of values, $R_2 = \dfrac{R_1 X_1^2}{R_1^2 + X_1^2}$ and $X_2 = X_1 \dfrac{R_1^2}{R_1^2 + X_1^2}$, where X_1 is the reactance of the L.C. arm. By a suitable choice of L_1 and C_1, X_1, and hence R_2, may be made zero at the highest tuning frequency. As the frequency is decreased X_1, and consequently R_2, increases, and this is the requirement for the compensating curve. By inserting this network in series with the coil it is possible to preserve an almost

constant resistance in the inductive arm over the tuning range. The variation of R_2 from the network cannot be made to give the exact R_2 curve required, but since there are three variables R_1, L_1 and C_1, three points of intersection are possible. The reactance term produced by the compensating network is usually very small

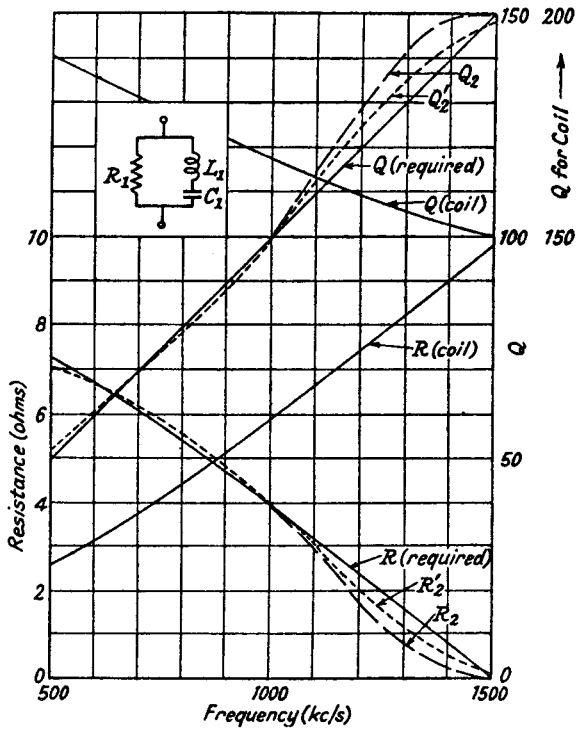


FIG. 4.4.—Correction Curves for a Tuned Circuit of Constant Q Value.

and can be neglected in comparison with the coil reactance. The reactance X_1 in the above expressions may be written

$$X_1 = j\omega L_1\left(1 - \frac{\omega_r^2}{\omega^2}\right)$$

$$= j\omega L_1\left(1 - \frac{f_r^2}{f^2}\right)$$

where

$$\omega_r = \frac{1}{\sqrt{L_1 C_1}}$$

To calculate the values of R_1, L_1 and C_1, we must fix the frequencies at which $R_2 = R_{required}$. Let us assume these are 600,

1,000 and 1,500 kc/s when $R_{required} = 6{\cdot}63$, $3{\cdot}94$ and 0 ohms respectively.

From $R_2 = 0$ at $f = 1{,}500$ kc/s, $f_r = 1{,}500$ kc/s.

At $f = 1{,}000$ kc/s.

$$3{\cdot}94 = R_1 . \frac{[6{\cdot}28L_1(1 - (1{\cdot}5)^2)]^2}{R_1^2 + [6{\cdot}28L_1(1 - (1{\cdot}5)^2)]^2}$$

L_1 is in μH.

Hence
$$L_1^2 = \frac{3{\cdot}94R_1^2}{61{\cdot}8R_1 - 243} \quad . \quad . \quad . \quad . \quad . \quad 4.10.$$

For $f = 600$ kc/s

$$6{\cdot}63 = \frac{R_1\left[6{\cdot}28 \times 0{\cdot}6L_1\left(1 - \left(\frac{1{\cdot}5}{0{\cdot}6}\right)^2\right)\right]^2}{R_1^2 + \left[6{\cdot}28 \times 0{\cdot}6L_1\left(1 - \left(\frac{1{\cdot}5}{0{\cdot}6}\right)^2\right)\right]^2} . \quad 4.11.$$

Replacing L_1 in 4.11 by 4.10 we find

$$R_1 = 7{\cdot}58\Omega$$

and from 4.10

$$L_1 = 1{\cdot}01\ \mu\text{H}$$

$$C_1 = \frac{1}{\omega_r^2 L} = 0{\cdot}01116\ \mu\text{F}.$$

The variation of R_2 with frequency is plotted in Fig. 4.4 (curve R_2). The variation of Q is also shown, and it can be seen to be almost identical with the required Q curve from 500 to 1,000 kc/s. Above 1,000 kc/s the departure is quite noticeable. If a calculation is made assuming f_r to be 1,600 kc/s, the curves R_2' and Q_2' result, and the discrepancy from 1,000 to 1,500 kc/s is much reduced, but from 600 to 1,000 kc/s it is less satisfactory. The values of the circuit constants are then

$$R_1 = 7{\cdot}8 \text{ ohms}, \quad L_1 = 0{\cdot}81\ \mu\text{H} \text{ and } C_1 = 0{\cdot}01225\ \mu\text{F}.$$

by making f_r still higher, the Q curve may be made to approximate even closer to the required up to 1,400 kc/s, but above this frequency the error is increased. The series reactance term (X_2) for the above R_1, L_1 and C_1 values has maximum effect at the lowest frequency (500 kc/s), and it reduces the inductance of the arm in which it is connected by approximately 0·5%, a negligible amount.

4.3. Coil Characteristics at Radio Frequencies.

4.3.1. Introduction. A coil possesses three characteristics of major importance when it is used for radio frequency amplifica-

tion purposes; they are its inductance, A.C. resistance and self-capacitance. To prevent undesirable pick-up and interaction between input and output circuits a metal screen is required round the coil, and its effect on the coil characteristics needs to be known.

4.3.2. The Inductance of a Coil. Calculation of the inductance of various types of air cored coils has formed the subject of many papers and it is possible to achieve very high accuracy of computation [1, 2] with some of the more common shapes of coil. For most practical purposes, however, a computation accuracy from 1 to 5% is sufficient, and the following simple formulae due to Wheeler [4] will be found most useful. Taking first the single-

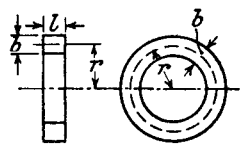


FIG. 4.5.—A Cylindrical Coil with Rectangular Section Winding Area.

layer helical solenoid—this is the coil of Fig. 4.5 with b = diameter of the wire forming the single layer—an approximate formula is

$$L = \frac{r^2N^2}{9r+10l}\ \mu\text{H} \quad . \quad . \quad . \quad . \quad 4.12$$

where l = length of the winding in inches
,, N = total number of turns on coil
and r = radius of the coil in inches to the centre of the wire layer. The error involved is less than 1% so long as $l > 0{\cdot}8r$. For values of l between $0{\cdot}8r$ to $0{\cdot}2r$ the following formula may be used with an error not exceeding 5%.

$$L = \frac{r^2N^2}{8r+11l} \quad . \quad . \quad . \quad . \quad 4.13.$$

An error not exceeding 2% over a range of r from 0 to $10l$ is obtained from

$$L = \frac{r^2N^2}{9r - \frac{r^2}{5l} + 10l} \quad . \quad . \quad . \quad 4.14.$$

Expression 4.13 is applicable to the thin spiral coil of one-wire thickness by replacing l by the breadth b, and taking r as the radius to the centre of the spiral face. The error is not greater than 5% so long as $b > 0{\cdot}2r$.

As an example of the single layer coil let us calculate the in-

ductance of a coil of 26 S.W.G. wire wound on a former of external radius 1 inch and 2 inches winding length.

Radius of wire = 0·009 inch.
Turns per inch of winding length with S.C.C. insulation = 50
Total number of turns = 100

using 4.12.

$$L = \frac{(1{\cdot}009)^2 100^2}{9(1{\cdot}009)+10(2)}$$
$$= 350\ \mu\text{H}.$$

A single-wire spiral of 1 inch breadth and inner radius of 1 inch using the same wire gives

$$b = 1,\ r = 1{\cdot}5,\ N = 50$$

and from 4.13.

$$L = \frac{(1{\cdot}5)^2 50^2}{12+11}$$
$$= 244\ \mu\text{H}.$$

Multilayer coils of approximately square section winding area may be calculated from

$$L = \frac{0{\cdot}8r^2N^2}{6r+9l+10b} \quad . \quad . \quad . \quad . \quad 4.15$$

with an error not exceeding 1% when l and b are approximately equal to $\frac{r}{2}$. A more accurate formula for small coil widths is

$$L = \frac{rN^2}{13\,5}\log_{10}\left(\frac{4{\cdot}9r}{l+b}\right)\mu\text{H} \quad . \quad . \quad . \quad 4.16$$

when the error is less than 3% if $(b+l) \leq r$.

Using 4.15, we will calculate the inductance of a coil of 26 S.W.G. S.C.C. wire of inner radius 1 inch, breadth 0·75 inch, and length 1 inch.

$$b = 0{\cdot}75 \text{ inch},\ l = 1 \text{ inch},\ r = 1+\frac{0{\cdot}75}{2} = 1{\cdot}375 \text{ inch}.$$

Total turns = winding area × (turns per inch)²
$= 0{\cdot}75 \times 1 \times 2{,}500$
$= 1{,}875.$

$$L = \frac{0{\cdot}8(1{\cdot}375)^2(1875)^2}{6(1{\cdot}375)+9+7{\cdot}5}$$
$$= 215{,}000\ \mu\text{H}.$$

The accuracy of all the above formulae is reduced if the turn apacing is large and the equivalent inductance value at any par-

ticular frequency is influenced by skin effect and the coil distributed capacitance.

The inductance of single and multilayer coils has been also set out in the form of a series of abacs.[34]

For iron-cored coils the inductance depends on the permeability of the core material, and details are usually given by the makers of the latter.

4.3.3. The A.C. Resistance of a Coil. The effective resistance of a straight wire to A.C. currents is greater than to D.C. and it increases as frequency is increased. This is due to the fact that the magnetic flux, which produces a circular field around the conductor, also exists inside the conductor. The latter can be considered as made up of a series of concentric cylinders of elemental thickness, and the magnetic field surrounding each progressively decreases as the outside of the wire is approached. The surrounding field introduces an inductance component in each cylindrical element of the conductor and this component is greatest for that cylinder having the greatest surrounding magnetic field, i.e., the one nearest the centre of the wire. This inductance offers an impedance to the flow of current and there is therefore a tendency for the current to crowd to the outer edges or skin of the conductor. Increasing frequency means increasing effective inductive reactance and less current flow through the centre of the wire. At radio frequencies of about 1,000 kc/s, the current at the centre of the wire may be practically zero, and there is then little difference between a solid conductor and a thin hollow tube of the same external diameter. This concentration of current causes the resistance of a conductor at radio frequencies to be many times greater than its D.C. resistance.

If the wire is coiled, the current area is further restricted to the inside face nearest the coil former due to the magnetic field from adjacent turns, and there is usually a considerable increase in R.F. resistance. Coiling a wire initially increases the inductance at a much greater rate than resistance and there are optimum coil dimensions and wire diameters for minimum resistance with a given inductance. The classic paper on the subject is that [3, 18] by Butterworth. The ratio of A.C. to D.C. resistance is proportional to $\sqrt{\frac{\mu f}{\rho}}$, where μ is the average permeability of the material (conductor and core) in the magnetic field produced by the current in the conductor, f is the frequency, and ρ is the resistivity of the conductor. It is interesting to note that a magnetic material in the circuit (this

means increased μ) increases the A.C./D.C. coil resistance ratio as does decrease of ρ, i.e., a good conductor shows much greater change of resistance under A.C. operation than a poor conductor, and an iron-cored coil shows a much greater ratio change of resistance than does an air-cored coil.

General conclusions [28] on realizing minimum resistance depend on what parameters are fixed. There is an optimum wire diameter to winding pitch ratio, and it is given by $d_w \simeq \frac{0{\cdot}7l}{N}$, where l = length of coil and N = turns per layer. For a given diameter of wire, minimum resistance is obtained in single layer coils with a length to radius ratio $\left(\frac{l}{r}\right)$ of 0·6 to 0·9, single wire spirals require a breadth to mean radius ratio of approximately 1, and for multilayer coils the following relationship should be satisfied $5b+3l \simeq 2(b+r)$. Resistance is decreased for a given inductance by using larger gauge wire, and when there are no restrictions on wire size, resistance, for a given inductance value, coil radius and optimum wire diameter, is decreased by increasing the length of a single layer solenoid—this really means using a larger gauge wire as length is increased—but there is little advantage in making $l > 4r$. This gives a more economical shape than that cited above for a fixed wire diameter and is the one most used in practice. Multi-strand litz wire is effective for reducing resistance at frequencies up to about 5 Mc/s.

Additional sources of increased coil resistance are to be found in dielectric losses in the insulating material (these are greatly increased if the insulation is subject to moisture absorption, and it is essential to treat the coil with a non-hygroscopic varnish), coil self-capacitance, and eddy currents in the shield surrounding the coil. Coil self-capacitance is shown in Section 4.3.4 to increase the effective resistance of a coil. Losses in the coil shield are reduced by reducing the coupling to the screen, i.e., by increasing the ratio of shield-to-coil diameter (the ratio[28] should not be less than 2/1) and by using screen material of lowest resistivity.

4.3.4. Coil Self-Capacitance. Coil self-capacitance is caused by electrostatic coupling between turns, and between turns and earth. It is highest in multi-layer coils, in which there is a high potential between layers. To reduce self-capacitance it is essential to separate high potential points, thus a two-layer coil is preferably wound with interleaved turns—the third and fifth turns being wound on top of, and between the first and second and second and

fourth turns, etc. Winding in two separate layers, one on top of the other, causes a large self-capacitance. Large inductances, e.g., for R.F. chokes, are best separated into a number of independent sections, so reducing the potential between the ends of each section. The self-capacitance acts in parallel with the coil and tends to increase the equivalent inductance and series resistance. Referring to Fig. 4.1, let us suppose that the actual inductance L and series resistance R of the coil are paralleled by a self-capacitance C_0 (shown dotted). The impedance of the coil circuit is

$$\begin{aligned} Z &= \frac{\frac{R+j\omega L}{j\omega C_0}}{R+j\omega L+\frac{1}{j\omega C_0}} \\ &= \frac{R+j\omega L}{1-\omega^2 LC_0+j\omega C_0 R} \\ &= \frac{(R+j\omega L)(1-\omega^2 LC_0-j\omega C_0 R)}{(1-\omega^2 LC_0)^2+(\omega C_0 R)^2} \\ &= \frac{R}{(1-\omega^2 LC_0)^2+(\omega C_0 R)^2}+\frac{j\omega L(1-\omega^2 LC_0)-\omega C_0 R^2}{(1-\omega^2 LC_0)^2+(\omega C_0 R)^2} \qquad 4.17 \\ &= R_A+j\omega L_A \end{aligned}$$

where R_A and L_A are the apparent resistance and inductance of the coil.

If C_0 is small, the natural frequency of the coil is high and $(\omega C_0 R)^2$ and $\omega C_0 R^2$ can be neglected in comparison with the other factors.

$$\therefore\ R_A = \frac{R}{(1-\omega^2 LC_0)^2} \quad . \quad . \quad . \quad . \quad . \quad 4.18a$$

$$L_A = \frac{L}{(1-\omega^2 LC_0)} \quad . \quad . \quad . \quad . \quad . \quad 4.18b$$

and

$$\begin{aligned} Q_A &= \frac{\omega L_A}{R_A} = \frac{\omega L}{R}(1-\omega^2 LC_0) \quad . \quad . \quad . \quad 4.18c \\ &= Q(1-\omega^2 LC_0). \end{aligned}$$

Since the natural frequency due to resonance of L and C_0 can be assumed to be much higher than the highest tuning frequency in the range over which the coil is to be used, $\omega^2 LC_0 \ll 1$ and $(1-\omega^2 LC_0) < 1$. Hence the apparent series resistance and inductance of the coil are increased by the self capacitance; R_A is increased at a much greater rate than L_A, so that the apparent or effective Q of the circuit is reduced. The ratios of R_A/R, L_A/L and Q_A/Q vary with frequency, the first two increasing and the third

decreasing as the tuning frequency is increased. The tuning frequency condition is given by

$$\omega L_A = \frac{1}{\omega C}$$

where C = tuning capacitance setting required for resonance.

$$\omega L_A = \frac{\omega L}{1 - \omega^2 L C_0} = \frac{1}{\omega C}$$

$$\omega^2 LC = 1 - \omega^2 L C_0$$

or

$$\omega^2 = \frac{1}{L(C+C_0)}.$$

Replacing ω^2 in 4.18*a*, *b* and *c* by this value gives

$$R_A = R\left(1+\frac{C_0}{C}\right)^2. \quad . \quad . \quad . \quad 4.19a$$

$$L_A = L\left(1+\frac{C_0}{C}\right) . \quad . \quad . \quad . \quad 4.19b$$

$$Q_A = Q/\left(1+\frac{C_0}{C}\right) . \quad . \quad . \quad . \quad 4.19c.$$

The above expressions show quite clearly that the effect of self-capacitance is most pronounced at the highest tuning frequencies, for which C is small.

Average values of self-capacitance are 10 — 13 $\mu\mu$F, 7 to 10 $\mu\mu$F and 4 to 7 $\mu\mu$F for the long, medium and short wave ranges.

4.3.5. The Effect of Screening on the Inductance and Resistance of a Coil.[20] A shield surrounding a coil acts as a short-circuited turn coupled to the coil and reflects an impedance into the coil in the same manner as the aerial reflects an impedance into the first tuned circuit of a receiver (Section 3.4). The reflected impedance is equivalent to a resistance and negative inductance of values

$$\frac{\omega^2 M^2 R_s}{|Z_s|^2} \text{ and } -\frac{\omega^2 M^2 L_s}{|Z_s|^2} \text{ (expression 3.18)}$$

in series with the inductance of the coil, where $Z_s = R_s + j\omega L_s$ and R_s and L_s are the equivalent resistance and inductance of the short-circuited turn representing the shield. The effective resistance of the coil is therefore increased to

$$R\left[1+\frac{\omega^2 M^2 R_s}{|Z_s|^2 R}\right]$$

and its effective inductance reduced to

$$L\left[1 - \frac{\omega^2 M^2 L_s}{|Z_s|^2 L}\right]$$

As estimation of the effect of enclosing a coil in a screen has been made by Kaden [17] and more recently by Bogle,[33] who gives some very useful empirical formulae for the change in inductance and resistance. The following formula, applicable to single-layer coils, gives the ratio of the change in inductance (the difference between the unscreened and screened inductance values, L_1 and L_2 respectively) to the unscreened inductance value when the screen is of non-magnetic material of good electrical conductivity.

$$\frac{\Delta L}{L_1} = \frac{L_1 - L_2}{L_1} = \frac{\frac{l}{g}}{\left(\frac{l}{g} + 1{\cdot}55\right)} \left(\frac{r_c}{r_s}\right)^2 . \quad . \quad 4.20$$

where l = length of coil
$g = r_s - r_c$ = radial gap between the screen and coil
r_s = radius of screen.
r_c = radius of the coil.

Dimensions may be in inches or centimetres, as the result is in ratio form.

This formula has an error not exceeding 2% so long as the distance between the ends of the coil and the screen is not less than $2g$.

Thus if we place a single layer coil of winding length 1 inch and radius 1 inch inside a screen of radius 1·5 inches the ratio reduction of inductance is

$$\frac{\Delta L}{L_1} = \frac{\frac{1}{0{\cdot}5}}{\left(\frac{1}{0{\cdot}5} + 1{\cdot}55\right)} \left(\frac{1{\cdot}0}{1{\cdot}5}\right)^2$$

$$= \frac{2}{3{\cdot}55} \times 0{\cdot}445 = 0{\cdot}251.$$

Under most practical conditions the thickness of the shield is determined by mechanical rather than electrical considerations, as the eddy currents do not penetrate deeply. The minimum thickness for adequate shielding is a function of the resistivity ρ of the material and the frequency f, the actual relationship being that the thickness (inches)

$$t > \frac{2}{2{\cdot}54\pi}\sqrt{\frac{\rho}{f}}$$

$$> 0{\cdot}251\sqrt{\frac{\rho}{f}}$$

where ρ is in micro-ohms per cubic centimetre

f is in kilocycles per second.

For copper $\rho = 1{\cdot}68\ \mu\Omega$/c.c., so that for a minimum frequency of 100 kc/s, the thickness must be greater than $0{\cdot}251\sqrt{\frac{1{\cdot}68}{100}} = 0{\cdot}0325$ inch.

Most coil screens for mechanical rigidity require to be made of at least 20 S.W.G., which gives a thickness of 0·036 inch, a value greater than that required for electrical screening at this low frequency. At higher frequencies the required thickness decreases.

Bogle also gives the following formula for the ratio of the change in inductance to the screened inductance due to eccentricity of the axes of the coil and screen.

$$\frac{\delta L}{L_2} = 0{\cdot}535\,\frac{e^2}{g^2}\,\frac{g}{l+g}\cdot\frac{r_s}{r_c}\cdot\frac{\Delta L}{L_2} \qquad . \qquad . \qquad . \qquad 4.21$$

where e = the eccentricity of the axes.

Let us suppose the coil in the first example has an eccentricity of 0·02 inch.

$$\frac{\delta L}{L_2} = 0{\cdot}535\cdot\left(\frac{0{\cdot}02}{0{\cdot}5}\right)^2\cdot\frac{0{\cdot}5}{1{\cdot}5}\cdot\frac{1{\cdot}5}{1}\cdot\frac{\Delta L}{L_2}$$

$$= 4{\cdot}28 \times 10^{-4}\frac{\Delta L}{L_2}$$

$$\frac{\Delta L}{L_2} = \frac{L_1 - L_2}{L_2} = \frac{L_1}{L_2} - 1.$$

From the previous example

$$\frac{\Delta L}{L_1} = 1 - \frac{L_2}{L_1} = 0{\cdot}251$$

$$\therefore \frac{L_2}{L_1} = 0{\cdot}749$$

and

$$\frac{L_1}{L_2} = 1{\cdot}336$$

$$\therefore \frac{\delta L}{L_2} = 4{\cdot}28 \times 10^{-4} \times 0{\cdot}336$$

$$= 1{\cdot}43 \times 10^{-4}.$$

It is clear from this that eccentricity is a second-order effect and for most practical purposes may be neglected.

The increase in resistance to be expected due to placing the coil in a screen can be stated empirically as

$$\Delta R = \frac{\Delta L}{L_1}\left[\frac{N}{l}\,\frac{r_c}{r_s}\right]^2 . 4\pi^2 r_s l\sqrt{\rho f} . 10^{-6} \text{ ohms} \quad . \quad . \quad 4.22a$$

$$= 3{\cdot}95 \times 10^{-5}\frac{\Delta L}{L_1}\,\frac{N^2 r_c^{\,2}}{l r_s}\sqrt{\rho f} \text{ ohms} \quad . \quad . \quad 4.22b$$

where N = total turns of coil

ρ = resistivity of the screen material (μ ohms per cubic centimetre)

f = frequency in kc/s.

Expression 4.22*b* gives the increment in resistance due to losses in the screen, and this will not necessarily be the actual change in resistance that would be measured for a given coil. Referring to Expression 4.19*a* in the previous section, we note that owing to self-capacitance the apparent resistance of the coil is increased by $R\left[\frac{2C_0}{C}+\frac{C_0^2}{C^2}\right]$, where C is the value of the tuning capacitance. Now the inductance of the coil is reduced by placing it in a screen so that C must be increased for a given tuning frequency ; this reduces the value of apparent resistance, thus offsetting the increase in resistance due to screen losses. In a particular case we may therefore find that screening reduces the apparent coil resistance ; however, the apparent value of Q for the screened coil will be less than that of the unscreened coil owing to the reduction in inductance by the screening.

4.4. Types of R.F. Coupling Circuits.

4.4.1. The Tapped Parallel Tuned Circuit.[29] A generator supplying a parallel-tuned circuit may be connected across the complete coil or across only a part of it. Certain advantages may be gained by tapping down the coil. If the internal resistance of the generator is less than the dynamic resistance (R_D) of the tuned circuit, the output voltage across the latter is increased and the selectivity characteristic improved by tapping down. There is actually an optimum tapping point for maximum amplification, but maximum selectivity is obtained with the lowest tap.

When the generator resistance is greater than R_D, tapping down decreases the output voltage. This has advantages in the case of a valve generator having high g_m and comparatively high anode-grid capacitance. Feedback through this capacitance may cause instability if the voltage amplification of the anode circuit is high. Tapping down enables the latter to be reduced.

To estimate the performance of the tapped circuit the impedance between the tapping points, and the overall amplification must be calculated. The equivalent circuit is shown in Fig. 4.6, and the

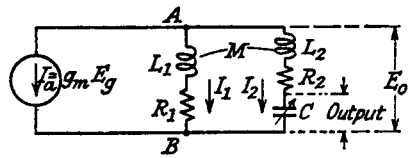


FIG. 4.6.—The Tapped Tuned Circuit.

current-voltage equations are given below. It is assumed that the valve has infinite resistance—a correction to the overall amplification is made later for a valve resistance R_a comparable with R_D—and that the resistances (R_1 and R_2) of the two sections of the coil are divided in proportion to the inductances L_1 and L_2.

$$I_a = I_1 + I_2 \quad . \quad . \quad . \quad . \quad 4.23a$$

$$I_1(R_1+j\omega L_1) - I_2 j\omega M = I_2\left[R_2+j\left(\omega L_2 - \frac{1}{\omega C}\right)\right] - I_1 j\omega M \quad 4.23b$$

$$\therefore\ I_1[R_1+j\omega(L_1+M)] = I_2\left[R_2+j\left[\omega(L_2+M) - \frac{1}{\omega C}\right]\right].$$

But the total coil inductance $L_T = L_1+L_2+2M$, so that

$$I_1[R_1+j\omega(L_1+M)] = I_2\left[R_2 - j\omega(L_1+M)+j\left(\omega L_T - \frac{1}{\omega C}\right)\right] \quad 4.23c.$$

Since $$\omega L_T - \frac{1}{\omega C} = \omega_r L_T\left(\frac{\omega}{\omega_r} - \frac{\omega_r}{\omega}\right) = \omega_r L_T F.$$

$$I_1 = \frac{I_2[R_2 - j\omega(L_1+M)+j\omega_r L_T F]}{R_1+j\omega(L_1+M)} \quad . \quad . \quad 4.23d.$$

Replacing 4.23*d* in 4.23*a*

$$I_a = \frac{I_2[R_T+j\omega_r L_T F)}{R_1+j\omega(L_1+M)} \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.23e$$

where $R_T = R_1+R_2$, the total coil resistance.

$$E_0 = I_1(R_1+j\omega L_1) - I_2 j\omega M.$$

$$= \frac{I_2.[(R_1+j\omega L_1)(R_2 - j\omega(L_1+M)+j\omega_r L_T F) - j\omega M(R_1+j\omega(L_1+M)]}{R_1+j\omega(L_1+M)} \quad 4.24$$

$$= \frac{I_a[(R_1+j\omega L_1)(R_2 - j\omega(L_1+M)+j\omega_r L_T F) - j\omega M(R_1+j\omega(L_1+M))]}{R_T+j\omega_r L_T F}.$$

The impedance between the tapping points AB is

$$Z_{AB} = \frac{E_0}{I_a} = \frac{(R_1+j\omega L_1)(R_2 - j\omega(L_1+M)+j\omega_r L_T F) - j\omega M(R_1+j\omega(L_1+M))}{R_T+j\omega_r L_T F}$$

$$= \frac{(R_1+R_2+j\omega_r L_T F)(R_1+j\omega L_1) - [R_1+j\omega(L_1+M)]^2}{R_T+j\omega_r L_T F}$$

$$= R_1+j\omega L_1 - \frac{[R_1+j\omega(L_1+M)]^2}{R_T+j\omega_r L_T F} \quad . \quad . \quad . \quad . \quad 4.25a.$$

If $$R_1 \ll j\omega L_1 \ll \frac{[R_1+j\omega(L_1+M)]^2}{R_T+j\omega_r L_T F}$$

$$Z_{AB} \simeq \frac{\omega^2(L_1+M)^2}{R_T+j\omega_r L_T F} = \frac{\omega^2(L_1+M)^2}{R_T(1+jQF)} \quad . \quad . \quad 4.25b$$

where $$Q = \frac{\omega_r L}{R_T}.$$

At resonance $$F = 0 \text{ and}$$

$$Z_{AB} \simeq \frac{\omega_r{}^2(L_1+M)^2}{R_T} \quad . \quad . \quad . \quad 4.25c.$$

The overall amplification (A) of the stage is

$$A = \frac{I_2}{E_g j\omega C}.$$

Replacing I_2 with the value from 4.23*e*

$$A = \frac{I_a[R_1+j\omega(L_1+M)]}{jE_g\omega CR_T(1+jQF)}$$

$$\simeq g_m \frac{(L_1+M)}{CR_T(1+jQF)}$$

where $$I_a = g_m E_g \text{ and } R_1 \ll \omega(L_1+M)$$

$$|A| = g_m \frac{L_1+M}{CR_T\sqrt{1+(QF)^2}} \quad . \quad . \quad . \quad 4.26a.$$

At resonance

$$|A_r| = g_m \frac{(L_1+M)}{CR_T}$$

$$= g_m R_D \frac{(L_1+M)}{L_T} \quad . \quad . \quad . \quad 4.26b$$

where $$R_D = \frac{L_T}{CR_T}.$$

The selectivity, given by the ratio of $|A_r|$ to $|A|$ is $\sqrt{1+(QF)^2}$, and this is identical with that of the untapped circuit. This is

expected since R_a is assumed to be infinite. The overall gain is thus reduced in the ratio $\frac{L_1+M}{L_T}$ by tapping down but the selectivity is unaltered. When R_a is comparable with Z_{AB} the expression for (I_1+I_2) becomes

$$I_1+I_2 = g_m E_g \cdot \frac{R_a}{R_a+Z_{AB}}$$

$$A = \frac{g_m R_a}{R_a+\frac{\omega^2(L_1+M)^2}{R_T(1+jQF)}} \cdot \frac{R_D(L_1+M)}{(1+jQF)L_T} \quad . \quad 4.27a$$

and

$$A_r = \frac{g_m R_a}{R_a+R_D\left(\frac{L_1+M}{L_T}\right)^2} \cdot \frac{R_D(L_1+M)}{L_T} \quad . \quad 4.27b.$$

Differentiating 4.27*b* with respect to (L_1+M) and equating to 0 gives for a maximum A_r

$$R_a = \frac{R_D(L_1+M)^2}{{L_T}^2} = \frac{\omega^2(L_1+M)^2}{R_T}$$

so that the optimum tapping point is at $R_a = Z_{AB(resonance)}$ and

$$A_r = \frac{1}{2} \frac{g_m R_D(L_1+M)}{L_T}.$$

The Selectivity Ratio

$$\frac{|A_r|}{|A|} = \frac{\sqrt{(R_aQF)^2+\left[R_a+R_D\left(\frac{L_1+M}{L_T}\right)^2\right]^2}}{R_a+R_D\left(\frac{L_1+M}{L_T}\right)^2}$$

$$= \sqrt{1+(QF)^2\left[\frac{R_a}{R_a+R_D\left(\frac{L_1+M}{L_T}\right)^2}\right]^2}$$

is a maximum when R_a is infinite or (L_1+M) is zero, and is $\sqrt{1+\left(\frac{QF}{2}\right)^2}$ when $R_a = R_D \cdot \left(\frac{L_1+M}{L_T}\right)^2$. Hence selectivity is halved at the optimum amplification tapping point.

4.4.2. The Transformer Coupled Tuned Circuit.[9] The tapped tuned circuit is not an ideal practical arrangement and a better method is to use transformer coupling. This has two advantages ; a coupling coil is more easily adjusted than a tapping point,

and the tuned circuit is isolated from the positive H.T. voltage. The circuit is given in Fig. 4.7 and the equations are

$$E_0 = I(R_1+j\omega L_1)+I_2 j\omega M \quad . \quad . \quad . \quad 4.28a$$

$$0 = I_2[R_T+j\omega_r L_T F]+Ij\omega M \quad . \quad . \quad 4.28b$$

where $\omega_r L_T F = \left(\omega L_T - \frac{1}{\omega C}\right)$.

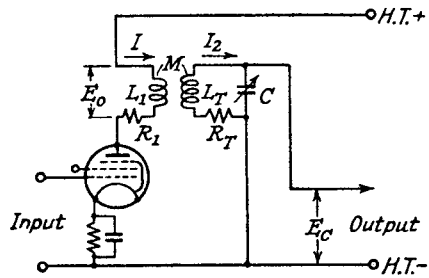


FIG. 4.7.—The Transformer Coupled Circuit.

From 4.28*b*

$$I_2 = -\frac{Ij\omega M}{R_T+j\omega_r L_T F} = -\frac{Ij\omega M}{R_T(1+jQF)}$$

$$\therefore\ E_0 = I\left[R_1+j\omega L_1+\frac{\omega^2 M^2}{R_T+j\omega_r L_T F}\right]$$

and

$$Z_0 = \frac{E_0}{I} = R_1+j\omega L_1+\frac{\omega^2 M^2}{R_T+j\omega_r L_T F}.$$

If

$$R_1+j\omega L_1 \ll \frac{\omega^2 M^2}{R_T+j\omega_r L_T F}$$

$$Z_0 = \frac{\omega^2 M^2}{R_T+j\omega_r L_T F} \quad . \quad . \quad . \quad . \quad 4.29a$$

and at resonance

$$Z_{0r} = \frac{\omega_r{}^2 M^2}{R_T} = R_D \frac{M^2}{L_T{}^2} \quad . \quad . \quad . \quad 4.29b.$$

The overall amplification A is

$$A = \frac{E_c}{E_g} = \frac{I_2}{E_g j\omega C} = \frac{-I}{j\omega C}\cdot\frac{j\omega M}{R_T(1+jQF)Eg}.$$

If R_a is very large $\mathrm{I} = \mathrm{I}_a = g_m E_g$

$$\therefore\ A = -g_m \frac{M}{CR_T(1+jQF)}$$

$$= -g_m \frac{R_D}{1+jQF}\cdot\frac{M}{L_T} \quad . \quad . \quad . \quad 4.30a.$$

The negative sign indicates that the sign of the mutual inductance coupling between primary and tuned secondary is negative, and it can be ignored when there is no other coupling between L_1 and L_T.

At resonance $$A_r = \frac{g_m M}{C R_T} = \frac{g_m R_D M}{L_T} \quad . \quad . \quad . \quad 4.30b.$$

When R_a is infinite (the assumption made above), amplification is reduced but selectivity is unchanged by reducing the coupling. For R_a comparable with Z_0, amplification is

$$A = \frac{g_m R_a}{R_a + \dfrac{\omega^2 M^2}{R_T(1+jQF)}} \cdot \frac{M}{C R_T(1+jQF)} \quad . \quad . \quad 4.31a$$

and the maximum value at resonance is obtained when $R_a = \dfrac{\omega^2 M^2}{R_T}$

Hence $$A_{r(max.)} = \frac{1}{2} g_m R_D \frac{M}{L_T} \quad . \quad . \quad . \quad 4.31b.$$

Selectivity is greatest when the coupling is least, and for maximum gain it is equivalent to that of a circuit of $\dfrac{Q}{2}$. Almost all the results are identical with those for a tapped circuit if M is replaced by (L_1+M).

It is clear from the above expressions that the primary coil inductance and resistance should be as small as possible and the coupling to the tuned circuit as large as possible. The position of the primary coil with respect to the tuned circuit is important as stray capacitive coupling between the two seriously modifies the selectivity characteristic. The primary must therefore be wound over the earthed end of the tuned coil.

4.4.3. The Choke-Capacitance Coupled Tuned Circuit. Another possible form of coupling to the tuned circuit is by R.F. choke and capacitance. Resistance-capacitance cannot be considered since a very high resistance is required for low damping of the tuned circuit, and this results in a very low D.C. anode voltage. The chief advantage of choke coupling is that the tuned circuit can be isolated from the positive H.T. voltage. The R.F. choke needs careful design; its inductance should be much higher (100 times at least) than the tuning coil and its self-capacitance should be low. The natural frequency of the choke must be well outside the frequency range of the tuned circuit if uneven frequency response and amplification characteristics are to be avoided. Generally the

natural frequency is much lower than the lowest tuned circuit frequency, and it therefore presents a high capacitive reactance over the tuning frequency range. This small equivalent capacitance can be compensated, for ganging purposes, by adjustment of the tuning capacitor trimmer. The coupling capacitance from the anode to the tuned circuit is not critical and a value of 0·0001 μF. is generally suitable. The minimum value of coupling capacitance is determined by R_D, and since R_D and this reactance are in phase quadrature a value of $X_c = \frac{R_D}{10}$ can give a voltage transfer of approximately 98%.

4.5. Band-Pass Tuned Circuits.[8, 10, 13, 15, 21, 22, 23, 24, 26]

4.5.1. Introduction. Band-pass tuned circuits are employed in R.F. amplifiers to obtain a wide and flat pass-band with sharp cut-offs. They consist usually of two tuned circuits coupled together by inductive or capacitive reactance, or a mixture of both. The theory of such band-pass filters is detailed in Chapter 7, and it is only necessary to indicate the special modifications which result from the need for variable tuning. The ideal tunable band-pass filter should maintain constant band width at all tuning frequencies, and we must examine the effect of different forms of coupling on the pass-band. The Howe method, detailed in Section 7.2, enables a preliminary examination of the pass-band to be made for a filter, having no resistance elements, as the tuned circuit resonant frequency is varied. Self-inductance coupling is not considered in the analysis, because it is but rarely employed in tunable filters.

4.5.2. Shunt Capacitance Coupling. The terms used to designate the position of the coupling are those normally employed in network theory as stated in Section 3.4.3. Taking first shunt (sometimes called common) capacitance coupling, the circuit is that

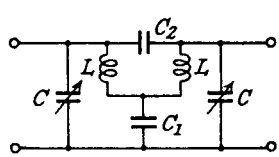


FIG. 4.8*a*.—Combined Series and Shunt Capacitance Coupling.

of Fig. 4.8*a* (showing combined shunt and series coupling) with the series capacitance C_2 open circuited. The peak frequencies are

$$f_2 = \frac{1}{2\pi\sqrt{\frac{LCC_1}{2C+C_1}}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{LC}}$$

and for the trough or mid frequency

$$f_m = \frac{1}{2\pi\sqrt{\frac{LCC_1}{C+C_1}}}.$$

The pass-band is determined by $f_2 - f_1$ and

$$f_2 - f_1 = \frac{1}{2\pi}\left(\frac{1}{\sqrt{\frac{LCC_1}{2C+C_1}}} - \frac{1}{\sqrt{LC}}\right)$$

$$= \frac{1}{2\pi\sqrt{LC}}\left(\sqrt{\frac{2C}{C_1}+1} - 1\right) \quad . \quad . \quad 4.32a.$$

Expanding $\left(\frac{2C}{C_1}+1\right)^{\frac{1}{2}}$ by the Binomial Theorem

$$\left(\frac{2C}{C_1}+1\right)^{\frac{1}{2}} = 1+\frac{1}{2}\frac{2C}{C_1} - \frac{1}{8}\left(\frac{2C}{C_1}\right)^2 + \ldots$$

$$= 1+\frac{C}{C_1} - \frac{1}{2}\left(\frac{C}{C_1}\right)^2 + \ldots$$

$= 1+\frac{C}{C_1}$, when $C \ll C_1$, as is generally the case.

Thus $$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}} \cdot \frac{C}{C_1} = f_1\frac{C}{C_1} \quad . \quad . \quad 4.32b.$$

The semi-band width is

$$\Delta f = \frac{f_2 - f_1}{2} = \frac{f_1}{2}\frac{C}{C_1} \simeq \frac{f_m}{2}\frac{C}{C_1} \quad . \quad . \quad 4.33$$

for f_1 is very nearly equal to f_m, when Δf is small compared with f_m. Since $C \simeq \frac{1}{\omega_m{}^2 L}$, $\Delta f \simeq \frac{1}{8\pi^2 LC_1 f_m}$, so that Δf decreases as f_m is increased. It is plotted against f_m as curve 1, Fig. 4.9, for $L = 156$ μH and $C_1 = 0{\cdot}0162$ μF. The latter gives $\Delta f = 10$ kc/s at $f_m = 500$ kc/s falling to $\Delta f = 3{\cdot}33$ kc/s at $f_m = 1{,}500$ kc/s.

4.5.3. Series Capacitance Coupling. The circuit diagram for series capacitance coupling is that of Fig. 4.8*a*, with the shunt capacitance C_1 short circuited. The band-width is

$$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}}\left[1 - \frac{1}{\sqrt{1+\frac{2C_2}{C}}}\right] = f_1\frac{C_2}{C}$$

and the semi-band-width

$$\Delta f = \frac{f_1}{2}\frac{C_2}{C} \simeq \frac{f_m C_2}{2C} \quad . \quad . \quad . \quad 4.34.$$

Replacing C by $\frac{1}{\omega_m^2 L}$, $\Delta f \simeq 2\pi^2 L C_2 f_m^3$, and it increases as f_m increases.

It is plotted against f_m as curve 2, Fig. 4.9, for $L = 156$ μH and $C_2 = 0{\cdot}963$ $\mu\mu$F. The value of Δf is 10 kc/s at $f_m = 1{,}500$ kc/s, falling to 0·37 kc/s at $f_m = 500$ kc/s. Curve 2 is in the reverse direction to curve 1 and it would appear possible by combining the two curves to obtain a more constant band-width over the tuning range.

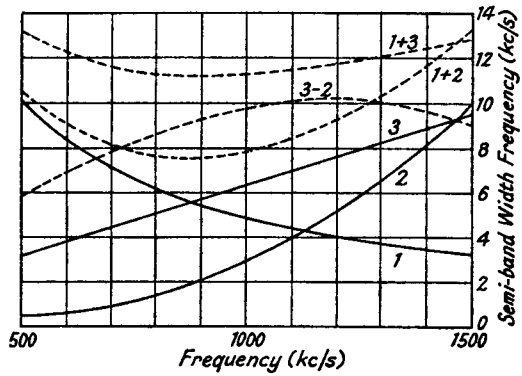


FIG. 4.9.—The Variation of Semi-Band-Width for Different Types of Coupling.
Curve 1.—Shunt Capacitance Coupling.
Curve 2.—Series Capacitance Coupling.
Curve 3.—Mutual Inductance Coupling.

4.5.4. Combined Shunt and Series Capacitance Coupling. The peak frequencies for combined capacitance coupling (Fig. 4.8*a*) are

$$f_2 = \frac{1}{2\pi\sqrt{L\left(\frac{CC_1}{2C+C_1}\right)}}; \text{ and } f_1 = \frac{1}{2\pi\sqrt{L(C+2C_2)}}$$

$$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}}\left[\sqrt{\frac{2C}{C_1}+1} - \frac{1}{\sqrt{\frac{2C_2}{C}+1}}\right]$$

$$\Delta f \simeq \frac{f_m}{2}\left[\frac{C}{C_1}+\frac{C_2}{C}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.35.$$

Hence by adding curves 1 and 2, we obtain the semi-band-width variation for the combined couplings over the tuning frequency

range. The result is curve (1+2) Fig. 4.9, showing a much reduced semi-band-width variation.

4.5.5. Mutual Inductance Coupling. The semi-band-width for mutual inductance coupling is

$$\Delta f = \frac{1}{\pi\sqrt{LC}}\left[\frac{1}{\sqrt{1-\frac{M}{L}}} - \frac{1}{\sqrt{1+\frac{M}{L}}}\right]$$

$$\Delta f \simeq \frac{f_m}{2}\frac{M}{L} \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.36.$$

Δf is plotted against f_m as curve 3, Fig. 4.9, and it rises linearly from 3·2 kc/s at 500 to 9·6 kc/s at 1,500 kc/s for $M = 2$ μH.

4.5.6. Combined Mutual Inductance and Shunt Capacitance Coupling.[5] Let us suppose that we have a combination

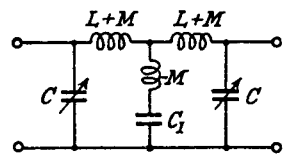


FIG. 4.8*b*.—Combined Mutual Inductance and Shunt Capacitance Coupling.

of negative mutual inductance and shunt capacitance coupling (Fig. 4.8*b*). The peak frequencies are

$$f_2 = \frac{1}{2\pi\sqrt{\frac{(L-M)CC_1}{2C+C_1}}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{(L+M)C}}$$

$$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}}\left[\sqrt{\frac{\frac{2C}{C_1}+1}{1-\frac{M}{L}}} - \frac{1}{\sqrt{1+\frac{M}{L}}}\right] \quad . \quad . \quad 4.37a.$$

Expanding by the Binomial Theorem and assuming second-order terms to be negligible

$$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}}\left[\left(1+\frac{M}{2L}+\frac{C}{C_1}\right) - \left(1-\frac{M}{2L}\right)\right]$$

$$\simeq f_m\left[\frac{M}{L}+\frac{C}{C_1}\right] \quad . \quad . \quad . \quad . \quad . \quad 4.37b$$

or

$$\Delta f = \frac{f_m}{2}\left[\frac{M}{L}+\frac{C}{C_1}\right] \quad . \quad . \quad . \quad . \quad . \quad 4.37c.$$

For combined coupling we must add the curves 1 and 3 of Fig. 4.9,

and the result is curve (1+3). It is seen that this curve is an improvement on curves 1 and 3. As a general rule we shall find that combined shunt capacitance and negative mutual inductance gives a more constant band-width than combined shunt and series capacitance coupling.

If M is positive

$$\Delta f = \frac{f_m}{2}\left[\frac{C}{C_1} - \frac{M}{L}\right] \quad . \quad . \quad . \quad 4.37d$$

and curve 3 is subtracted from curve 1, thereby increasing the band-width variation.

4.5.7. Combined Positive Mutual Inductance and Series Capacitance Coupling. The circuit diagram for this coupling is shown in Fig. 4.8c.

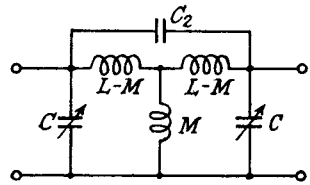


FIG. 4.8c.—Combined Mutual Inductance and Series Capacitance Coupling.

$$f_2 = \frac{1}{2\pi\sqrt{(L+M)C}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{(L-M)(C+2C_2)}}$$

$$f_2 - f_1 = \frac{1}{2\pi\sqrt{LC}}\left[\frac{1}{\sqrt{1+\frac{M}{L}}} - \frac{1}{\sqrt{\left(1-\frac{M}{L}\right)\left(1+\frac{2C_2}{C}\right)}}\right]$$

$$\simeq f_m\left[1 - \frac{M}{2L} - \left(1+\frac{M}{2L}\right)\left(1-\frac{C_2}{C}\right)\right]$$

$$\simeq f_m\left[\frac{C_2}{C} - \frac{M}{L}\right]$$

$$\Delta f \simeq \frac{f_m}{2}\left[\frac{C_2}{C} - \frac{M}{L}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.38.$$

The result of combining positive mutual inductance with series capacitance is shown in curve (3 − 2), Fig. 4.9, which is obtained by subtracting curve 2 from curve 3 multiplied by 2; i.e., M is assumed to be 4 μH instead of 2, so that the semi-band-width may have a maximum of the order of 10 kc/s. The result is not as satisfactory as curve (1+3).

Negative mutual inductance gives $\Delta f = \frac{f_m}{2}\left[\frac{C_2}{C} + \frac{M}{L}\right]$ and this

accentuates the band-width variation because curves 2 and 3, Fig. 4.9, are added.

4.6. The Design of a Tunable Band-Pass Filter. The analysis of 4.5 is most useful as a preliminary investigation and we have now to consider the practical filter, containing resistive as well as reactive components. The generalized curves for a two-circuit band-pass filter are given in Fig. 7.7, and in Section 7.5 it is shown that they are equally applicable to shunt or series reactance coupling. For the tunable filter, in which a mixture of shunt and series coupling is used, we have not discussed the application of the curves. If it is possible to replace either form of coupling by the other—the series reactance may, for example, be changed to a shunt reactance—then the curves are obviously applicable. Let us now take the series coupling reactance and try to convert it to the shunt coupling reactance. The first circuit is a symmetrical π section, whilst the second is a symmetrical T section network. The rules for conversion from one to the other are as follows: suppose the π section consists of two shunt arms of Z_1 and a series arm of Z_2, and the T section of two series arms of Z_a and a shunt arm of Z_b, then

$$Z_a = \frac{Z_1 Z_2}{2Z_1+Z_2} \text{ and } Z_b = \frac{Z_1^2}{2Z_1+Z_2}$$

If $Z_1 = \frac{1}{j\omega C}$, $Z_2 = \frac{1}{j\omega C_2}$ and $Z_2 \gg Z_1$ (this is usually true)

then $$Z_a = Z_1 \text{ and } Z_b = \frac{Z_1^2}{Z_2} = \frac{C_2}{j\omega C^2}.$$

We may therefore replace the series capacitance coupling by a shunt capacitance coupling of $\frac{C^2}{C_2}$, so that the combined coupling circuit of Fig. 4.8*a* may be replaced by a shunt coupling circuit having a shunt capacitance of C_1 and $\frac{C^2}{C_2}$ in series, i.e., $\frac{C_1 C^2}{C_1 C_2 + C^2}$.

In Section 7.4 the generalized selectivity response for the two circuit band-pass filter is shown to be

$$20 \log_{10} \frac{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}}{2Qk}$$

and in Section 7.3 the meaning of Qk is indicated to be $\frac{\text{coupling reactance}}{\text{coil resistance}}$. Hence for combined series and shunt capacitance coupling we have $Qk = \frac{\omega C^2 C_1}{R}$, for mixed positive mutual

$$C_1 C_2 + C^2$$

inductance and series capacitance coupling $Qk = \dfrac{\left(\omega M - \dfrac{C_2}{\omega C^2}\right)}{R}$, and for mixed negative mutual inductance and shunt capacitance coupling $Qk = \dfrac{\left(\omega M + \dfrac{1}{\omega C_1}\right)}{R}$.

The particular value to be assigned to Qk at any frequency is determined by Q and the maximum band-width required, and it may be obtained from the curves of Fig. 7.7. Let us imagine that a band-pass filter is to be designed to operate over the medium wave range to give a semi-band-width Δf of approximately 8 kc/s. The inductance of the coils is 156 μH, and the Q values at three frequencies are as tabulated below. Since Q is fixed, the Δf scale is automatically fixed and it is adjusted for each frequency so that $\Delta f = \dfrac{f_m}{2Q}$ is immediately under $QF = 1$. By choosing the most suitable curves to give the required semi-band-width of 8 kc/s the values of Qk are as set out in Table 4.1. The curves are redrawn to the correct off-tune frequency scale in Fig. 4.10.

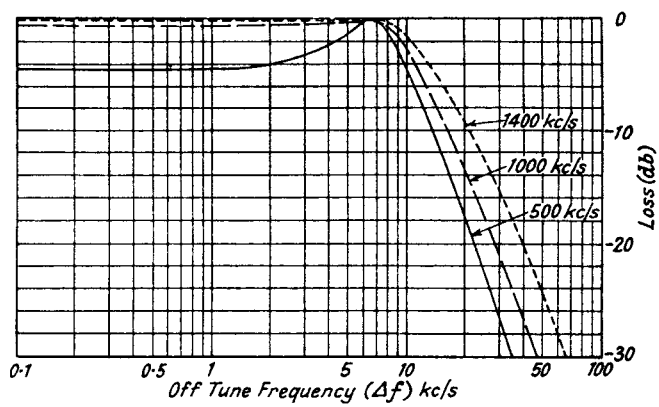


FIG. 4.10.—Selectivity Curves for a Tunable Band-Pass Filter.
[*Note.*—Read 600 kc/s for 500 kc/s in the above.]

TABLE 4.1

f_m(kc/s)	600	1,000	1,400
Q	120	100	80
$\Delta f(QF = 1)$(kc/s)	2·5	5	8·75
Qk	3	1·5	1·0
$R(\Omega)$	4·89	9·75	17·2
X coupling $= QkR(\Omega)$	14·67	14·63	17·2

Now we have the problem of finding the most suitable values of mixed coupling reactances. Let us first consider combined negative mutual inductance and shunt capacitance coupling. The coupling reactance is $\omega M + \frac{1}{\omega C_1}$, and as there are only two variables we cannot have the correct reactance at more than two frequencies. It is necessary to choose M and C_1, so as to give the nearest approach to the desired reactance over the tuning range.

Let us take some particular reference frequency f_0 at which the coupling $X = X_0 = \omega_0 M + \frac{1}{\omega_0 C_1}$ and assume $\omega_0 M$ to be some fraction A of X_0.

Thus $$\omega_0 M = AX_0$$

and $$\frac{1}{\omega_0 C_1} = (1 - A)X_0.$$

At any other frequency f

$$X = \omega M + \frac{1}{\omega C_1} = AX_0 \cdot \frac{\omega}{\omega_0} + (1 - A)X_0 \frac{\omega_0}{\omega}$$

or $$\frac{X}{X_0} = A\frac{f}{f_0} + (1 - A)\frac{f_0}{f} \quad . \quad . \quad . \quad . \quad . \quad 4.39.$$

The above expression, due to Beatty,[14] enables us to draw generalized curves, from which may be quickly estimated the values of M and C_1 to give a particular variation of coupling reactance over a given tuning range. Generalized curves of $\frac{X}{X_0}$ against $\frac{f}{f_0}$ for different values of A are plotted in Fig. 4.11. Both axes have logarithmic scales, as this allows the actual variation of X against f to be read directly by correctly registering similar logarithmic scales in X and f placed parallel to the two axes. For example, from the second column of Table 4.1. $X_0 = 14{\cdot}63\Omega$ and $f_0 = 1{,}000$ kc/s. The logarithmic reactance scale is set with $X = 14{\cdot}63\Omega$ registering with $\frac{X}{X_0} = 1$ and the frequency scale with $f = 1{,}000$ kc/s registering with $\frac{f}{f_0} = 1$ as shown in Fig. 4.11. If we choose any particular value of M such as $1{\cdot}165\ \mu$H, the value of $A = \frac{\omega_0 M}{X_0} = \frac{7{\cdot}325}{14{\cdot}63} = 0{\cdot}5$, and we can read from the curve

$A = 0{\cdot}5$ the values of coupling reactance for any desired frequency. Thus at $f = 600$ and 1,400 kc/s the coupling reactance is 16·7 and 15·4 ohms respectively. The shunt capacitance C_1 is 0·0217 μF obtained from $\dfrac{1}{\omega_0 C_1} = (1 - A)X_0$. Generally we shall wish to determine the values of M and C_1 to produce a coupling reactance variation over the tuning range as close as possible to values giving the desired band-pass characteristics. This is realized by plotting the required coupling reactance values, given in Table 4.1, against

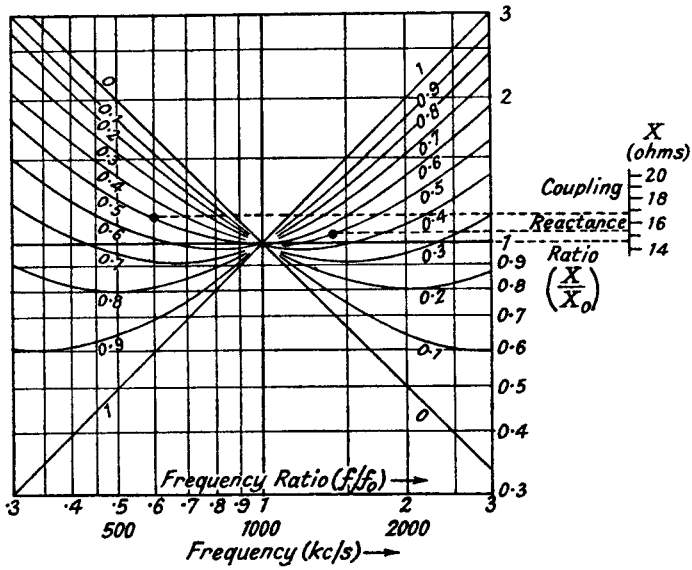


FIG. 4.11.—The Variation of Reactance Ratio Against Tuning Frequency Ratio for Combined Mutual Inductance and Shunt Capacitance Coupling.

frequency on tracing paper with X and f logarithmic scales identical with those of the generalized curves in Fig. 4.11. This curve is moved over the curves in Fig. 4.11, keeping the scale axes parallel, until the generalized curve most closely coinciding with the required coupling reactance curve is found. This occurs in our example for curve $A = 0{\cdot}6$, when $\dfrac{f}{f_0} = 1$ registers with $f = 920$ kc/s and $\dfrac{X}{X_0} = 1$ with $X = 14{\cdot}4\Omega$. The desired and generalized coupling reactance curves are curves 1 and 2 respectively in Fig. 4.12. Now we have all the data necessary to calculate M and C_1.

Thus

$$\omega_0 M = AX_0$$

$$M = \frac{AX_0}{\omega_0} = \frac{0{\cdot}6 \times 14{\cdot}4 \times 10^6}{6{\cdot}28 \times {\cdot}92 \times 10^6}\ \mu\text{H} = 1{\cdot}495\ \mu\text{H}$$

and $$\frac{1}{\omega_0 C_1} = (1 - A)X_0$$

or $$C_1 = 0{\cdot}03\ \mu\text{F}.$$

The same procedure can be applied to combined shunt and series capacitance coupling. The equivalent shunt coupling circuit gives a coupling reactance of $\left(\frac{C_2}{\omega C^2} + \frac{1}{\omega C_1}\right)$. C is a variable dependent

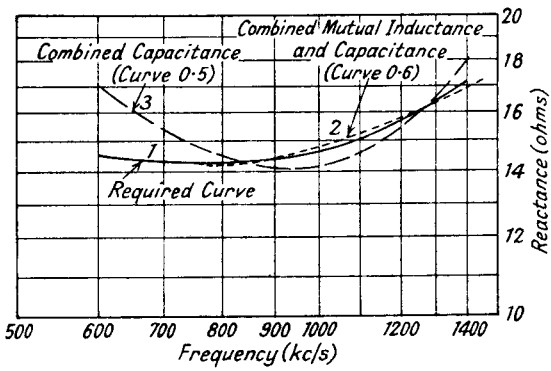


FIG. 4.12.—Required and Approximate Coupling Reactance-Frequency Curves.

on frequency and it may be eliminated by noting that $\omega = \frac{1}{\sqrt{LC}}$

$$\therefore\ X = \omega^3 L^2 C_2 + \frac{1}{\omega C_1}$$

$$= AX_0 \frac{f^3}{f_0{}^3} + (1 - A)X_0 \frac{f_0}{f} \quad . \quad . \quad 4.40$$

where $$\omega_0{}^3 L^2 C_2 = AX_0 \text{ and } \frac{1}{\omega_0 C_1} = (1 - A)X_0.$$

Generalized curves of $\frac{X}{X_0}$ against $\frac{f}{f_0}$ may be plotted as in Fig. 4.13.

The nearest curve is again found to the required $X - f$ curve from the tabulated figures. The $A = 0{\cdot}5$ curve gives the nearest approach, and this curve is curve 3 of Fig. 4.12. The agreement between the curves 1 and 3 is not so satisfactory as for MC_1 coupling

and this we should expect from the preliminary examination as illustrated in Fig. 4.9, which shows a much greater semi-band-width variation for C_1C_2 coupling. The reference point for curve 3 is $f_0 = 1{,}250$ kc/s and $X_0 = 16{\cdot}0\Omega$ and this gives

$$C_2 = \frac{AX_0}{\omega_0{}^3L^2} = \frac{0{\cdot}5 \times 16{\cdot}0 \times 10^{12} \times 10^{12}}{(6{\cdot}28 \times 1250 \times 10^3)^3.(156)^2}$$
$$= 0{\cdot}676\ \mu\mu\text{F}$$

and
$$C_1 = \frac{10^6}{(1-A)X_0.\omega_0} = \frac{10^6}{0{\cdot}5 \times 16{\cdot}0 \times 6{\cdot}28 \times 1250 \times 10^3}$$
$$= 0{\cdot}0159\ \mu\text{F}.$$

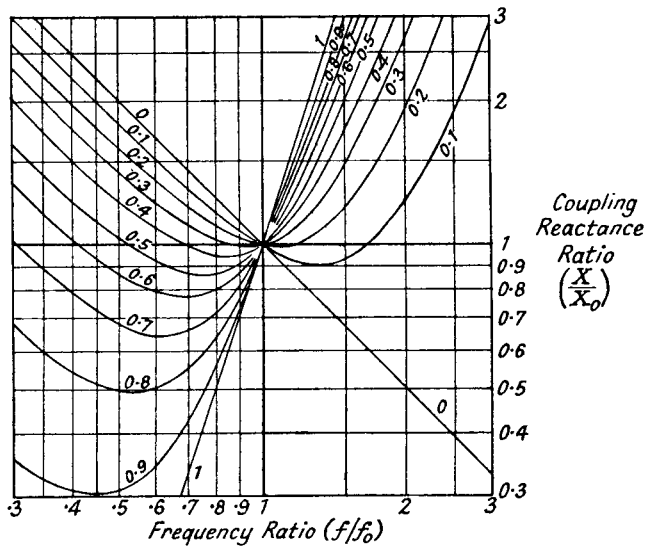


FIG. 4.13.—The Variation of Reactance Ratio Against Tuning Frequency Ratio for Combined Shunt and Series Capacitance Coupling.

Accurate adjustment of C_2 is not easily obtainable owing to its very low value. In addition, the coupling reactance variation cannot be made to follow closely the required law, so that mixed M and C_1 coupling is to be preferred.

In the above calculation we have chosen the values of Qk to give the required band-width by examination of the generalized curves. If we place a more explicit meaning to the term band-width, Qk can be calculated directly without reference to the curves. Using the definition of Section 7.7.4, the band-width becomes the difference between the frequencies at which the response is equal to

that at the minimum or trough. The transfer impedance from 7.2*c* is

$$|Z_T| = \frac{R_D Qk}{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}}.$$

For the overcoupled case the minimum value $|Z_T|$ occurs for $F = 0$, i.e.,

$$|Z_T| = \frac{QkR_D}{1+Q^2k^2}$$

and the frequencies at the edge of the pass-band are obtained by equating the two expressions, thus

$$1+Q^2k^2 = \sqrt{[1+Q^2(k^2-F_p{}^2)]^2+4Q^2F_p{}^2}$$

where $$F_p = \frac{2\Delta f_p \text{ (to edge of pass-band)}}{f_m}.$$

squaring both sides

$$1+2Q^2k^2+Q^4k^4 = 1+2Q^2(k^2-F_p{}^2)+Q^4(k^2-F_p{}^2)^2+4Q^2F_p{}^2$$

$$Q^4F_p{}^4+Q^2F_p{}^2[2-2Q^2k^2] = 0$$

$$QF_p = \pm\sqrt{2(Q^2k^2-1)}.$$

$$F_p = \pm\sqrt{2\left(k^2-\frac{1}{Q^2}\right)}.$$

We may note that the frequency of maximum response is (Section 7.3).

$$F_{max.} = \pm\sqrt{k^2-\frac{1}{Q^2}}$$

$$\therefore\ F_p = \sqrt{2}\,.\,F_{max.}$$

or $$F_p{}^2 = \left(\frac{2\Delta f_p}{f_m}\right)^2 = 2\left(k^2-\frac{1}{Q^2}\right)$$

and $$k = \sqrt{2\left(\frac{\Delta f_p}{f_m}\right)^2+\frac{1}{Q^2}}.$$

Replacing Δf_p by 8 kc/s, and f_m and Q by the values tabulated above, the following values of k are found.

f_m(kc/s)	600	1,000	1,400
k	0·0206	0·0151	0·0149
Qk	2·48	1·51	1·19

These values do not differ greatly from those given in the first table except for $f_m = 600$ kc/s.

4.7. Distortion due to the R.F. Valve Characteristic.[11, 12]

4.7.1. Modulation Envelope Distortion and its Measurement. The characteristics required of a valve for R.F. ampli-

fication differ from those for A.F. amplification owing to the different nature of the anode load impedance. Tuned circuits discriminate in favour of a relatively narrow band of frequencies and require a high valve slope resistance for satisfactory operation. The constant-current generator circuit (Fig. 2.12*b*) indicates this very clearly, for the valve resistance is in parallel with and so damps the tuned circuit. The impedance of the latter is usually high so that a large stage gain is possible. Difficulties due to feedback are therefore greatly increased and the anode-grid capacitance must be as low as possible if instability is to be prevented. Another difference between R.F. and A.F. amplification is that, owing to the selective properties of R.F. tuned circuits, second harmonic R.F. distortion due to valve curvature is permissible and has no distorting effect on the modulation envelope. The most suitable valve for the purpose is the tetrode or pentode valve, since either can have a high R_a and low anode-grid capacitance. The tetrode having screened grid I_aE_a characteristics (Fig. 2.5) is less suitable when large output voltages are required because secondary emission limits the minimum anode voltage.

To see how the distortion effects occur in an R.F. amplifier we will take a simplified mathematical treatment, in which it is assumed that R_a is infinite and E_g is the only variable influencing I_a. The anode load is assumed to be zero to all frequencies outside the required pass-band and resistive to frequencies inside. We need therefore only consider components of anode current at frequencies within the pass-band. If the I_aE_g characteristic is represented by

$$I_a = a_0 + a_1E_g \quad . \quad . \quad . \quad . \quad 4.41$$

and $$E_g = \hat{E} \cos \omega t(1 + M \cos pt) - E_b$$

where $\hat{E}$ = carrier peak voltage

M = modulation ratio

$-E_b$ = grid-bias voltage,

no distortion can occur, and the output voltage is an amplified replica of the input and the valve functions as a linear amplifier. Let us next take an I_aE_g characteristic of the form.

$$I_a = a_0 + a_1E_g + a_2E_g^2 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 4.42$$

$$= a_0 + a_1[\hat{E} \cos \omega t(1 + M \cos pt) - E_b] + a_2[\hat{E} \cos \omega t(1 + M \cos pt) - E_b]^2$$

$$= a_0 + a_1[\hat{E} \cos \omega t(1 + M \cos pt) - E_b] + a_2[\hat{E}^2 \cos^2 \omega t(1 + 2M \cos pt + M^2 \cos^2 pt) - 2E_b\hat{E} \cos \omega t(1 + M \cos pt) + E_b^2] \quad . \quad 4.43$$

By noting that $\cos^2\theta = \frac{1+\cos 2\theta}{2}$, we may separate the anode current into the following components:

D.C. $$a_0 - a_1E_b + a_2\left[\frac{\hat{E}^2}{2}\left(1+\frac{M^2}{2}\right)+E_b^2\right]$$

Modulated R.F. fundamental $(a_1\hat{E} - 2a_2\hat{E}E_b)\cos\omega t(1+M\cos pt)$.

Distorted modulated R.F. 2nd Harmonic

$$\frac{a_2\hat{E}^2}{2}\cos 2\omega t\left(1+\frac{M^2}{2}+2M\cos pt+\frac{M^2}{2}\cos 2pt\right).$$

A.F. modulation fundamental $a_2\hat{E}^2M\cos pt$

A.F. ,, 2nd Harmonic $\frac{a_2\hat{E}^2M^2}{4}\cos 2pt$.

Since the anode-load impedance is assumed to be zero to all frequencies outside $\frac{\omega+p}{2\pi}$ to $\frac{\omega-p}{2\pi}$ the output voltage from the valve is

$$I_aZ_0 = (a_1\hat{E} - 2a_2\hat{E}E_b)\cos\omega t(1+M\cos pt)Z_0.$$

The voltage is dependent on E_b, decreasing as E_b is increased, and we see that a parabolic I_aE_g characteristic may be used to produce variable gain with variable bias without distortion of the modulation envelope. This effect is entirely due to the selective properties of the external anode impedance. With such a characteristic, aperiodic anode circuits would allow A.F. modulation distortion to appear at the detector output. The distortion products occur from the second power term in the I_aE_g characteristic, and consist of the distorted modulated R.F. second harmonic, the A.F. modulation and its second harmonic. A point to note is that the D.C. anode current increases as the input signal is increased, and it is also affected by the modulation ratio. Small variations of D.C. anode current are therefore to be expected in a variable mu valve when the signal is modulated.

Unfortunately it is difficult to achieve a parabolic I_aE_g curve, and a variable-mu characteristic is more nearly represented by a power series having a very large number of terms. If we add another term $a_3E_g^3$ to 4.42 we have, in addition to the component frequencies listed above, other frequencies obtained as follows:

$$a_3E_g^3 = a_3[\hat{E}^3\cos^3\omega t(1+M\cos pt)^3 - 3E_b\hat{E}^2\cos^2\omega t(1+M\cos pt)^2 + 3E_b^2\hat{E}\cos\omega t(1+M\cos pt) - E_b^3]. \quad . \quad 4.44.$$

Replacing $\cos^3\theta$ by $\frac{3\cos\theta+\cos 3\theta}{4}$ the following components can be separated

D.C. $$-a_3E_b\left[\frac{3\hat{E}^2}{2}\left(1+\frac{M^2}{2}\right)+{E_b}^2\right]$$

Distorted modulated R.F. fundamental

$$a_3\hat{E}\cos\omega t\left[\frac{3\hat{E}^2}{4}\left(1+\frac{3M^2}{2}\right)+3{E_b}^2\right.$$
$$+\cos pt\left(\frac{9\hat{E}^2M}{4}+\frac{9\hat{E}^2M^3}{16}+3{E_b}^2M\right)+\cos 2pt\left(\frac{9M^2\hat{E}^2}{8}\right)$$
$$\left.+\cos 3pt\left(\frac{3\hat{E}^2M^3}{16}\right)\right].$$

Distorted modulated R.F. second and third harmonics, and A.F. modulation fundamental and second harmonic components are also present, but are not important because they do not contribute to the output voltage. The most important component is the distorted modulated R.F. fundamental, which is passed on through the amplifier to the detector and results in a distorted A.F. output. Power terms higher than the third all contribute to modulation envelope distortion.

It is therefore essential to reduce the factors, a_3, a_4, etc., in the I_aE_g power series to the smallest possible values, and rapid changes of curvature must be avoided. This is usually achieved by using a continuously variable pitch winding for the grid electrode as described in Section 2.4.

Distortion of the modulation envelope arising from these higher power terms limits the maximum modulated signal which can be accepted by an R.F. valve. The maximum signal generally increases as the negative bias voltage is increased, and the type of curve is shown in Fig. 7.19. A method of measuring directly the signal handling capacity of a R.F. valve is described in Section 7.11. The reason for the increase in maximum signal as the bias is increased is not very clearly shown by the power series, but it is actually due to a reduction in the rate of change of the g_m curve (see curve 2 in Fig. 2.8). Point A in Fig. 2.8 shows greatest rate of change of g_m and corresponds to the slight dip at $E_g = -12{\cdot}5$ volts in the input signal curve of Fig. 7.19.

A simpler method [32] of determining the signal handling capacity has been developed from the fact that the percentage second harmonic envelope distortion in a modulated R.F. output from the valve is directly related to the percentage third harmonic distortion produced by the same valve when an undistorted sinusoidal voltage is applied to its grid circuit. Expression 4.44 gives the amplitude of the second harmonic envelope distortion as

$$\frac{9a_3\hat{E}^3M^2}{8}$$

and by combining 4.43 and 4.44 the amplitude of the fundamental modulation envelope component is

$$[a_1 - 2a_2E_b + 3a_3E_b^2]\hat{E}M + \frac{9}{4}a_3\left(1+\frac{M^2}{4}\right)\hat{E}^3M.$$

Hence the second harmonic distortion percentage is

$$\frac{9a_3\hat{E}^2M \times 100}{8\left[[a_1 - 2a_2E_b + 3a_3E_b^2] + \frac{9}{4}a_3\left(1+\frac{M^2}{4}\right)\hat{E}^2\right]} \quad . \quad 4.45a.$$

By replacing $\frac{3a_3\hat{E}^2}{a_1 - 2a_2E_b + 3a_3E_b^2}$ by k, 4.45a becomes

$$\text{percentage second harmonic} = \frac{\frac{3}{8}kM100}{1+\frac{3}{4}k\left(1+\frac{M^2}{4}\right)} \quad . \quad 4.45b.$$

Applying a fundamental input voltage $\hat{E}\cos pt$ to

$$I_a = a_0 + a_1E_g + a_2E_g^2 + a_3E_g^3 \quad . \quad . \quad . \quad 4.46.$$

gives a fundamental amplitude of

$$(a_1 \quad 2a_2E_b + 3a_3E_b^2)\hat{E} \quad \tfrac{3}{4}a_3\hat{E}^3$$

and a third harmonic amplitude of $\frac{a_3\hat{E}^3}{4}$

$$\therefore \text{ percentage third harmonic} = \frac{\frac{a_3\hat{E}^2}{4}100}{a_1 - 2a_2E_b + 3a_3E_b^2 + \frac{3}{4}a_3\hat{E}^3}$$

$$= \frac{100k}{12\left(1+\frac{k}{4}\right)} \quad . \quad . \quad . \quad 4.47.$$

For any given percentage of second harmonic envelope distortion and percentage modulation, expression 4.45b gives a particular value of k, which inserted in 4.47, gives the corresponding value of third harmonic distortion. If therefore the valve is connected as an L.F. amplifier, and the input voltage adjusted to produce the calculated value of third harmonic percentage, the input voltage represents the signal handling capacity of the valve, i.e., it is equal to the carrier voltage which, modulated at the specified modulation percentage, gives the specified second harmonic distortion percentage of the modulation envelope. The values of k and percentage third harmonic distortion for different values of distortion and modulation percentage are tabulated below.

TABLE 4.2

Percentage 2nd Harmonic	30% M.		60% M.		80% M.	
	k.	Percentage 3rd.	k.	Percentage 3rd.	k.	Percentage 3rd.
2·5	0·2685	2·095	0·122	1·01	0·09	0·735
5	0·674	4·81	0·271	2·115	0·195	1·55
7·5	1·365	8·47	0·458	3·43	0·319	2·46
10	2·795	13·7	0·699	4·96	0·47	3·5

Since terms above the third power in the I_aE_g characteristic have been neglected the signal handling capacity determined by this method is not strictly accurate; as a general rule the error is quite small except at a grid-bias voltage where the characteristic has a rapid change of curvature or near cut-off of anode current.

The circuit recommended for this indirect measurement of signal handling capacity is shown in Fig. 4.14. The fundamental

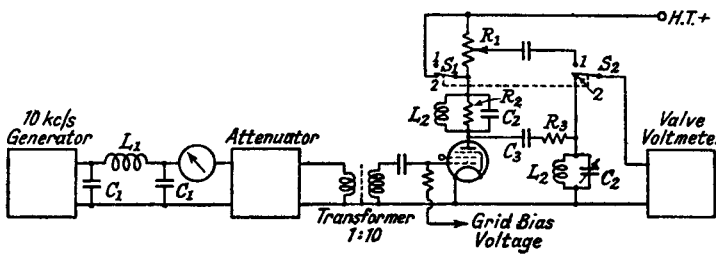


FIG. 4.14.—The Circuit Diagram for the Indirect Measurement of the Signal Handling Capacity of a R.F. Valve.

input voltage (frequency = 10 kc/s) is connected to the grid of the test valve through a low-pass filter ($L_1 = 3{\cdot}18$ mH, $C_1 = 0{\cdot}16$ μF), terminated by an attenuator of constant impedance 100 ohms and followed by a 1 : 10 step-up transformer. The anode circuit of the test valve contains a potentiometer R_1 (600Ω), across which the fundamental voltage is produced, and a frequency discriminating circuit L_2C_2 tuned to the third harmonic (30 kc/s). Ganged switches S_1 and S_2 enable the 30 kc/s output (position 2) to be compared with a proportion of the fundamental output (position 1), measured by a valve voltmeter. Potentiometer R_1 is adjusted so that the valve voltmeter indicates no change of reading between the two positions of the switches, and it is calibrated in terms of k by inserting a standard valve, the signal handling capacity of which

has either been calculated as described in Section 4.7.2 or measured as described in Section 7.11. The procedure for any valve is then to set R_1 to the required value of k, e.g., 0·271 for 60% modulation and 5% second harmonic envelope distortion, and to adjust the attenuator setting until the valve voltmeter indicates no change between the two switch positions. The signal handling capacity is then ten times the attenuator setting. An alternative method of calibrating R_1 in terms of k is to apply known amplitudes of 10 and 30 kc/s frequencies to the grid circuit of the valve in accordance with Table 4.2; thus for the condition set out above the amplitude ratio of 30 to 10 kc/s would be 2·115 to 100. Care must be taken to see that the amplitude of the 10 kc/s is small so that distortion of this frequency by the valve may be small.

The frequency discriminating network consists of two similar parallel tuned circuits of $L_2 = 10$ mH and $C_2 = 0{\cdot}01$ μF connected by a capacitor $C_3 = 0{\cdot}01$ μF in series with a high resistance R_3 (0·5 MΩ) to increase the selective properties of the second-tuned circuit. The first-tuned circuit is damped by a resistance R_2 (30,000Ω) to ensure that its dynamic resistance is much less than the anode impedance of any valve likely to be tested.

4.7.2. Calculation of Signal Handling Capacity from the g_mE_g Characteristic.[25] If 4.46 is differentiated with respect to E_g an expression for g_m in terms of E_g is obtained.

$$\frac{dI_a}{dE_g} = g_m = a_1 + 2a_2E_g + 3a_3E_g^2.$$

Now suppose $E_g = \hat{E}\cos\omega t - E_b$, then E_g varies between $\hat{E} - E_b$ and $-\hat{E} - E_b$. Denoting mutual conductance at these points by $g_m\,(max.)$ and $g_m\,(min.)$, and at $E_g = -E_b$ by g_{m0}, we have

$$\begin{aligned}
g_m\,(max.) &= a_1 + 2a_2(\hat{E} - E_b) + 3a_3(\hat{E} - E_b)^2 \\
g_m\,(min.) &= a_1 - 2a_2(\hat{E} + E_b) + 3a_3(\hat{E} + E_b)^2 \\
g_{m0} &= a_1 - 2a_2E_b + 3a_3E_b^2 \\
g_m\,(average) &= \frac{g_m\,(max.) + g_m\,(min.)}{2} \\
&= a_1 - 2a_2E_b + 3a_3E_b^2 + 3a_3\hat{E}^2 \\
&= g_{m0} + 3a_3\hat{E}^2 \\
&= g_{m0}\left[1 + \frac{3a_3\hat{E}^2}{a_1 - 2a_2E_b + 3a_3E_b^2}\right] \\
&= g_{m0}(1+k) \quad . \quad . \quad . \quad . \quad . \quad 4.48
\end{aligned}$$

where k has the same meaning as in Section 4.7.1.

Thus for 60% modulation and 5% second harmonic envelope distortion, $k = 0{\cdot}271$ and $g_m\,(aver.) = g_{m0}.1{\cdot}271$. Expression 4.48 may be used to calculate the signal handling capacity at any given bias $-E_b$, and the procedure is as follows. Using the g_mE_g curve draw a straight line through the point $g_m = g_{m0}(1+k)$, $E_g = -E_b$ to intersect the g_mE_g curve at two values, E_{g1} and E_{g2} such that $E_{g1} - E_b = E_b - E_{g2}$. The signal handling capacity is then $(E_{g1} - E_b)$ peak (unmodulated) carrier voltage or $0{\cdot}707\ (E_{g1} - E_b)$ R.M.S. carrier voltage.

4.7.3. Cross-Modulation Distortion. One of the most undesirable forms of distortion in R.F. amplifiers is that known as cross-modulation. This effect can occur when a modulated undesired signal is applied to the grid of a valve at the same time as the desired signal. If the I_aE_g power series has terms above E_g^2, the undesired modulation is transferred to the desired carrier, and discrimination against the undesired carrier after this R.F. valve produces no reduction of undesired modulation. It can be shown by taking the $a_3E_g^3$ term and replacing E_g by

$$(\hat{E}_d \cos \omega_d t + \hat{E}_u \cos \omega_u t(1 + M \cos pt) - E_b),$$

where the suffix d denotes desired and u undesired signal. For simplicity the former is not modulated. Expanding

$$a_3(\hat{E}_d \cos \omega_d t + \hat{E}_u \cos \omega_u t(1 + M \cos pt) - E_b)^3$$

shows a term of the form

$$3a_3\hat{E}_d \cos \omega_d t \hat{E}_u{}^2 \cos {}^2\omega_u t(1 + M \cos pt)^2,$$

which equals

$$3a_3\hat{E}_d \cos \omega_d t \hat{E}_u{}^2\left(\frac{1 + \cos 2\omega_u t}{2}\right)(1 + M \cos pt)^2$$

and this contains a component

$$\frac{3a_3}{2}\hat{E}_d\hat{E}_u{}^2 \cos \omega_d t\left(1 + \frac{M^2}{2} + 2M \cos pt + \frac{M^2}{2} \cos 2pt\right).$$

The desired carrier is now modulated by the undesired modulation and its second harmonic. The remedy for this cross-modulation effect is to increase the selectivity of the circuits preceding the first R.F. amplifier so as to reduce $\hat{E}_u$, and also to decrease all the factors a_3, a_4, etc., in the power series for the I_aE_g curve. Thus the methods applied to reduce distortion of the desired modulation envelope also reduce cross-modulation. It will be noticed that the cross-modulating term is dependent on the square of $\hat{E}_u$ and independent of its frequency. In a well-designed receiver the

cross-modulating effect due to a strong local station signal should be confined to a band of tuning frequencies in its immediate neighbourhood, for the tuned circuits preceding the R.F. valve should quickly reduce E_u to a small value as the receiver is tuned away from the local station position. Special aerial rejector circuits may be fitted when a receiver operates in close proximity to a powerful transmitter.

4.8. Instability in R.F. Amplifiers.[7] Instability is always possible in an amplifier when coupling exists between input and output circuits. Energy feedback may occur via output to input leads, through the common supply voltages, or via the anode-grid capacitance. The first two causes are under the designer's control, and careful placing of the leads and adequate decoupling with short lead capacitors can prevent regeneration. The total anode-grid capacitance can be reduced by eliminating stray coupling external to the valve, but it cannot be made less than the inter-electrode capacitance. The only possible method of preventing instability from this source is to limit the gain of the stage. From Section 2.8 we note that feedback through the anode-grid capacitance produces, in conjunction with the anode load, a grid input admittance of resistance and capacitance in parallel. The sign of the resistive term depends on the reactance of the anode load, being positive for a capacitance and negative for an inductance. If the anode load is a tuned circuit we have for the input grid resistance R_g (see Section 7.8, expression 7.26*a*)

$$R_g = \frac{(QF)^2 + \left(1 + \frac{R_D}{R_a}\right)^2}{g_m B_{g_a} QF R_D},$$

where $B_{g_a} = \omega C_{g_a}$ = anode-grid interelectrode susceptance, and the other terms have their usual meaning.

This is a minimum when $QF = \pm\left(1 + \frac{R_D}{R_a}\right)$ so that the minimum negative value of R_g is

$$R_g = \frac{-2\left(1 + \frac{R_D}{R_a}\right)}{g_m R_D B_{g_a}}$$

$$\simeq \frac{-2}{g_m R_D B_{g_a}}$$

when $$R_a \gg R_D.$$

Instability occurs when R_g is less than the parallel resistance component of the grid tuned circuit. If the anode and grid circuits are identical, the condition for stable operation is defined by

$$R_g > R_D$$

or

$$g_m R_D{}^2 B_{g_a} < 2$$

and for a given B_{g_a} either g_m or R_D must be limited. Generally it is better to reduce g_m by increasing grid bias, rather than to reduce R_D because selectivity is often reduced by decreasing R_D. However, the use of a tapped tuning coil allows the equivalent R_D to be reduced without loss of selectivity (see Section 4.4.1).

In estimating the condition for stability in a two-valve amplifier we will assume that all tuned circuits are identical. Owing to feedback, the dynamic impedance of the tuned circuit in the grid of the last valve is increased, and as this is the anode circuit of the next valve we have the following results:

Dynamic resistance of the last tuned circuit $= R_0' = R_D$

Grid input resistance of the last valve $= R_g' = \dfrac{-2}{g_m R_D B_{g_a}}$

Resultant dynamic resistance of the grid tuned circuit $= R_0''$

$$R_0'' = \frac{R_D R_g'}{R_D + R_g'} = \frac{2}{g_m B_{g_a}\left(\dfrac{2}{g_m R_D B_{g_a}} - R_D\right)}$$

Grid input resistance of the second valve $= R_g'' = \dfrac{-2}{g_m R_0'' B_{g_a}}$

$$= -\left[\frac{2}{g_m R_D B_{g_a}} - R_D\right].$$

The condition for stability is that

$$-R_g'' = \left[\frac{2}{g_m R_D B_{g_a}} - R_D\right] > R_D.$$

i.e.,

$$g_m R_D{}^2 B_{g_a} < 1.$$

For a three-valve amplifier the following results are obtained:
Resultant dynamic resistance of the third tuned circuit $= R_0'''$

$$R_0''' = \frac{R_g'' R_D}{R_D + R_g''} = \frac{R_D{}^2 - \dfrac{2}{g_m B_{g_a}}}{2R_D - \dfrac{2}{g_m R_D B_{g_a}}}.$$

$$\text{Grid input resistance of the valve} = R_g''' = -\frac{2}{g_m R_0''' . B_{g_a}}$$

$$= \frac{4\left[g_m R_D^2 - \frac{1}{B_{g_a}}\right]}{g_m R_D B_{g_a}\left(g_m R_D^2 - \frac{2}{B_{g_a}}\right)}.$$

For stability

$$R_D < -R_g'''$$

i.e., $$(g_m R_D^2 B_{g_a})^2 - 6 g_m R_D^2 B_{g_a} + 4 < 0$$

$$g_m R_D^2 B_{g_a} < \frac{+6 \pm \sqrt{20}}{2} < 0{\cdot}764.$$

Summarizing the results in the form of a table:

TABLE 4.3

No. of Amplifier Stages.	Maximum value of $g_m R_D^2 B_{g_a}$
1	2
2	1
3	0·764

In practice the maximum value of $g_m R_D^2 B_{g_a}$ will have to be much less than the value given in Table 4.3, otherwise feedback, though insufficient to cause oscillation, will have a serious effect on the overall frequency response of the amplifier. It causes greater amplification of frequencies below resonance and less amplification above resonance. The result is a lop-sided selectivity curve as discussed in Section 7.8.

4.9. Noise Limitation to Maximum Amplification.[35]

4.9.1. Introduction. By careful design a receiver may be constructed with very high amplification, and in the absence of a limiting factor it would be possible to obtain adequate output even from the weakest signal. The limiting factor is noise. Noise may be produced outside the receiver and be picked up in association with the desired signal or it may occur in the receiver itself. The desired signal must therefore be large enough to give with average modulation (about 30%) an A.F. output very much greater than that contributed by noise. Generally a signal-to-noise ratio of 15 db. is regarded as the minimum satisfactory level. External noise may be due to atmospheric disturbances (thunderstorm and magnetic storms), or to interference radiated from electrical

machinery. Over neither of these has the designer any control, though he can mitigate the effects of the latter, which is usually transmitted over the electric power lines, by including suitable R.F. filter chokes and by-pass capacitors between the supply and the receiver mains leads, and by locating the aerial as far as possible from the interference field of the supply lines (away from buildings) and connecting it to the receiver by screened leads (Section 3.5).

Noise produced in the receiver may be classed as accidental and inherent. The former is caused by faulty components—volume controls, variable capacitors, joints, can all contribute—and can be eliminated, but the latter cannot. Inherent noise is due to thermal effects in the conductors and shot noise in the valves. In both instances it is usually only the first stages of the receiver that have to be considered, since in later stages the signal is amplified and is much greater than any noise likely to be produced in these stages.

4.9.2. Thermal Noise. In Chapter 2 it was stated that all conductors contain free electrons, which are in a state of random motion. The average velocity of these free electrons is directly proportional to the absolute temperature and is only zero at 0° absolute. Each electron in motion constitutes a minute current and the sum total of these currents over a long period is zero. At any given instant, however, this will not necessarily be true, and there may be a net current in one direction or the other. These transient currents produce, across the ends of the conductor, voltages, the frequency components of which cover an infinite band. Nyquist assumes thermal noise to be equivalent to a voltage in series with the resistance of the conductor and computes its mean square value to be

$$E_{n_i}^2 = 4RkT(f_1 - f_2) \quad . \quad . \quad . \quad 4.49$$

where R = resistance of the conductor

k = Boltzmann's constant

T = Absolute temperature

$f_1 - f_2$ = Pass-band width of the receiver = $2\Delta f$.

The pass-band of the receiver must obviously affect the noise voltage since the wider this is the more noise frequency components are brought in. The pass-band is defined as the range of frequencies over which the response is greater than 70% of the maximum. Actually frequencies outside the normal pass-band of the stages preceding any non-linear device, such as a frequency changer or detector, have a noise-producing effect because the non-linear device can produce intermodulation frequencies in the overall pass-

band from noise frequencies normally outside it. Boltzmann's constant relates absolute temperature and electron energy, and the resistance R is the actual or equivalent shunt resistance of the first circuit. If the latter is a tuned circuit the resistance is the dynamic resistance (R_D) of the tuned circuit. The actual volume of noise produced at the output of the receiver—it appears generally as a hiss because of the greater sensitivity of the ear to the higher audio frequency components (500 to 5,000 c.p.s.)—is dependent on whether a carrier is being received. If there is no carrier the noise must provide its own, and as this is small in value the noise output is small. Section 8.2.8 states that detectors normally have a parabolic characteristic, and this makes the detection efficiency or sensitivity low for small signals. When a carrier is present detection sensitivity is increased and the noise voltages act as sidebands giving greater output. With no A.G.C. action in the receiver, increase of carrier up to a certain level increases the noise output, but beyond this level, at which the detector has reached the linear condition of maximum sensitivity, any further increase in carrier does not alter the noise output. If the receiver has A.G.C., as the carrier is increased beyond the point of operation of A.G.C., the receiver sensitivity is decreased and noise output reduced. The most important thermal noise voltage is that produced in the first tuned circuit, and if we assume normal operating temperature to be 63° F. (290° abs.) expression 4.49 becomes

$$E_{n_t} = 1{\cdot}25 \times 10^{-10}\sqrt{R_D(f_1 - f_2)} \quad . \quad . \quad 4.50.$$

Taking $R_D = 100{,}000\Omega$ and $(f_1 - f_2) = 10$ kc/s as typical of the medium wave band the noise R.M.S. voltage is

$$E = 1{\cdot}25 \times 10^{-10}\sqrt{10^5 \times 10^4} = 3{\cdot}95\ \mu\text{V}.$$

Hence a carrier (modulated 30%) of R.M.S. value 74 μV would be required to give the necessary 15 db. signal-to-noise ratio. The noise is assumed to have 1 sideband of the above value, whilst the carrier has two of 11·1 μV.

4.9.3. Shot Noise. A second important source of noise in a receiver is shot noise, produced by the flow of electrons from cathode to anode of the valve. The electrons making up the anode current have random motion, and the number arriving at the anode varies from one time-instant to another. This is equivalent to a small variable current of infinite number of frequency components superimposed upon the mean D.C. current. A mechanical analogy for shot noise is sand falling upon a gong; the resultant noise is dependent upon the resonant properties of the gong, the quantity

of sand falling per second and the size of the sand particles. In the same way shot noise is dependent on the external anode impedance (Z_0), the mean anode current I_a, and the electron charge e. For a saturated diode the mean square value of noise current has been shown to be

$$I_{n_s}^2 = 2I_a e(f_1 - f_2).$$

Replacing e by its actual value

$$I_{n_s} = 5{\cdot}54 \times 10^{-4}\sqrt{I_a(f_1 - f_2)}\ \mu\text{A}.$$

The noise voltage developed in the anode load Z_0 is

$$E_{n_s} = I_{n_s}Z_0 = 5{\cdot}54 \times 10^{-10} \times Z_0\sqrt{I_a(f_1 - f_2)} \quad . \quad 4.51.$$

This formula is not applicable to an amplifier valve operating under non-saturated space-charge conditions, for the space-charge reservoir of electrons acts as a cushion to smooth out the random variations, and the measured noise voltage is usually much less than this. Another theory assumes the valve to be replaced by its internal resistance at half the cathode temperature, and calculates the thermal noise voltage to be expected in the anode from this resistance. Calculations based on this assumption usually give too low a noise voltage. A convenient method of expressing shot noise is as the equivalent resistance, between grid and cathode of the valve, which would give at room temperature a thermal noise voltage in the anode circuit equal to that produced by the shot noise. This method has the advantage of allowing the relative magnitudes of shot and thermal noise to be compared by comparing the equivalent shot resistance with that of the input circuit. Since the thermal noise voltage across the anode circuit is proportional to $g_m Z_0\sqrt{R_D(f_1 - f_2)}$ and the shot noise voltage to $Z_0\sqrt{I_a(f_1 - f_2)}$, it follows that the equivalent shot noise resistance is proportional to $\frac{I_a}{g_m^2}$. Hence the best type of amplifier valve is one having a high value of g_m and low value of I_a. If the total space current exceeds the anode current, as in all multi-electrode valves except the triode, the equivalent shot noise is multiplied by a factor greater than 1 and is proportional to the ratio of total-to-anode current.

The following are average values of shot-noise resistance for different types of valves:

Valve Type.	Shot Noise Equivalent Resistance.
Triode	200 to 500Ω
Special beam tetrode	4,000 to 5,000Ω
Ordinary screened grid and pentode	20,000 to 50,000Ω
Frequency changer	50,000 to 100,000Ω

Frequency-changer valves are worst because g_c is never greater than $\frac{g_m}{4}$, i.e., signal gain is low for a given I_a. Ordinary screened grid or pentode valves are poor because of secondary emission and screen current. By concentrating the electrons into a beam and reducing screen current, shot noise can be much reduced, as shown by the beam tetrode value. The beam tetrode has the smallest shot noise of the multi-electrode valves because secondary emission is small and the space-charge reduces electron fluctuations.

If an R.F. amplifying stage is incorporated in a receiver, the impedance of the first-tuned circuit, except at ultra-high frequencies, is greater than the shot-noise resistance, so that thermal noise is the limitation. If, however, the first stage consists of a frequency changer the reverse is generally true and shot noise is greater than thermal. It should be noted that in a frequency changer stage shot noise is introduced by the oscillator valve as well ; shot and thermal noise components in the image or second channel region also add their quota. For a given overall gain a receiver without an R.F. valve before the frequency changer has normally at least twice the noise voltage of a receiver with an R.F. amplifier stage.

4.10. Problems in Short Wave and Ultra Short Wave Amplification.

4.10.1. Introduction. Amplification on the short wave range is chiefly complicated by the fact that selectivity is considerably reduced as the signal frequency is increased, and also that in many receivers the complete range from 6 to 15 Mc/s is covered by one coil using the same tuning capacitance, without modification, for short wave as for the medium and long wave ranges. As far as broadcast reception is concerned this method has two disadvantages : it makes tuning of the broadcast stations very sharp and difficult : it calls for a very small value of inductance which results in a low dynamic impedance and consequently low amplification. Much improved performance can be realized by band-spreading (selecting certain comparatively narrow bands in the short wave range). For communication receivers, required to cover completely the short wave band, the latter can be split up into a number of much smaller overlapping ranges. The tuning capacitance range is reduced by the use of a series or shunt capacitance (or combination of both), or in special cases much smaller tuning capacitances may be employed. The latter usually have the advantage of lower

losses, an important point at the higher frequencies, e.g., at 15 Mc/s.

The same problems also present themselves, though in a more acute form, at ultra high frequencies, and there are added complications due to low input valve conductance, stray inductance and capacitance.

4.10.2. Short Wave Amplification. In section 4.2.3 selectivity is shown to be a function of Q and the resonant frequency, and the pass-band is defined by

$$2\Delta f = \frac{f_r}{Q},$$

so that as f_r is increased the band-width is increased unless Q rises. The Q of short wave coils is often higher than corresponding coils for long and medium waves (values from 150 to 200 are common), but dielectric losses, valve input conductance, etc., lower considerably the effective Q of the tuned circuit, and a probable maximum value under favourable conditions is 50. This gives band-widths of 120 and 300 kc/s at 6 and 15 Mc/s respectively. The R.F. amplifier therefore only discriminates against undesired signal frequencies (reacting with harmonics of the oscillator—Section 5.4.3) well separated from the desired. Adjacent channel rejection is achieved in the I.F. amplifier. At an I.F. of 465 kc/s, the image signal (assuming $Q = 50$) is reduced by approximately 24 db. and 16·0 db. at 6 and 15 Mc/s respectively (see Fig. 4.3 for $QF = 15{\cdot}5$ and 6·2), and additional image rejection is really necessary. Circuits have been developed for application to band-spread receivers with preset signal tuning and these are described in Section 5.9.4. Little can be done to improve selectivity unless regeneration is employed, but this lack of selectivity may be used to advantage in band-spread broadcast receivers. Broadcast transmissions occur over certain quite narrow bands, rarely exceeding 200 kc/s width, and centred at 6·1, 9·6, 11·9, 15·2, 17·8 and 21·6 Mc/s. By reducing Q to 30 at 6·1 Mc/s, transmissions from 6 to 6·2 Mc/s can be accepted with a maximum loss of 3 db., i.e., the signal circuit can be preset tuned to 6·1 Mc/s and selection obtained by variation of oscillator frequency. At 15·2 Mc/s no decrease of Q is necessary. To obtain maximum amplification the signal-tuning capacitance should be as low as possible consistent with stray and valve capacitances, a minimum value being about 60 $\mu\mu$F. The required values of Q for a pass-band width of 200 kc/s, and the tuning inductance L for $C = 60$ $\mu\mu$F are tabulated below for the central frequencies listed.

Frequency.	Q	L
6·1 Mc/s	30·5	11·3 μH
9·6 ,,	48·0	4·57 ,,
11·9 ,,	59·5	2·97 ,,
15·2 ,,	76	1·82 ,,
17·8 ,,	89	1·32 ,,
21·6 ,,	108	0·9 ,,

The dynamic impedance is constant at 13,200Ω since $Q \propto f$. In actual fact it is unlikely that a Q value greater than 50 could be realized, and under these conditions the pass-band width increases as the frequency increases, whilst dynamic impedance and therefore amplification decrease as set out below.

Frequency.	R_D	Amplification Ratio. $\left(\frac{R_D \text{ at } f \text{ Mc/s}}{R_D \text{ at } 6{\cdot}1 \text{ Mc/s}}\right)$
11·9 Mc/s	11,100 Ω	0·841
15·2 ,,	8,700 ,,	0·66
17·8 ,,	7,380 ,,	0·588
21·6 ,,	6,120 ,,	0·464

Let us now consider the communication receiver required to cover with overlaps the short wave range from 6 to 25 Mc/s. Arbitrarily dividing into four ranges with overlaps gives

Range 1	6 to 9 Mc/s
,, 2	8·66 to 13 Mc/s
,, 3	12 to 18 Mc/s
,, 4	17 to 25·5 Mc/s

a frequency ratio change in each instance of 1 to 1·5, so that a capacitance change of 2·25 to 1 is required. We will assume that the frequency scale on the medium wave range is linear, and that we wish if possible to preserve this relationship while restricting the equivalent tuning capacitance change. If the signal-tuning inductance is 156 μH on the medium wave range and its self-capacitance is 10 $\mu\mu$F, the following tuning capacitance values are obtained at the equally spaced frequencies from 550 to 1,500 kc/s.

Frequency (kc/s) .	.	550	788·5	1,025	1,263·5	1,500
Capacitance ($\mu\mu$F)	.	526·8	252	145	92	62·17

and the problem is to obtain frequencies on the short wave ranges separated by a constant amount for these capacitance settings.

Restricting the capacitance range by a series padding capacitance gives the tuning frequency as

$$f = \frac{1}{2\pi\sqrt{L.\frac{C_p C}{C_p + C}}} \quad . \quad . \quad . \quad 4.52$$

where C_p = the series padding capacitance. By replacing f by 6 and 9 Mc/s and C by 526·8 and 62·17 $\mu\mu$F respectively in expression 4.52, the value of C_p is found to be 106 $\mu\mu$F. The frequencies corresponding to the chosen settings of C are

C ($\mu\mu$F)	526·8	252	145	92	62·17
Frequency (Mc/s)	6	6·52	7·2	8·02	9

whilst the frequencies for a linear frequency scale should be 6, 6·75, 7·5, 8·25 and 9 Mc/s. Hence the series padding capacitance has produced a reasonably satisfactory scale, which is slightly cramped at the high frequencies.

For a shunt trimmer capacitance the tuning frequency is

$$f = \frac{1}{2\pi\sqrt{L(C_t+C)}} \quad . \quad . \quad 4.53$$

and using the same tuning capacitance values C_t is calculated to be 310·5 $\mu\mu$F. The frequencies corresponding to the chosen settings of C are

C ($\mu\mu$F)	526·8	252	145	92	62·17
Frequency (Mc/s)	6	7·32	8·13	8·67	9

This means a very unsatisfactory scale with appreciable cramping at the low frequencies.

Since series and shunt capacitance restriction of tuning have opposite effects, it is possible to obtain a better approach to the linear frequency scale by a combination of both. Thus for C_p = 250 $\mu\mu$F and C_t = 46·5 $\mu\mu$F (C_p being between C and C_t), the frequencies are 6, 6·73, 7·51, 8·29 and 9 Mc/s, and the frequency scale is very nearly linear. From the point of view of maximum amplification over a tuning range, the series padding capacitance would be preferred to combined series and shunt capacitance because, although the frequency scale may be less satisfactory, the equivalent tuning capacitance is less and so the dynamic impedance is increased.

The values of C_p and C_t given above have the same effect on all ranges because the maximum-to-minimum frequency ratios are the same; thus for C_p = 250 and C_t = 46·5 $\mu\mu$F, the frequencies for range 4 are 17, 19·1, 21·3, 23·5 and 25·5 Mc/s.

4.10.3. Ultra Short Wave Amplification. In the previous section the difficulties of obtaining reasonable amplification in short wave amplifiers are outlined, and these are multiplied at ultra high frequencies. In spite of low stage amplification (about 5 times for general purpose receiving valves at the television frequency,

45 Mc/s and 12 for acorn valves [30]), there are definite advantages in including a R.F. stage between the aerial and frequency changer of a receiver. Overall amplification, signal-to-noise ratio, and selectivity against image and spurious I.F. responses due to interaction between undesired signals and the oscillator (Section 5.4) are all increased by the addition of a R.F. stage. The heavy damping from the R.F. valve (Section 2.8.3) prevents the realization of a high degree of selectivity against adjacent channels, but the I.F. amplifier discriminates satisfactorily against these. This fact, together with the probability of transmission being confined to comparatively narrow bands at certain selected positions in the ultra high frequency range, makes it possible to consider preset signal tuning to the centre of the band, tuning over it being accomplished by varying the oscillator frequency. Although the losses due to the valve input conductance, aerial connection and coil resistance predominate over all others, it is important to remember that components [27] such as the tuning capacitance, the trimmer, valve-holder, valve-base and any switches, which at lower frequencies generally have little effect, can contribute their quota. Expressing the losses as conductances, since they are circuits in parallel with the coil, typical values (at 45 Mc/s) for the components listed above are

Component	Bakelite Insulation.	Ceramic.
Tuning capacitance (minimum)	120 micromhos	30 micromhos
,, ,, (maximum)	20 ,,	5 ,,
Valve-holder	5 ,,	1 ,,
Valve-base	5 ,,	1 ,,
Range switch . . .	5 ,,	1 ,,

Let us consider the case of an amplifier stage operating at 45 Mc/s, and assume that the valve has characteristics identical with the one in Section 2.8.3, viz.,

$$g_k = 3 \text{ mA/volt}, \; C_{g_k} = 3 \; \mu\mu\text{F}, \; L_k = 0{\cdot}2 \; \mu\text{H}.$$

From expression 2.21c

$$R_g = -\frac{g_k{}^2+(B_k+B_{g_k})^2}{g_k B_{g_k} B_k}$$

$$B_{g_k} = 8{\cdot}48 \times 10^{-4} \text{ mhos}$$

$$B_k = -1{\cdot}768 \times 10^{-2} \text{ mhos}$$

$$R_g = \frac{(3 \times 10^{-3})^2+(-1{\cdot}683 \times 10^{-2})^2}{3 \times 10^{-3} \times 8{\cdot}48 \times 10^{-4} \times 1{\cdot}768 \times 10^{-2}}$$

$$= 6{,}480\Omega,$$

converting this to a conductance $G_g = 154$ μmhos. For maximum amplification the tuning capacitance should be as small as possible, and a probable minimum value is 30 $\mu\mu$F (valve and stray capacitance prevent it being less). For $C = 30$ $\mu\mu$F, $f = 45$ Mc/s, the tuning inductance is 0·416 μH, and taking Q as 150 for the coil, the dynamic resistance of the tuned circuit if C has no losses is

$$R_D = Q\omega L = 150 \times 6{\cdot}28 \times 45 \times 0{\cdot}416$$
$$= 17{,}700\Omega$$

or as a conductance

$$G_D = 56{\cdot}5 \ \mu\text{mhos}.$$

We will assume tuning capacitance loss to be 15 μmhos so that the overall total conductance including feedback loss due to the valve is made up as follows:

Part of the Circuit.	Conductance.
Coil	56·5 μmhos
Tuning capacitance	15·0 ,,
Valve-holder	2 ,,
Valve-base	2 ,,
Wiring	2·5 ,,
Range switch	2·0 ,,
Electron transit time in the valve	30·0 ,,
Total excluding feedback loss	110·0 ,,
Feedback loss due to L_k	154 ,,
Total conductance	264 ,,

To estimate the amplification from the aerial to the grid of the first R.F. valve we will assume that preset signal tuning is employed and that optimum coupling is therefore possible. Section 3.4.2 shows that the transfer voltage ratio is given by (expression 3.20c)

$$T_R = \frac{1}{2\omega C_2 \sqrt{R_2 R_{a1}}}$$

where C_2 is the tuning capacitance for the first-tuned circuit

R_{a1} is the total series resistance of the aerial circuit

R_2 is the total equivalent series resistance of the tuned circuit.

Taking R_{a1} as 80 ohms, the half-wave resonant impedance of a dipole aerial, the type most likely to be used at ultra high frequencies, and noting from expression 4.7 in section 4.2.2 that

$$R_2 = \frac{\omega^2 L^2}{R_D} = \omega^2 L^2 G_D$$
$$= (6{\cdot}28 \times 45 \times 0{\cdot}416)^2 \times 264 \times 10^{-6}$$
$$= 3{\cdot}65 \text{ ohms}$$

we find that

$$T_R = \frac{10^6}{2 \times 6{\cdot}28 \times 45 \times 30\sqrt{80 \times 3{\cdot}65}}$$
$$= 3{\cdot}45.$$

It will be seen from expression 3.20*c* quoted above, that $T_R \propto \frac{1}{\sqrt{R_2}}$ i.e., $\propto \frac{1}{\sqrt{G_D}}$, so that a decrease of G_D increases T_R. The overall Q of the tuned circuit in the absence of the aerial connection is $\frac{\omega L}{R_2} = \frac{117{\cdot}5}{3{\cdot}65} = 32{\cdot}2$, and for optimum coupling to the aerial, Q is halved, i.e., equals 16·1 (see expression 3.22*b* in Section 3.4.2). If we assume the band-width to be the frequency range over which the loss does not exceed 3 db., this condition is satisfied by

$$QF = \frac{Q2\Delta f}{f_r} = 1.$$

Hence the band-width $2\Delta f = \frac{45}{16{\cdot}1} = 2{\cdot}8$ Mc/s, and it is clear that preset signal tuning is a possibility when the band-width of the required transmissions is limited to the range 42·2 to 47·8 Mc/s. In Section 2.8.3 it is shown that a resistance R_k inserted in the cathode lead decreases the feedback conductance component of the valve. If $R_k = 150\Omega$ and $C_k = 5\ \mu\mu$F (this is stray capacitance, but it is essential in order to realize decreased input conductance).

$$Z_k = \frac{R_k}{1+\omega^2 C^2 R_k{}^2} + j\left(\omega L_k - \frac{\omega C_k R_k{}^2}{1+\omega^2 C_k{}^2 R_k{}^2}\right)$$
$$= 143{\cdot}3 + j(56{\cdot}5 - 30{\cdot}3)$$
$$= 143{\cdot}3 + j26{\cdot}2.$$

Thus $G_k = 6{\cdot}78 \times 10^{-3}$
$B_k = -1{\cdot}23 \times 10^{-3}$.

From expression 2.21*a*, Section 2.8.3

$$R_g = \frac{(G_k+g_k)^2+(B_k+B_{g_k})^2}{B_{g_k}(G_k B_{g_k} - g_k B_k)}$$
$$= \frac{(9{\cdot}78\times10^{-3})^2+(-0{\cdot}38\times10^{-3})^2}{8{\cdot}48\times10^{-4}(6{\cdot}78\times10^{-3}\times8{\cdot}48\times10^{-4}+3\times10^{-3}\times1{\cdot}23\times10^{-3})}$$
$$= 11{,}950\Omega$$
$G_g = 83{\cdot}6\ \mu$mhos.
Total $G_D = 193{\cdot}6\ \mu$mhos.

$$\frac{T_R(R_k = 150)}{T_R(R_k = 0)} = \sqrt{\frac{264}{193{\cdot}6}} = 1{\cdot}168.$$

This increase in transfer voltage ratio from the aerial to the grid of the first valve is, however, offset by the decrease in valve amplification due to negative feedback from the added cathode resistance. Section 2.7, expression 2.5*e*, indicates that the equivalent mutual conductance of the valve is reduced in the ratio $\frac{1}{1+g_k Z_k}$. Thus the valve equivalent g_m when $R_k = 0$ is

$$\frac{g_m}{1+g_k j\omega L_k} = \frac{g_m}{1+3 \times 10^{-3} \times 56{\cdot}5j} = \frac{g_m}{\sqrt{1+(0{\cdot}1695)^2}}$$
$$= 0{\cdot}972\ g_m$$

whereas for $R_k = 150\Omega$, $C_k = 5\ \mu\mu$F,

$$g_m' = \frac{g_m}{1+3 \times 10^{-3}(143{\cdot}3+j26{\cdot}2)}$$
$$= \frac{g_m}{\sqrt{(1+0{\cdot}4299)^2+(0{\cdot}0786)^2}}$$
$$= 0{\cdot}7\ g_m.$$

The actual reduction ratio of $g_m = \frac{0{\cdot}7}{0{\cdot}972} = 0{\cdot}72$ so that overall amplification, including the valve, has been reduced in the ratio 0·84 by inserting the resistance R_K. The advantages gained are:

(1) increase of selectivity in the ratio $\frac{264}{193{\cdot}6} = 1{\cdot}362$, and the band-width for a loss of 3 db. is reduced to 2·05 Mc/s; this is more important when variable tuned signal circuits are employed and may be a disadvantage with preset signal tuning.

(2) negative feedback tends to decrease distortion in the valve.

(3) there is less change of input conductance under A.G.C. operation, e.g., when $g_k = 0$, the valve input conductance = 106·2 μmhos ($R_g = 9{,}400\Omega$). It is practically independent of g_k in the particular example chosen.

Under certain conditions overall amplification can be increased by the use of a cathode resistance; for example, let us consider $R_k = 200$ ohms and $C_k = 15\ \mu\mu$F.

$$Z_k = \frac{200}{1{\cdot}715} + j\left(56{\cdot}5 - \frac{169{\cdot}0}{1{\cdot}715}\right)$$

$$= 116{\cdot}5 - j42.$$

$$G_k = 7{\cdot}58 \times 10^{-3},\ B_k = +2{\cdot}73 \times 10^{-3}$$

$$R_g = \frac{(10{\cdot}58 \times 10^{-3})^2 + (3{\cdot}58 \times 10^{-3})^2}{8{\cdot}48 \times 10^{-4}(7{\cdot}58 \times 10^{-3} \times 8{\cdot}48 \times 10^{-4} - 3 \times 10^{-3} \times 2{\cdot}73 \times 10^{-3})}$$

$$= -83{,}500\Omega$$

$$G_g = -11{\cdot}95\ \mu\text{mhos}.$$

Thus $G_D = 110 - 11{\cdot}95 = 98{\cdot}05\ \mu$mhos

and $$\frac{T_R(R_k = 200)}{T_R(R_k = 0)} = \sqrt{\frac{264}{98{\cdot}05}} = 1{\cdot}64.$$

The equivalent mutual conductance of the valve is

$$g_m' = \frac{g_m}{1 + 3 \times 10^{-3}(116{\cdot}5 - j42)} = \frac{g_m}{\sqrt{(1{\cdot}348)^2 + (0{\cdot}126)^2}}$$

$$= 0{\cdot}74\ g_m$$

and the ratio reduction of $g_m = \frac{0{\cdot}74}{0{\cdot}972} = 0{\cdot}76.$

The overall amplification ratio change is $1{\cdot}64 \times 0{\cdot}76 = 1{\cdot}245$. In this particular instance the overall amplification is increased by including R_k, but it will be noted that it is due to the cathode stray capacitance C_k overcompensating for the lead inductance L_k and so producing a negative input conductance. Selectivity is considerably increased, the ratio change being $\frac{264}{98{\cdot}05} = 2{\cdot}69$, and the bandwidth is reduced to 1·04 Mc/s from 2·8 Mc/s. Input conductance now varies considerably as g_k varies, from $-11{\cdot}95\ \mu$mhos at 3 mA/volt to 0 at 2·36 mA/volt and finally $+77{\cdot}5\ \mu$mhos at $g_k = 0$. This variation will help to improve the A.G.C. action and it has the merit of giving greatest selectivity for weak signals (lowest bias). An interesting feature of the particular component values chosen is the comparatively small change of conductance which occurs when the signal frequency is changed, e.g., at 40 Mc/s the conductance is $-15{\cdot}35\ \mu$mhos and it only decreases to $-5{\cdot}23\ \mu$mhos at 50 Mc/s. Care must be exercised to ensure that the negative input conductance is not increased sufficiently to cause an approach to oscillation.

Double-tuned transformers [31] may be employed for preset signal tuning between the R.F. valve and frequency changer, and they have the advantage of a flatter pass-band and sharper cut-off than a single circuit, though overall amplification is generally less. It is shown in Section 7.3 that maximum amplification with two tuned

circuits is only half that with a single-tuned circuit of identical L, C and Q values. However, a higher impedance may be possible with two tuned circuits since stray capacitance is approximately halved across each circuit, and the tuning inductance of either circuit may be increased upon that of a single circuit.

Adjustment of tuning at ultra high frequencies is preferably by variation of inductance, as fixed capacitors are less susceptible to ageing and temperature effects than variable ones, and variable inductances are more stable than variable capacitances. The tuning inductance usually consists of a few turns of copper wire with a metal plunger, which can be screwed into the axis of the coil. The plunger, which acts as a short-circuited turn to reduce inductance, must be of high conductivity material if it is not to alter appreciably the Q of the coil. The usual precautions appropriate to ultra high frequency operation must be taken in constructing the R.F. amplifier; leads must be as short as possible, all earth connections taken to the same point on the chassis, adequate decoupling, by small mica capacitors, of electrodes normally carrying only D.C. or A.C. supply voltages (screen, cathode and heaters).

BIBLIOGRAPHY

1. The Inductance Coefficients of Solenoids. Nagoka, *Journal of College of Science*, Tokyo, Aug. 15th, 1909, p. 1.
2. Bureau of Standards Circular No. 74 (1924).
3. The Effective Resistance of Inductance Coils at Radio Frequency. S. Butterworth, *Wireless Engineer*, April (p. 203), May (p. 309), July (p. 417), Aug. (p. 483), 1926.
4. Simple Inductance Formulas for Radio Coils. H. A. Wheeler, *Proc. I.R.E.*, Oct. 1928, p. 1398. Discussion on above (R. A. Batcher), *Proc. I.R.E.*, March 1929, p. 580.
5. An Improved Pre-Selector Circuit for Radio Receivers. E. A. Uehling, *Electronics*, Sept. 1930, p. 279.
6. The Design of Tuned Cicuits to fulfil Predetermined Conditions. A. L. M. Sowerby, *Wireless Engineer*, Jan. 1931, p. 23. Discussion (S. Butterworth), *Wireless Engineer*, April 1931, p. 199.
7. Oscillation in Tuned R.F. Amplifiers. B. J. Thompson, *Proc. I.R.E.*, March 1931, p. 421. Discussion, *Proc. I.R.E.*, July 1931, p. 1,281.
8. Theory and Operation of Tuned R.F. Coupling Systems. H. A. Wheeler and W. A. Macdonald, *Proc. I.R.E.*, May 1931, p. 738.
9. The Design of H.F. Transformers. M. Reed, *Wireless Engineer*, July 1931, p. 349.
10. The Design of the Band Pass Filter. N. R. Bligh, *Wireless Engineer*, Feb. 1932, p. 61.
11. Distortion in Screen-Grid Valves. R. O. Carter, *Wireless Engineer*, March 1932, p. 123.
12. The Theory of Distortion in Screen-Grid Valves. R. O. Carter, *Wireless Engineer*, Aug. 1932, p. 429.

13. Resistance in Band Pass Filters. G. H. Buffery, *Wireless Engineer*, Sept. 1932, p. 504.
14. Two Element Band Pass Filters. R. T. Beatty, *Wireless Engineer*, Oct. 1932, p. 546.
15. The Theory of Band Pass Filters for Radio Receivers. C. W. Oatley, *Wireless Engineer*, Nov. 1932, p. 608.
16. Ferro Inductors and Permeability Tuning. W. J. Polydoroff, *Proc. I.R.E.*, May 1933, p. 690.
17. Die Rückwirkung Metallischer Spulenkapseln auf Verluste Induktivität und Aussenfeld einer Spule. H. Kaden, *Elektrische Nachrichten Technik.*, July 1933, p. 277.
18. The Effective Resistance of Inductance Coils at Radio Frequency. B. B. Austin, *Wireless Engineer*, Jan. 1934, p. 12.
19. A Study of the Possibilities of R.F. Voltage Amplification with Screened Grid and with Triode Valves. F. M. Colebrook, *Journal I.E.E.*, Feb. 1934, p. 187.
20. The Effect of Screening Cans on the Inductance and Resistance of Coils. G. W. O. Howe, *Wireless Engineer*, March 1934, p. 115.
21. Band Pass Filter Characteristics. H. N. Jaderholm, *Electronics*, July 1936, p. 33.
22. Nomograms for the Design of Band Pass R.F. Circuits. C. P. Nachod, *Radio Engineering*, Dec. 1936, p. 13.
23. Two Mesh Tuned Coupled Circuit Filters. C. B. Aiken, *Proc. I.R.E.*, Feb. 1937, p. 230.
24. Universal Performance Curves for Tuned Transformers. J. E. Maynard, *Electronics*, Feb. 1937, p. 15.
25. A Graphical Estimation of the Signal Handling Capacity of Screened Grid Valves. R. W. Sloane, *Phil. Mag.*, April 1937, p. 529.
26. Using the R.F. Charts. C. P. Nachod, *Radio Engineering*, June 1937, p. 19.
27. The Resultant Q of Tuned Circuits. A. W. Barber, *Radio Engineering*, July 1937, p. 5.
28. The Design of Inductances for Frequencies Between 4 and 25 Mc/s. D. Pollack, *R.C.A. Review*, Oct. 1937, p. 184.
29. The Impedance of a Tapped Resonant Circuit. K. R. Sturley, *Marconi Review*, Oct.–Dec. 1938, p. 1.
30. A Receiver for Frequency Modulation. J. R. Day, *Electronics*, June 1939, p. 32.
31. Television Signal Frequency Circuits. G. Mountjoy, *R.C.A. Review*, Oct. 1939, p. 204.
32. The Signal Handling Capacity of H.F. Valves. R. W. Sloane, *Wireless Engineer*, Nov. 1939, p. 543.
33. The Effective Inductance and Resistance of Screened Coils. A. G. Bogle, *Journal I.E.E.*, Sept. 1940, p. 299.
34. *Radio Data Charts.* R. T. Beatty. Messrs. Iliffe.
35. *Spontaneous Fluctuations of Voltage.* E. B. Moullin, Oxford University Press. Text-book.
36. British Patent No. 518,969. J. D. Brailsford and Marconi's Wireless Telegraph Coy.

CHAPTER 5

FREQUENCY CHANGING

5.1. Problems in Frequency Changing.

5.1.1. Introduction.[1, 2, 3] The number of signal frequency amplifier stages in receivers covering a range of frequencies is generally limited, because ganging is difficult and, unless special precautions are taken, a considerable variation of sensitivity and selectivity occurs over the tuning range. For short wave reception it is almost impossible to obtain a high gain from an R.F. amplifier owing to the low value of tuning inductance (the dynamic resistance of a tuned circuit is $\frac{L}{CR}$) which must be employed, and at the high-frequency end of the range, cathode lead inductance and electron transit time combine to produce a high grid input admittance (Section 2.8). Many advantages are obtained if each signal frequency can be converted as desired to another fixed frequency and this is the principle involved in the superheterodyne method of reception. The frequency change is carried out by applying the signal and a local oscillator voltage (often known as the heterodyne * voltage) to a non-linear device, this term implying that frequencies, in addition to those applied at the input appear in the output voltage. Harmonics of, and the sum and difference frequencies of the signal and local oscillator, are produced in the output circuit, and any of these components may be selected by a suitable filter. The amplitudes of the sum ($f_h + f_s$) and difference ($f_h - f_s$) frequencies are equal, but the latter, called the intermediate frequency, is selected because possible amplification and selectivity are greater at the lower frequency. Any given signal can be converted to the intermediate frequency by a suitable choice of oscillator frequency.

In an ideal frequency changer the intermediate frequency amplitude varies directly as the signal frequency amplitude, and amplitude changes due to modulation of the signal frequency are reproduced without distortion with the intermediate frequency as carrier. It is, of course, assumed that no attenuation of the modula-

* All terms associated with the oscillator will be denoted by a suffix h, *e.g.*, $\hat{E}_h$ and f_h mean the local oscillator peak voltage and frequency respectively. The suffix h is used in preference to 0, as 0 is used to denote output circuit. The suffix s is used for the signal.

tion side-bands occurs in the intermediate frequency anode tuned circuit of the frequency changer.

5.1.2. The Advantages of Superheterodyne Reception. The chief advantages of superheterodyne reception are:

(1) The amplifier following the frequency changer can be designed for optimum performance because tuning controls are all preset.

(2) The overall sensitivity and selectivity of such a receiver is more constant over the tuning range because most is concentrated in the fixed frequency amplifier.

The intermediate frequency is generally, though not always, less than the signal frequency so that high sensitivity, selectivity and stability are obtainable. The gain of the frequency changer valve, itself, must be considered and it is therefore essential to define conversion conductance, g_c. This term, analogous to mutual conductance in an amplifier, is expressed as

$$g_c = \frac{\text{intermediate frequency component of } I_a \text{ for zero anode load}}{\text{signal voltage producing this component}}.$$

The output voltage from the frequency changer is

$$E_0 = g_c E_g Z_0$$

when Z_0, the external anode impedance, is much less than R_a, the frequency changer slope resistance.

Methods of measuring conversion conductance, and its relationship to mutual conductance, are discussed later. It should be noted, however, that g_c is generally less than 0·25 g_m. The lower value of g_c is offset to some extent by the higher anode impedance obtained at the intermediate frequency. This is made clear by Table 5.1, in which are listed typical values of coil constants used in radio receivers at intermediate frequencies of 110 and 465 kc/s, and at three signal frequencies of 200, 1,000, and 6,000 kc/s.

TABLE 5.1

Frequency (kc/s).	Inductance.	Q.	Impedance (Z_0).	Amplification.
110 (I.F.)*	9,000 μH	40	249,000	187
465 (I.F.)	1,000 μH	80	233,000	175
200 (S.F.)†	2,200 μH	45	124,000	372
1,000 (S.F.)	156 μH	100	98,000	293
6,000 (S.F.)	1·5 μH	200	11,300	34

* I.F.—intermediate frequency.
† S.F.—signal frequency.

For calculating amplification we have assumed that the mutual conductance of the amplifier is 3 mA/volt and that of the frequency changer 0·75 mA/volt. There is therefore a loss of amplification

by introducing frequency changing on the long and medium wave bands, but a considerable improvement on the short wave band—this is due to the low impedance of the short wave tuned circuit. The short wave inductance is low partly because it is assumed that the same tuning capacitor is used on all ranges. In receivers specially designed for short wave operation smaller capacitors (maximum value 150 $\mu\mu$F as compared with 500 $\mu\mu$F) are often employed. This allows an increase of inductance, which in turn increases the dynamic resistance of the signal tuned circuit and the amplification as an R.F. amplifier. The wave-range coverage is, of course, correspondingly reduced.

For medium and long wave operation the increased amplification obtained in the intermediate frequency amplifier tends to compensate for the lower amplification of the frequency changer.

5.1.3. The Principles of Frequency Changing.[16] Let us consider signal and oscillator voltages, represented by $\hat{E}_s \cos \omega_s t$ and $\hat{E}_h \cos \omega_h t$ respectively, applied to the grid circuit of a valve. If the valve has a linear $I_a E_g$ characteristic curve and the input voltages operate over the linear part, i.e., the valve is not biased near cut-off, we shall obtain the following result.

$$I_a = a_0 + a_1 E_g \quad . \quad . \quad . \quad . \quad . \quad 5.1a$$

but $$E_g = \hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b$$

where $-E_b$ is the bias voltage

thus $$I_a = a_0 + a_1(\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b) \quad . \quad 5.1b.$$

The output circuit therefore contains only the two frequencies present in the input circuit.

If the valve has a characteristic curve represented by

$$I_a = a_0 + a_1 E_g + a_2 E_g^2 \quad . \quad . \quad . \quad 5.2a$$

we have

$$\begin{aligned} I_a &= a_0 + a_1[\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b] \\ &\quad + a_2[\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b]^2 \\ &= a_0 + a_1[\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b] \\ &\quad + a_2[\hat{E}_s^2 \cos^2 \omega_s t + \hat{E}_h^2 \cos^2 \omega_h t + E_b^2 + 2\hat{E}_s \hat{E}_h \cos \omega_s t \cos \omega_h t \\ &\quad - 2\hat{E}_s E_b \cos \omega_s t - 2\hat{E}_h E_b \cos \omega_h t] \end{aligned}$$

and by noting that

$$\cos^2 \theta = \frac{1 + \cos 2\theta}{2}$$

and $$\cos \theta \cos \phi = \frac{\cos(\theta + \phi) + \cos(\theta - \phi)}{2}$$

we can separate the anode current into its seven components as follows :

D.C. component	$a_0 - a_1E_b + a_2\left[\frac{\hat{E}_s^2 + E_h^2}{2} + E_b^2\right]$
Fundamental Frequencies f_s	$(a_1\hat{E}_s - 2a_2\hat{E}_sE_b)\cos\omega_s t$
Fundamental Frequencies f_h	$(a_1\hat{E}_h - 2a_2\hat{E}_hE_b)\cos\omega_h t$
Harmonic Frequencies $2f_s$	$\frac{a_2\hat{E}_s^2}{2}\cos 2\omega_s t$
Harmonic Frequencies $2f_h$	$\frac{a_2\hat{E}_h^2}{2}\cos 2\omega_h t$
Sum frequency (f_h+f_s)	$a_2\hat{E}_s\hat{E}_h\cos(\omega_h+\omega_s)t$
Difference frequency (f_h-f_s)	$a_2\hat{E}_s\hat{E}_h\cos(\omega_h-\omega_s)t.$

The valve has produced four frequencies in addition to the signal and oscillator, and the insertion of a suitable filter in its anode circuit allows any one of these frequencies to be selected. Thus we may reject all except the difference frequency voltage component, which is then passed on to the intermediate frequency amplifier. The difference frequency $(f_h - f_s)$ naturally assumes that the oscillator frequency is greater than the signal. From a mathematical point of view the difference frequency term may be written as $a_2\hat{E}_s\hat{E}_h\cos(\omega_h - \omega_s)t$ or $a_2\hat{E}_s\hat{E}_h\cos(\omega_s - \omega_h)t$ since $\cos\theta = \cos(-\theta)$ and in practice we have $(\omega_h - \omega_s)$ if $f_h > f_s$ and $(\omega_s - \omega_h)$ if $f_s > f_h$. The frequency changer is quite unable to discriminate against the two signal frequencies which can produce with a given oscillator frequency a difference frequency equal to the intermediate frequency. Usually the lower frequency signal giving $(f_h - f_s')$ as the difference frequency is regarded as the desired signal—the reasons for this are set out in Section 5.1.5—and the higher giving $f_s'' - f_h$ is regarded as an undesired frequency, generally called the second channel or image frequency.[5]

If the signal frequency is modulated, e.g., $\hat{E}_s\cos\omega_s t(1+M\cos pt)$, all anode current components containing $\hat{E}_s$ are similarly modulated and the difference frequency component is

$$a_2\hat{E}_s\hat{E}_h(1+M\cos pt)\cos(\omega_h - \omega_s)t;$$

so that according to the definition in 5.1.2 conversion conductance

$$g_c = \frac{I_a}{\hat{E}_s} = a_2\hat{E}_h \quad . \quad . \quad . \quad . \quad 5.3a.$$

It may be pointed out that the sum frequency is similarly modulated but is never selected because it is higher in frequency than either signal or oscillator. Many of the advantages of super-

heterodyne reception are lost when the intermediate frequency is higher than the signal, though in certain instances, over the long wave band of an all-wave receiver, the difference frequency often does exceed the signal frequency.

In the above analysis it has been shown that a valve having a linear I_aE_g characteristic curve does not normally produce frequency changing. It should, however, be noted that such a valve, biased either to cut-off or to a point where the total grid voltage swing of oscillator and signal passes over the cut-off discontinuity, produces frequency changing. It does so because the condition of non-linearity is fulfilled and the valve characteristic may be represented by an infinite power series containing a term such as $a_2E_g^2$.

A valve having two control electrodes giving the I_aE_g relationship

$$I_a = (a_0+a_1E_{g1})(b_0+b_1E_{g3}) \quad . \quad . \quad . \quad 5.2b$$

will also act as a frequency changer so long as the signal voltage is applied to grid 1 and the oscillator to grid 3, or vice versa. The term $a_1b_1E_{g1}E_{g3}$ corresponds to $a_2E_g^2$ in expression 5.2*a* and conversion conductance is

$$g_c = \frac{a_1b_1}{2}E_h \quad . \quad . \quad . \quad . \quad 5.3b$$

i.e., $\frac{1}{2}a_1b_1$ corresponds to a_2. Any frequency changer such as a pentagrid, heptode, octode or hexode, in which the oscillator voltage is applied to an electrode other than that of the signal, operates on this principle.

5.1.4. Considerations Governing the Choice of the Intermediate Frequency.[4, 14] The choice of the intermediate frequency is first of all limited to a position in the frequency range where there is little chance of direct interference from transmitting stations. As a rule, frequencies in the long or medium wave range are avoided because interference may be produced due to direct signal amplification in the intermediate frequency amplifier. The intermediate frequency is most commonly located in the frequency range 450 to 475 kc/s, though a lower frequency range 100 to 125 kc/s has been used for superheterodyne receivers covering only the medium and long wave-bands, and a higher [7, 22] frequency range 1,500 to 2,000 kc/s is employed for "single span" [18] and double superheterodyne receivers. The intermediate frequency range 450 to 475 kc/s is preferred to the 100 to 125 kc/s range because it gives greater frequency separation between the desired and the image signal. An amplifier operating at the lower intermediate frequency is generally more stable and has greater amplification and selectivity

than one operating at 450 kc/s, but these advantages are far outweighed by the fact that locking or control of the oscillator frequency by the signal circuit tuning becomes serious on the short wave band when a low intermediate frequency is used. Furthermore, the higher intermediate frequency produces less interference whistles (see Section 5.4) over a given range of signal frequencies.

5.1.5. The Oscillator Frequency. In Section 5.1.3 it is shown that there are two signal frequencies which can react with a given oscillator frequency to produce the intermediate frequency and, conversely, there are two oscillator frequencies which can give the intermediate frequency with a certain signal frequency. The two oscillator frequencies are higher and lower than the signal frequency by an amount equal to the intermediate frequency. The higher oscillator frequency is almost invariably chosen because the ratio of its maximum to minimum value over a given range is less than that of the signal. This means that the ratio of the maximum to minimum values of oscillator tuning capacitor is less than that of the signal tuning capacitor. For example, a receiver having a range from 550 to 1,500 kc/s calls for a signal capacitor ratio change of $7{\cdot}43 : 1\left(C \propto \frac{1}{f^2}\right)$, whilst the oscillator capacitor ratio is only $3{\cdot}75 : 1$ for an I.F. of 465 kc/s (frequency range from 1,015 to 1,965 kc/s) when $f_h > f_s$. Constructional and ganging difficulties are reduced, because the same tuning capacitor unit may be used for the oscillator as for the signal circuit, its range being adjusted by series and parallel padding capacitances. (See Section 6.12.)

5.1.6. Interference Whistles. Frequency changing possesses the disadvantage that it increases the possibility of obtaining interference from undesired signals. The interference may make itself evident by superimposing an undesired programme and a whistle on the desired, but generally it is shown by a whistling note only, the frequency of which varies when slight changes of oscillator capacitance tuning are made. The most serious is that due to image interference. Special measures to increase the filtering properties of the signal-tuned circuits at this particular frequency are often undertaken and are described in Section 5.9. There are many other forms of interference and a whistle is always possible if

$$\pm mf_h \mp nf_u \simeq f_1$$

where m and n are positive whole numbers

f_u = undesired signal frequency
f_1 = intermediate frequency.

Let us take as an example a receiver tuned to a signal of 750 kc/s with an I.F. of 465 kc/s. The oscillator frequency, 1,215 kc/s, may react with an undesired signal frequency of 982 kc/s at the input of the frequency changer to produce a 1 kc/s whistle in the I.F. amplifier for

$$2 \times 1{,}215 - 2 \times 982 = 466 \text{ kc/s}.$$

Similarly an undesired signal of 1,967 kc/s can produce a 2 kc/s whistle because

$$2 \times 1{,}215 - 1{,}967 = 463 \text{ kc/s}.$$

A more detailed discussion of these interference frequencies is given later in Section 5.4.

5.2. Frequency Changer Circuits.

5.2.1. Introduction. Frequency changer circuits may conveniently be divided into two classes; in the first the signal and oscillator frequencies are applied to a common electrode, usually the grid-cathode circuit of a valve, and in the second the two frequencies are applied at different electrodes. There is no fundamental difference between the two types of circuit, and frequency changing results in both cases because the oscillator voltage controls the mutual conductance and hence the amplification of the valve. Common electrode coupling entails a curved or discontinuous I_aE_g curve because frequency changing is determined by the E_g^2 term shown in equation 5.2*a*. For separate electrode coupling, the oscillator voltage causes a variation in the slope of the $I_aE_{sig.\,grid}$ characteristic curve, which may itself be a straight line, and frequency changing occurs in accordance with expression 5.2*b*.

5.2.2. Oscillator Application to the Grid-Cathode Circuit. Let us imagine that the signal and oscillator voltages are applied in series to the grid circuit of an R.F. amplifier valve having a curved I_aE_g characteristic, and that there is no interaction between the voltage sources.

If we wish to find the operating condition for maximum conversion conductance we have two dependent variables to consider, viz., the oscillator and bias voltages. The signal voltage is not a dependent variable since conversion conductance, in the same manner as mutual conductance, is independent of the signal voltage under normal operating conditions. The bias voltage must always be such that grid current does not flow, because damping of the signal-tuned circuit is not permissible. It is preferable therefore to vary the bias voltage with the oscillator voltage, keeping the bias

voltage always at least 1 volt more negative than the peak value of the oscillator voltage. Typical conversion conductance-oscillator voltage curves, obtained as described in Sections 5.6.3 and 5.6.4 are shown in Fig. 5.1*a*. The negative bias values for the three curves are adjusted to be 1, 2 and 3 volts, respectively, greater than the peak value of the oscillator voltage. The maximum value of conversion conductance and the oscillator voltage required to reach it tend to fall as the bias is increased. For small values of oscillator voltage the conversion conductance actually increases with increase

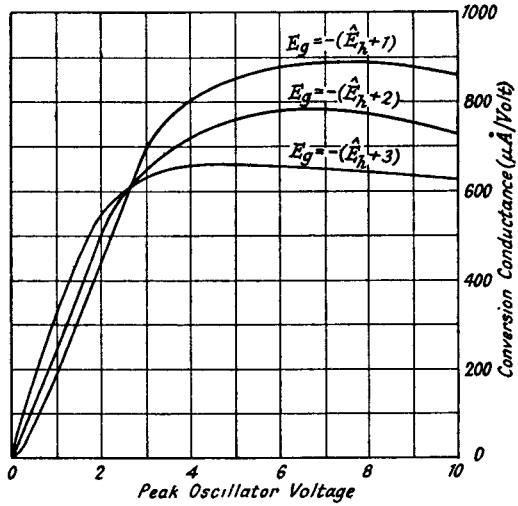


FIG. 5.1*a*.—Typical Conversion Conductance-Oscillator Voltage Curves for Signal and Oscillator Voltages in the Grid-Cathode Circuit.

of bias. If the latter is increased far enough the conversion conductance begins to fall and the effect is analogous to the optimum bias point for an anode bend detector as described in 8.4.1.

It is necessary now to consider the reason for maximum conversion conductance at a particular oscillator voltage, since the I.F. term $a_2\hat{E}_s\hat{E}_h \cos(\omega_h - \omega_s)t$ obtained from expression 5.2*a* would not suggest an optimum but an ever-increasing value as the oscillator is increased. This is due to the fact that a characteristic I_aE_g curve is never completely represented by the simple power series [16] in 5.2*a*, but more nearly by

$$I_a = a_0 + a_1E_g + a_2E_g^2 + a_3E_g^3 - a_4E_g^4 + \quad . \quad . \quad 5.4.$$

A negative sign for the a_4 term agrees with the known fact that

as E_g is increased positively I_a does not increase indefinitely but reaches some saturation value. Replacing E_g by

$$\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t - E_b$$

in $-a_4 E_g^4$ gives a term

$$\begin{aligned} & -4a_4 \hat{E}_s \hat{E}_h^3 \cos \omega_s t \cos^3 \omega_h t \\ = & -4a_4 \hat{E}_s \hat{E}_h^3 \cos \omega_s t (\tfrac{3}{4} \cos \omega_h t + \tfrac{1}{4} \cos 3\omega_h t) \\ = & -\tfrac{3}{2} a_4 \hat{E}_s \hat{E}_h^3 [\cos (\omega_h - \omega_s)t + \cos (\omega_h + \omega_s)t] \\ & -\frac{a_4}{2} \hat{E}_s \hat{E}_h^3 [\cos (3\omega_h - \omega_s)t + \cos (3\omega_h + \omega_s)t]. \end{aligned}$$

The first term, $-\tfrac{3}{2} a_4 \hat{E}_s \hat{E}_h^3 \cos (\omega_h - \omega_s)t$, increases rapidly as the oscillator voltage is increased and subtracts from the I.F. component produced by $a_2 E_g^2$. There is also a term $-\tfrac{3}{2} a_4 \hat{E}_s^3 \hat{E}_h \cos (\omega_h - \omega_s)t$, but its effect is much less marked because $\hat{E}_s$ is usually much less than $\hat{E}_h$. Hence an optimum oscillator voltage value is obtained as shown by Fig. 5.1*a*, which gives curves typical of a valve of the variable mu type. For a non-variable mu valve the optimum oscillator voltage has a much more pronounced maximum. It will be noticed that the curves, after the initial rapid rise, reach a maximum rather slowly and it is preferable to operate at an oscillator voltage lower than maximum, e.g., 4 volts for $E_g = -(\hat{E}_h + 1)$ because the amplitude of the interference whistles generated by the frequency changer increases rapidly with increase of oscillator voltage (see Section 5.4.1).

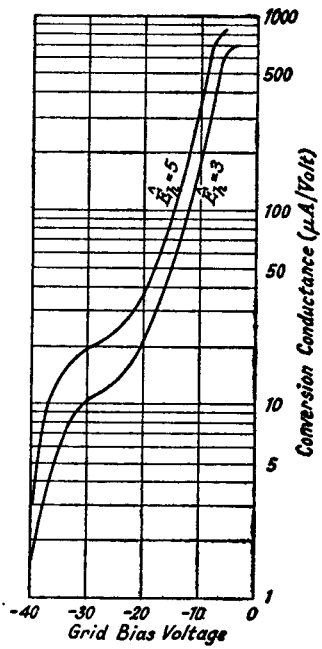


FIG. 5.1*b*.—Conversion Conductance-Grid Voltage Curves for Oscillator Voltage Application to the Cathode Circuit.

For good automatic gain control characteristics it is essential to control as many R.F. stages as possible so that the frequency changer valve may possess variable mu characteristics. Typical $g_c E_g$ curves are illustrated in Fig. 5.1*b* for two values of oscillator voltage. The larger gives greater conversion conductance and a longer cut-off (approximately longer by the difference between the two oscillator peak voltage values, i.e., 2 volts).

The oscillator voltage may be applied to the grid-cathode circuit

in a number of ways. One possible method is via a coil connected between the grid and the signal-tuned circuit. This has the disadvantage that the tuning range of the signal circuit is restricted by the coil capacitance to earth, and ganging with the preceding signal amplifier stages is rendered almost impossible. The alternative position for the coil, on the earthed side of the tuning circuit, necessitates insulation of rotor as well as stator plates of the tuning capacitance and is therefore impracticable. It can however be used when signal tuning is preset as for push-button operated receivers or band-spread receivers at high and ultra-high frequencies.

A second method is to apply the oscillator voltage via a coil coupled to the signal tuning coil. This suffers from the disadvantage that the actual value of the oscillator voltage applied to the grid circuit now depends on the selectivity of the signal circuit, and the frequency separation between oscillator and signal. The signal-tuned circuit is directly connected to the pick-up coil and interaction tends to cause a variation of oscillator frequency with variation of the signal-tuned circuit.

Coupling by capacitance to the grid of the frequency changer is a possibility, but it possesses the disadvantages of both the first two methods.

The best method of inserting the oscillator voltage in the grid-cathode circuit is by means of a coil in the cathode-earth lead. The coupling between the signal and oscillator circuits is limited to the grid-cathode capacitance and interaction is very much reduced. This circuit is very suitable for signal frequencies in the medium and long wave band, but is not satisfactory for short wave operation because the grid-cathode capacitance is sufficient to cause the signal circuit tuning to influence the amplitude and frequency of the oscillator voltage. The effect of this coupling at medium and long wave frequencies is chiefly to develop across the signal circuit an oscillator voltage in opposition to the cathode oscillator voltage, so that the net voltage applied between the grid and cathode of the frequency changer is decreased. A high I.F. reduces the feedback voltage developed in the signal circuit by increasing the frequency separation between signal and oscillator circuits. For an I.F. of 465 kc/s, the oscillator voltage developed in a signal circuit tuned to the centre of the medium wave-band is of the order of $\frac{1}{30}$ of that in the cathode circuit.

It is quite usual to combine the oscillator and frequency changer in one valve, and a typical circuit for grid-cathode coupling is shown in Fig. 5.2*a*. The valve is a triode-pentode with a cathode

common to both. The pick-up coil in the cathode is also the feedback coil for the oscillator, and this necessitates decoupling of the anode and screen circuits of the pentode to the cathode and not to earth. If these two circuits are decoupled to earth, current varia-

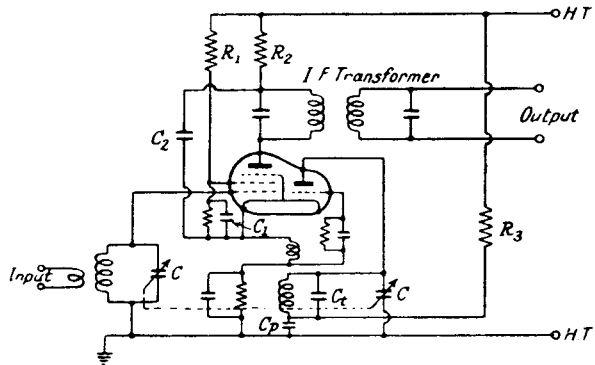


FIG. 5.2*a*.—The Circuit for Oscillator Voltage in the Cathode of a Triode-Pentode Valve.

tions in the pentode section at the oscillator frequency develop a voltage across the coil in opposition to that generated by the oscillator. This is more clearly demonstrated by considering the equivalent circuit in Fig. 5.2*b*. The I.F. circuit impedance is low to the oscillator frequency and is assumed to be zero. For convenience the cathode coil is replaced by a separate generator of

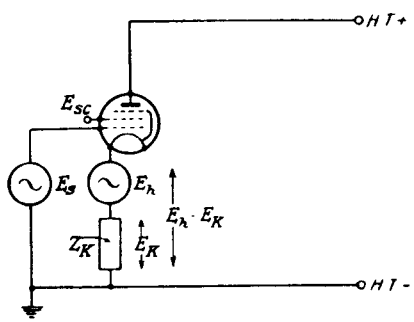


FIG. 5.2*b*.—The Equivalent Circuit for the Oscillator Voltage in the Cathode Circuit of a Pentode Valve.

open circuit voltage E_h (the oscillator voltage) with an internal impedance Z_k, the coil impedance at the oscillator frequency. The voltage applied to the grid-cathode electrodes of the pentode is $E_h - E_k$, where E_k is the voltage produced across the coil imped-

ance Z_k by the oscillator frequency variations of the pentode current. The value of E_k is

$$E_k = g_m(E_h - E_k)Z_k$$

where g_m = mutual conductance of the pentode.

Thus

$$E_k = E_h \frac{g_m Z_k}{1 + g_m Z_k}$$

and the net voltage between cathode and earth is

$$E_h - E_k = \frac{E_h}{1 + g_m Z_k} \quad . \quad . \quad . \quad 5.5.$$

The effective voltage is therefore reduced, and since this is also the driving voltage for the oscillator it means that there is a serious degenerative effect on the latter. This defect, which may even be sufficient to prevent oscillation altogether, can be eliminated by decoupling the screen and anode circuits to cathode. The resistances R_1 and R_2 and capacitances C_1 and C_2 in Fig. 5.2*a* perform this function, diverting pentode current variations at the oscillator frequency direct to cathode instead of via earth through Z_k to the cathode. I.F. current variations are diverted from Z_k at the same time. Capacitors C are ganged together and C_p acts as a padding capacitance for ganging purposes.

5.2.3. Oscillator Application to the Screen Circuit. Frequency changing may be accomplished by applying the oscillator voltage in the screen circuit of a multigrid valve. Owing to the low

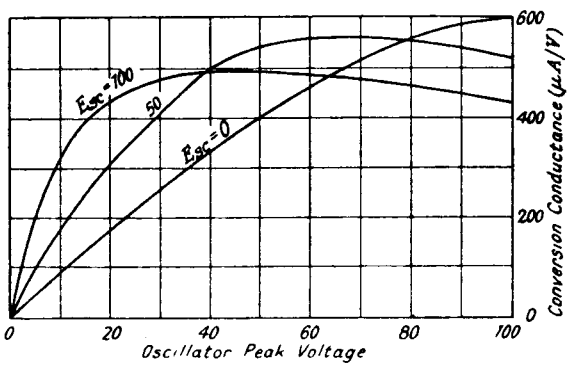


FIG. 5.3.—Typical Conversion Conductance-Oscillator Voltage Curves for Oscillator Voltage in the Screen Circuit of a Pentode Valve.

amplification factor between screen and anode a much larger voltage is required than for the cathode circuit. Typical conversion conductance-oscillator voltage curves are shown in Fig. 5.3 for a

pentode valve. Similar curves may be obtained for a screened-grid valve, but they are of little practical value since conversion gain under normal operating conditions is limited by the secondary emission characteristic, and falls off rapidly when the sum of the D.C. screen voltage and oscillator peak voltage approaches the minimum voltage developed in the anode under normal operating conditions. This voltage, which is equal to the H.T. voltage minus the peak value of the output voltage E_0, is dependent on the I F. external anode impedance and the signal voltage. In a pentode the effect does not occur since the minimum anode voltage is practically independent of screen voltage and can be very nearly zero.

An interesting point shown by the curves is that maximum g_c is obtainable with zero D.C. screen volts. The valve then operates as a half-wave frequency changer on the positive half of the oscillator voltage wave, the anode current being cut off for the negative half-cycle. Section 5.5 indicates that this is the most efficient method of frequency changing.

A serious disadvantage of screen application is the large oscillator voltages required. The screen-control grid capacitance is almost equal to that of the control grid-cathode so that the proportion of oscillator voltage transferred to the signal circuit is the same. Since the required value is so large (ten to twenty times greater than for cathode application) the voltage appearing across the signal circuit is much greater. Also the oscillator-tuned circuit has to be used to provide the screen voltage, and the signal-circuit tuning consequently influences the oscillator frequency to a great extent. For these reasons screen frequency changing is never used.

5.2.4. Oscillator Application to the Suppressor Grid. We have already shown in Section 2.5 that bias on the suppressor grid of a pentode may be used for controlling the amplification of the valve. Frequency changing can therefore be accomplished by applying the oscillator voltage to the suppressor grid circuit as in Fig. 5.4, and it takes place in accordance with expression 5.2*b*. The current taken by the suppressor grid on application of the oscillator provides the self-biasing voltage across R_s (about 1 MΩ). The optimum oscillator voltage depends on the suppressor grid cut-off bias voltage, and its peak value is usually 2 to 3 volts greater than half the latter. A typical curve of conversion conductance is shown in Fig. 5.4 for the self-bias condition. The maximum conversion gain $\left(g_c \cdot \frac{Z_0}{R_a+Z_0}\right)$ is generally little more than half that for cathode application because of the reduction in valve slope

resistance brought about by biasing the suppressor grid (Section 2.5). This resistance reduction causes loss of selectivity as well as of amplification when the valve is associated with an I.F. tuned transformer.

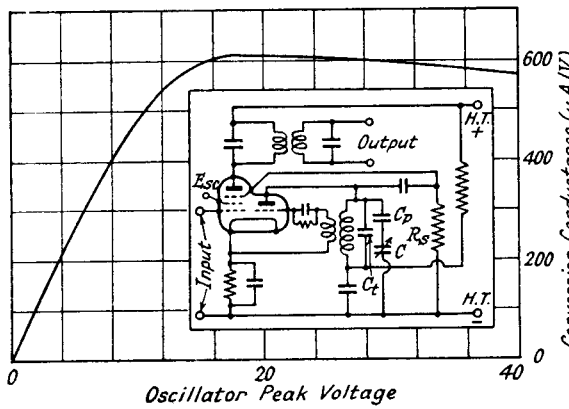


FIG. 5.4.—Typical Conversion Conductance-Oscillator Voltage Curves for Suppressor Grid Oscillator Voltage and Self-Bias on a Pentode Valve.

The chief advantage of suppressor grid application is the additional screening provided between signal and oscillator circuits by the screen-grid electrode and it therefore can be used for short wave operation.

5.2.5. Oscillator Application to the Anode Circuit.[9] Oscillator application to the anode circuit of a tetrode or pentode valve is not employed because the anode voltage has little effect on anode current unless the former is very low. For large conversion conductance the D.C. anode voltage must be low and the oscillator voltage must carry the anode voltage over the bend of the I_aE_a curve. Under these conditions the valve slope resistance is low and the I.F. transformer characteristics are adversely affected.

5.2.6. Frequency Changing and Oscillation from a Single Valve.[8] It is possible to combine the functions of frequency changer and local oscillator in one valve, though it is preferable to employ a separate oscillator as it is not easy to obtain high efficiency frequency changing with satisfactory oscillation and little interaction over a range of signal frequencies. Oscillation may be obtained by feedback coupling between cathode and anode or screen and anode, or by using the dynatron, or negative resistance characteristic (Section 2.4) of the screen or anode of a screened-grid valve. In one method a coil, inserted in the cathode lead of a tetrode or

pentode, is coupled to the oscillator-tuned circuit connected by capacitance to the anode, which contains an I.F. transformer with untuned primary, the latter acting as a choke to the local oscillator frequency. An alternative is to connect the feedback coil in the screen circuit, but this requires much tighter coupling as the amplification between screen and anode is generally low, and screen application requires a large oscillator voltage for satisfactory frequency changing. The position of the oscillator tuned circuit and feedback coil can be reversed, the latter being placed in series with the I.F. transformer primary, which can then be tuned. Oscillations may be generated by inserting the oscillator tuned circuit in either screen or anode of a screened grid valve and suitably biasing both electrodes to bring the dynatron characteristic into play, but this method is not very satisfactory and adjustments are critical. Combined single valve operation of this kind is only possible over medium and long wave ranges since there is considerable interaction between signal and oscillator circuits.

5.3. Special Types of Frequency Changers.[42]

5.3.1. The Triode Hexode.[10, 32] The triode-hexode frequency changer is a logical development from the triode pentode with suppressor grid application. It will therefore be treated before the heptode or octode, although it was preceded by them. The low slope resistance due to biasing the suppressor grid of the pentode is restored to a high value by the inclusion of a positively biased

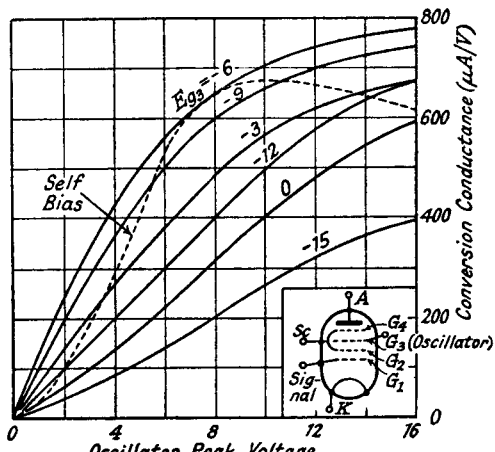


FIG. 5.5.—Typical Conversion Conductance Curves for a Hexode Valve.
(E_a = 250, E_{sc} = 70, E_{g1} = − 1·5.)

screen between the suppressor grid and anode. The effect of this screen, which converts the pentode to a hexode, is shown in Section 2.5 to be the conversion of the triode, formed by the virtual cathode round the first screen, the suppressor grid and anode, into a screen grid valve.

Typical conversion conductance-oscillator voltage curves are shown in Fig. 5.5. The dotted curve is the one obtained for self-bias of the oscillator grid. The rule for optimum oscillator peak voltage as given for self-biased suppressor grid application is again applicable, viz., that it is about half the cut-off bias voltage (— 16) plus 2 or 3. The additional screen (g_4) not only raises the valve slope resistance but also allows g_3 to have a closer mesh without seriously reducing anode current, and so the optimum oscillator voltage is much lower than for suppressor grid application. The signal grid g_1 is generally given variable mu characteristics, and variation of bias on this grid

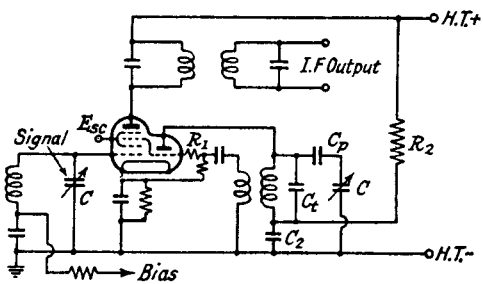


FIG. 5.6.—The Circuit for a Triode-Hexode Frequency Changer.

produces a curve similar in shape to those shown in Fig. 5.1*b*. A representative circuit for a triode-hexode frequency changer is shown in Fig. 5.6.

To avoid excessive interference whistle production it is advisable to operate the hexode at an oscillator voltage not exceeding the optimum and preferably less. Advantage should not be taken of the flat part of the self-biased conversion conductance curve to maintain constant g_c (in spite of large oscillator output variations over a tuning range), but rather attempts should be made to control oscillator output. The series grid resistance R_1 in Fig. 5.6 assists in preserving more constant voltage. (See Section 6.4.)

The hexode possesses the chief advantage of suppressor grid application, viz., low capacitance coupling between signal and oscillator circuits, and in addition gives high conversion conductance with high anode impedance. It is therefore particularly suitable for all-wave operation.

5.3.2. The Heptode.[13, 19, 29] The heptode (originally called pentagrid [5]) valve, which was an earlier development than the hexode, is a combined oscillator and frequency changer. It contains five grids, four are of normal construction and the fifth, so-called grid (G_2 in the circuit diagram of Fig. 5.7), consists of two rods,

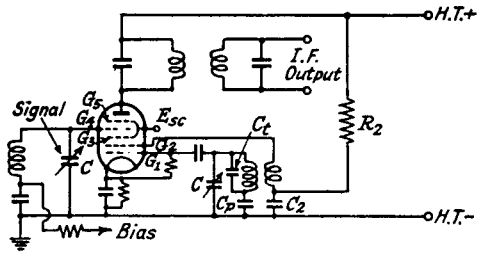


FIG. 5.7.—The Circuit Diagram for a Heptode Frequency Changer.

similar to grid support wires, placed between grids 1 and 3 These two rods constitute the anode of the oscillator, the grid of which is G_1 nearest the cathode. The grid G_1 is the active grid for frequency changing. Grids 3 and 5 are positively biased screening grids and grid 4 is the control grid. The valve, except for the oscillator anode, is similar to the hexode with the positions of the signal and oscillator voltages reversed. The rods forming the oscillator anode are only large enough just to maintain oscillation, and are set in line with the other grid support wires so that they are outside the main electron stream. Current for the oscillator anode, grid 2, is mainly derived by secondary emission from the screen, grid 3. It is essential that the oscillator anode should influence the main electron stream as little as possible, because changes of voltage on this electrode are in antiphase to those on grid 1 and tend to reduce the effective oscillator voltage applied for frequency changing.

The oscillator is of the tuned-grid type since the comparatively large voltage required on grid 1 can be obtained more easily from a tuned-grid than from a tuned-anode oscillator. A disadvantage of the valve is that variation of signal-grid bias (G_4), for amplification control purposes, affects the secondary emission from G_3 to the oscillator anode and causes amplitude and frequency variations of the oscillator. This can be overcome by the use of a separate oscillator section,[23] or greatly reduced by the constructions shown in Figs. 5.8*a* and 5.8*b*. In Fig. 5.8*a* [36] the first positively biased grid (G_3) is a pair of triangular plates, concentrating the main electron stream into a beam away from the oscillator anode (G_2). Variation of bias on G_4 has now very much less effect on the second-

ary emission from G_3 to G_2. Pentode I_aE_a characteristics are obtained by adding the suppressor grid (G_6). In Fig. 5.8*b* [35] the oscillator triode anode is separated from the frequency changing section by using half the cathode emission for the oscillator and half

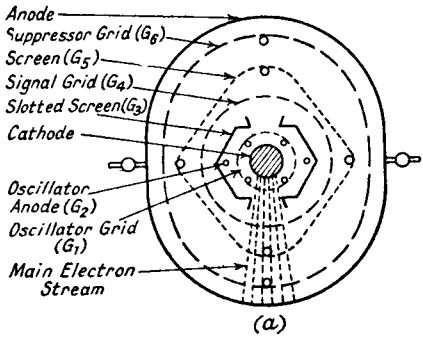


FIG. 5.8*a*.—The Electrode Structure of a Special Octode.

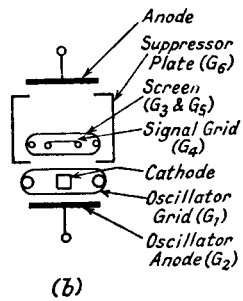


FIG. 5.8*b*.—An Improved Pentagrid Type of Frequency Changer.

for the frequency changer. The oscillator grid is, however, common to both circuits. The box-shaped suppressor plate helps to reduce capacitive coupling between oscillator and signal circuits and to suppress secondary emission from the frequency changer anode, thus raising the output impedance and giving pentode I_aE_a characteristics. Oscillator performance, especially on short waves, is greatly improved because of increased g_m, but conversion conductance suffers by using only one side of the cathode for frequency changing.

A disadvantage of a heptode operating over the short wave band is coupling between the oscillator and signal grids due to the common electron stream and this is discussed more fully in Section 5.8.3.

5.3.3. The Diode Frequency Changer. The detecting properties of the diode valve make its employment as a frequency changer possible, but it has not been used for this purpose to any extent because it has a low conversion gain (normally less than unity), a high oscillator harmonic response (Section 5.5 indicates that a high oscillator harmonic response is to be expected when a frequency changer valve operates into anode current cut-off), a comparatively poor signal-to-noise ratio,[41] grid current damping of the input signal circuit, and appreciable coupling between signal and oscillator circuits through the anode-cathode interelectrode capacitance. The property of high oscillator harmonic response (the oscillator second harmonic response ($2f_h - f_s'$) may be as much as

70% of the fundamental oscillator ($f_h - f_s$) conversion gain) may make it suitable as an ultra short wave frequency changer, since a lower oscillator frequency means increased oscillator frequency stability, a pressing problem on ultra short waves. Signal-to-noise ratio is low because of poor conversion gain and it is maximum for minimum diode conduction resistance (R_d of Section 8.2.5). Grid current damping can only be reduced by tapping down the signal-circuit coil and using a high conduction resistance diode with consequent loss of conversion efficiency. Interelectrode capacitance is decreased by increased electrode spacing (this, however, magnifies transit time effects at high frequencies) and decreased electrode area, both of which increase diode conduction resistance. A suitable circuit [30] for a diode frequency changer consists of an oscillator pick-up coil, tuned signal and intermediate frequency circuits, a diode valve, and a self-biasing resistance paralleled by a capacitance, all connected in series. The pick-up coil should be loosely coupled to the oscillator-tuned circuit so as to reduce interaction between it and the signal circuit, which is tapped down to reduce damping from the diode. The capacitance across the self-bias resistance must have a high enough value to bypass the I.F., but not too large to cause " non-tracking " distortion of the I.F. modulation envelope as described in Section 8.2.1. One point in favour of the diode is that when $\hat{E}_h$ is large, conversion gain is independent of the variation of $\hat{E}_h$, and I.F. output is linearly proportional to the input signal voltage.

5.4. Interference Whistle Production.[12, 17, 24, 27, 28]

5.4.1. Introduction. Interference whistles may be produced in a frequency changer whenever the signal circuit contains undesired frequencies which can combine amongst themselves or with the oscillator or its harmonics to produce a frequency near to the intermediate frequency. The former possibility is remote and need not be considered. Oscillator interference whistles, except that due to the image signal, against which even an ideal frequency changer cannot discriminate, are due mainly to distortion in the frequency changer. Harmonic distortion in the oscillator valve is rarely serious. To demonstrate the method by which interference may arise let us take the expression 5.4 for the $I_a E_g$ characteristic. Though this expression is concerned with common electrode application, similar results are obtained for separate electrode application by modifying the ideal expression of 5.2*b*. Using the same signal and oscillator designations but noting that the signal is an un-

desired frequency, i.e., $f_h - f_s \neq f_1$, and assuming for convenience that $E_b = 0$, expansion of the term $a_3 E_g^3$ in the series

$$I_a = a_0 + a_1 E_g + a_2 E_g^2 + a_3 E_g^3 - a_4 E_g^4$$

gives

$$\begin{aligned}
&a_3(\hat{E}_s \cos \omega_s t + \hat{E}_h \cos \omega_h t)^3 \\
&= a_3[\hat{E}_s \cos^3 \omega_s t + 3\hat{E}_s^2 \hat{E}_h \cos^2 \omega_s t \cos \omega_h t \\
&\qquad + 3\hat{E}_s \hat{E}_h^2 \cos \omega_s t \cos^2 \omega_h t + \hat{E}_h^3 \cos^3 \omega_h t] \\
&= a_3\left[\frac{\hat{E}_s^3}{4}(3 \cos \omega_s t + \cos 3\omega_s t) + \frac{3\hat{E}_s^2 \hat{E}_h}{2}(1 + \cos 2\omega_s t) \cos \omega_h t \right. \\
&\qquad \left. + \frac{3\hat{E}_s \hat{E}_h^2}{2}(1 + \cos 2\omega_h t) \cos \omega_s t + \frac{\hat{E}_h^3}{4}(3 \cos \omega_h t + \cos 3\omega_h t)\right] \\
&= a_3\left[\frac{\hat{E}_s^3}{4}(3 \cos \omega_s t + \cos 3\omega_s t)\right. \\
&\qquad + \frac{3\hat{E}_s^2 \hat{E}_h}{2}\left(\cos \omega_h t + \frac{\cos (2\omega_s + \omega_h)t + \cos (2\omega_s - \omega_h)t}{2}\right) \\
&\qquad + \frac{3\hat{E}_s \hat{E}_h^2}{2}\left(\cos \omega_s t + \frac{\cos (2\omega_h + \omega_s)t + \cos (2\omega_h - \omega_s)t}{2}\right) \\
&\qquad \left. + \frac{\hat{E}_h^3}{4}(3 \cos \omega_h t + \cos 3\omega_h t)\right] \qquad . \quad . \quad . \quad . \quad 5.6.
\end{aligned}$$

Eight frequencies are represented above by the pulsances ω_s, $3\omega_s$; ω_h, $3\omega_h$, $2\omega_s + \omega_h$, $2\omega_s - \omega_h$, $2\omega_h + \omega_s$, $2\omega_h - \omega_s$, and it is significant to note that they may be obtained by adding or subtracting the two fundamental pulsances ω_h and ω_s three at a time in all possible ways, but neglecting negative pulsances. For example

$$\begin{aligned}
&\omega_s + \omega_s - \omega_s = \omega_s, && \omega_h + \omega_h - \omega_h = \omega_h \\
&\omega_s + \omega_s + \omega_s = 3\omega_s, && \omega_h + \omega_h + \omega_h = 3\omega_h \\
&\omega_s + \omega_s + \omega_h = 2\omega_s + \omega_h && \omega_h + \omega_h + \omega_s = 2\omega_h + \omega_s. \\
&\omega_s + \omega_s - \omega_h = 2\omega_s - \omega_h && \omega_h + \omega_h - \omega_s = 2\omega_h - \omega_s
\end{aligned}$$

This rule may be extended to any term in the power series, and frequencies produced by the fourth power term can be obtained by taking the fundamental frequencies four at a time.

Most of the pulsances given by 5.6 do not contribute to the output voltage since the intermediate frequency circuit rejects them. However, if the signal frequency has certain values, the frequencies $\frac{2\omega_s - \omega_h}{2\pi}$ and $\frac{2\omega_h - \omega_s}{2\pi}$ may be very nearly equal to the intermediate frequency, and when this condition is fulfilled an interference whistle is produced. For example, for an I.F. of 465 kc/s and a signal frequency of 700 kc/s, the oscillator is

1,165 kc/s; undesired signal frequencies of 351, $(\omega_h - 2\omega_s)$, 816, $(2\omega_s - \omega_h)$, 1,867, $(2\omega_h - \omega_s)$ and 2,797 kc/s $(\omega_s - 2\omega_h)$, present at the grid of the frequency changer produce 2 kc/s interference tones if the $I_a E_g$ characteristic of the latter has a third power term $(a_3 E_g^3)$.

We can see from expression 5.6 that the amplitude of the terms bears a relationship to the frequency components; thus the frequency $\frac{2\omega_s - \omega_h}{2\pi}$ has an amplitude proportional to $\hat{E}_s^2 \hat{E}_h$. The factor multiplying the fundamental frequency indicates that the fundamental amplitude is raised to a power equal to the multiplying factor. Interference involving signal frequency harmonics usually has a smaller amplitude than that involving the oscillator harmonic since $\hat{E}_s$ is less than $\hat{E}_h$.

Interference whistles are conveniently classified as follows, the suffixes u, h and 1 denote the undesired signal, oscillator and intermediate frequency respectively.

(1) Image frequency, $f_u - f_h \simeq f_1$

(2) Combination of the signal and oscillator harmonics of different integers

$$\pm m f_h \mp n f_u \simeq f_1$$

(3) Combinations of harmonics of equal integers

$$m(\pm f_h \mp f_u) \simeq f_1$$

(4) Intermediate frequency harmonics

$$f_s \simeq n f_1$$

where f_s = the desired signal frequency.

The " approximately equals " sign is used to denote that the resultant of undesired signal and oscillator is within audio range of the I.F., i.e., equals $f_1 \pm 15$ kc/s.

5.4.2. Image Signal Interference. Image interference is generally located at certain points in the tuning range corresponding to signal frequencies equal to those of any local powerful transmissions minus twice the intermediate frequency. Since the frequency changer cannot discriminate against the image signal special precautions are necessary to improve the selectivity of the signal frequency circuits at a frequency $2f_1$ kc/s above the desired frequency, and circuits for achieving this are considered in Section 5.9.

5.4.3. Interference due to Combination of Different Harmonics of the Signal and Oscillator. Oscillator and

signal harmonic interference whistles of this type are spread over the tuning range in an irregular manner. Harmonics produced by the oscillator and R.F. stage can cause this interference with an ideal frequency changer having an I_aE_g characteristic in accordance with expression 5.2*a* or 5.2*b*, but this rarely happens. Distortion in the frequency changer is the more usual source, and the interference effect can be reduced by reducing the curvature (the higher power terms in the power series) of the I_aE_g characteristics for the signal and oscillator electrodes. Generally oscillator harmonic interference is the more serious—the signal tuned circuits attenuate undesired signals so that $E_u \ll E_h$—and it is reduced by making the oscillator amplitude as small as possible consistent with good conversion conductance.

5.4.4. Interference due to Combination of Equal Harmonics of the Signal and Oscillator. The characteristic of this type of interference is that the whistles are grouped around the oscillator frequency. For example, with an I.F. of 465 kc/s, a desired signal of 700 kc/s, and an oscillator of 1,165, possible interfering frequencies are 932, 1,009, 1,049, 1,398, 1,321, 1281 kc/s, i.e., $2(f_h - f_s)$, $3(f_h - f_s)$, $4(f_h - f_s)$, $2(f_s - f_h)$, etc.

5.4.5. Intermediate Frequency Harmonics. Harmonics of the intermediate frequency may cause interference if there is feedback from the I.F. amplifier to the signal circuits when the latter are tuned to nf_1. Serious interference and even instability is possible for an I.F. of 465 kc/s at signal tuning settings of 930, 1,395, 1,860 kc/s, etc. The usual cause of feedback is inadequate filtering of the A.G.C. bias, but it may also occur due to insufficient decoupling of the R.F., frequency changer and I.F. stages of the receiver, and to insufficient R.F. filtering between the detector and first A.F. amplifier. Improved A.G.C. bias, R.F. decoupling and screening usually reduces the effect to small proportions.

5.4.6. Interference Charts.[6] Interference frequencies are most conveniently expressed in the form of charts. A chart for an I.F. of 465 kc/s is shown in Fig. 5.9, and it is constructed in the following manner. The horizontal axis gives the signal tuning frequency of the receiver, and the vertical axis, with one exception, the interfering frequency. The exception is a line at 45° through the origin. This line gives the desired frequency on the vertical axis and is designated $(h - s)$, where h is the oscillator frequency and s, for all lines except $(h - s)$, is the interfering frequency. Lines due to harmonics of the signal and oscillator higher than the third are not shown on the chart. The image signal line

$(s - h)$ is a line parallel to $(h - s)$ and vertically above it by twice the intermediate frequency.

Interference lines due to unequal harmonics of signal and oscillator are distributed over the chart at different angles. For

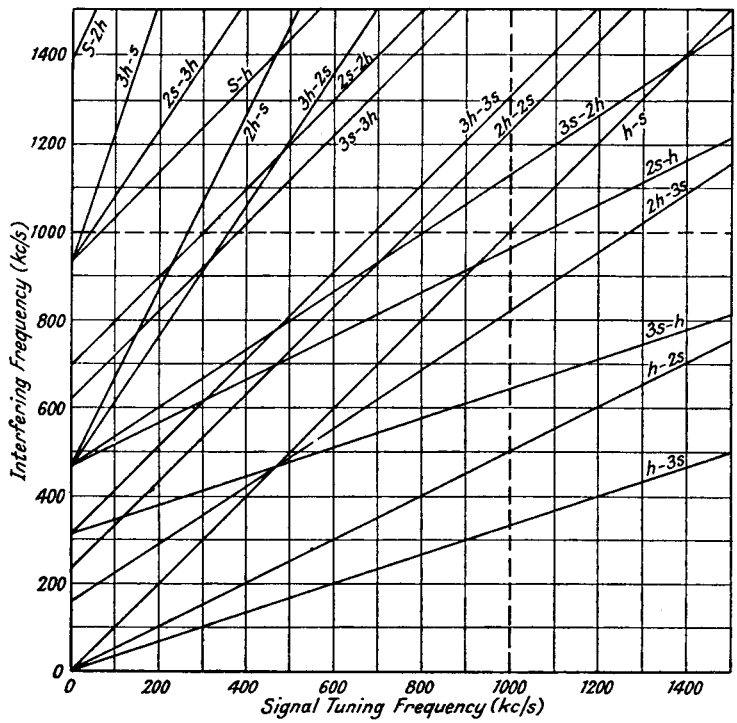


FIG. 5.9.—Interference Chart for an Intermediate Frequency of 465 kc/s.

example, the line corresponding to $(2h - s)$ passes through 100 kc/s (horizontal), 665 (vertical) and 500 (horizontal), 1,465 (vertical). The first point is calculated as follows:

$$\begin{aligned} \text{desired signal} &= 100 \text{ kc/s}; \quad \text{oscillator} = 565 \text{ kc/s} \\ \text{undesired signal} &= 2 \times \text{oscillator} - f_1 \\ &= 1{,}130 - 465 = 665 \text{ kc/s}. \end{aligned}$$

Lines corresponding to equal harmonics of signal and oscillator are parallel to the desired frequency line $(h - s)$ and spaced vertically above it by $\left(\frac{n-1}{n}\right)f_1$ for $h > s$ and $\left(\frac{n+1}{n}\right)f_1$ for $s > h$, where n is the number of the harmonic. Hence $3(h - s)$ is 310 kc/s and $3(s - h)$ is 620 kc/s above $(h - s)$.

To interpret the charts, let us consider a desired frequency of 1,000 kc/s. By drawing a vertical line from 1,000 kc/s on the horizontal axis we find eight intersections with the interference lines from 333·3 ($h - 3s$) to 1,310 kc/s ($3h - 3s$). If any of these undesired frequencies is present in the input to the frequency changer a whistle may be produced. Conversely, if 1,000 kc/s is the frequency of a strong local station its interference possibilities may be estimated by drawing a horizontal line from 1,000 kc/s on the vertical axis. The points of intersection (read on the horizontal axis) with the interference lines give the tuning points at which interference due to the local station may be expected. In Fig. 5.9 there are 12 tuning points from 0 to 1,500 kc/s at which interference might be expected.

5.5. The Maximum Value of Conversion Conductance.[20, 21] In Section 5.1.2 it is stated that the ratio of conversion conductance g_c, to mutual conductance g_m, rarely exceeds 0·25 in practice. We will now determine the maximum value of this ratio under various oscillator operating conditions. Differentiating the I_aE_g expression for grid-cathode application

$$I_a = a_0 + a_1E_g + a_2E_g^2$$

we have

$$\frac{dI_a}{dE_g} = a_1 + 2a_2E_g \quad . \quad . \quad . \quad . \quad 5.7a.$$

But $\frac{dI_a}{dE_g}$ is the slope of the I_aE_g characteristic curve, i.e., it is the mutual conductance g_m.

At a fixed bias voltage $-E_b$, we have

$$\left[\frac{dI_a}{dE_g}\right]_{E_g=-E_b} = g_m' = a_1 - 2a_2E_b \quad . \quad . \quad 5.7b.$$

The anode current component for a signal grid voltage of $\hat{E}_s \cos \omega_s t$ at the above bias voltage is

$$\begin{aligned} I_a &= g_m'\hat{E}_s \cos \omega_s t \\ &= (a_1 - 2a_2E_b)\hat{E}_s \cos \omega_s t \quad . \quad . \quad 5.8a. \end{aligned}$$

If an oscillator voltage $\hat{E}_h \cos \omega_h t$ is also applied, $-E_b$ in 5.8a becomes $(\hat{E}_h \cos \omega_h t - E_b)$ and

$$I_a = [a_1 + 2a_2(\hat{E}_h \cos \omega_h t - E_b)]\hat{E}_s \cos \omega_s t.$$

The I.F. component (I_1) in I_a is contained in the term

$$2a_2\hat{E}_h\hat{E}_s \cos \omega_h t \cos \omega_s t,$$

thus

$$I_1 = a_2\hat{E}_h\hat{E}_s \cos (\omega_h - \omega_s)t \quad . \quad . \quad . \quad 5.8b$$

but from 5.3*a*

$$g_c = \frac{I_1}{\hat{E}_s} = a_2\hat{E}_h.$$

To interpret $a_2\hat{E}_h$, and hence g_c, in terms of g_m, let us imagine that the bias is decreased to $(-E_b+\hat{E}_h)$, at which the mutual conductance is $g_{m(max.)}$. Similarly at a bias of $-(E_b+E_h)$ let mutual conductance be $g_{m(min.)}$ Now

$$g_{m(max.)} = a_1+2a_2(\hat{E}_h - E_b)$$

and

$$g_{m(min.)} = a_1 - 2a_2(\hat{E}_h+E_b)$$

so that

$$g_{m(max.)} - g_{m(min.)} = 4a_2E_h.$$

Hence

$$g_c = \frac{g_{m(max.)} - g_{m(min.)}}{4} \quad . \quad . \quad . \quad 5.9a$$

or the conversion conductance is 0·25 of the change in mutual conductance produced by the application of the oscillator voltage. Maximum change of mutual conductance is obtained when g_m is varied from 0 to the value at which grid current begins. Since a valve operating as an R.F. amplifier would be biased initially almost to the start of grid current- -its mutual conductance is then $g_{m(max.)}$— we may say that the ratio of conversion conductance as a frequency changer to mutual conductance of the same valve acting as an R.F. amplifier is not greater than 0·25 as long as the oscillator takes the valve only to cut-off and not beyond.

The above method is not easily applicable when the oscillator voltage causes the anode current to be cut off for a part of the cycle, but if $\hat{E}_h \gg \hat{E}_s$ the following analysis may be employed. Suppose we take the expression for g_m in 5.7*b* and replace $-E_b$ by $-E_b+\hat{E}_h \cos \omega_h t$. At the left-hand side of Fig. 5.10, the curve of g_m against bias voltage (it is a straight line as shown by 5.7*b*) is drawn, whilst on the right-hand side the variation of g_m against time is plotted for two values of oscillator voltage; for curve 2 cut-off point is passed. Taking first curve 1, the coefficient of $\cos \omega_h t$ in this curve may be obtained by Fourier analysis, thus:

Coefficient of $\cos \omega_h t$ in curve 1

$$= \frac{\omega_h}{\pi}\int_0^{\frac{2\pi}{\omega_h}} [a_1+2a_2(\hat{E}_h \cos \omega_h t - E_b)] \cos \omega_h t \, dt$$

$$= \frac{\omega_h}{\pi}\left[\frac{a_1}{\omega_h} \sin \omega_h t+2a_2\hat{E}_h\left[\frac{t}{2}+\frac{\sin 2\omega_h t}{4\omega_h}\right] - \frac{2a_2}{\omega_h}E_b. \sin \omega_h t\right]_0^{\frac{2\pi}{\omega_h}}$$

$$= \frac{\omega_h}{\pi}2a_2\hat{E}_h\frac{2\pi}{2\omega_h} = 2a_2\hat{E}_h$$

$$= 2g_c \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad 5.9b.$$

Multiplying this fundamental oscillator component in the g_m curve by $\hat{E}_s \cos \omega_s t$ gives

$$2a_2\hat{E}_h\hat{E}_s \cos \omega_h t \cos \omega_s t$$

and the I.F. component in this is $a_2\hat{E}_h\hat{E}_s \cos (\omega_h - \omega_s)t$. We therefore see that half the coefficient of $\cos \omega_h t$ in the g_m curve plotted against time gives the conversion conductance. This method of attack is applicable to any shape of g_m curve provided $\hat{E}_h \gg \hat{E}_s$.

Let us apply it to the case where the oscillator voltage is large enough to cut off the anode current during a part of the cycle. For

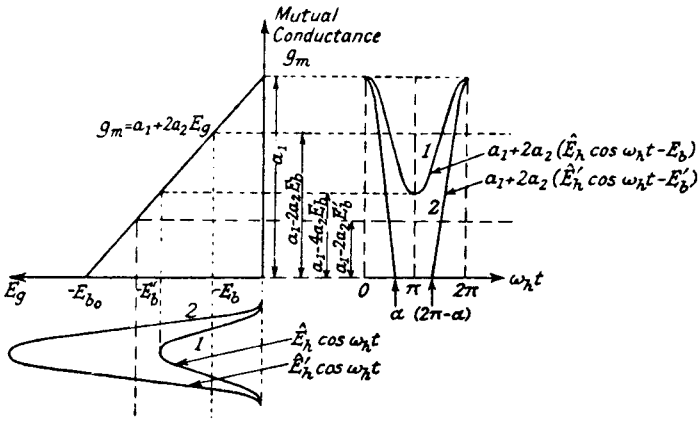


FIG. 5.10.—Linear Mutual Conductance Characteristic for the Normal Type of Frequency Changer.

simplification the maximum positive swing of the oscillator voltage will be assumed to be always 0 volts and the mutual conductance at this point to be $g_{m(max.)}$. The cut-off bias voltage, found by putting $g_m = 0$ in 5.7b, is $\dfrac{-a_1}{2a_2}$ and this we will designate by $-E_{b0}$. The equation to the g_mE_g curve in terms of $g_{m(max.)}$ and $-E_{b0}$ is

$$g_m = g_{m(max.)}\left[\frac{E_g}{E_{b0}} + 1\right] \quad . \quad . \quad . \quad 5.10a.$$

For oscillator and bias voltages of $\hat{E}_h' \cos \omega_h t$ and $-E_b'$ respectively we have

$$\begin{aligned} E_g &= \hat{E}_h' \cos \omega_h t - E_b' \\ &= \hat{E}_h'(\cos \omega_h t - 1) \end{aligned}$$

for $\hat{E}_h' \cos \omega_h t$ is assumed to take the bias to zero at the maximum

positive peak so that $E_b' = \hat{E}_h'$. To simplify let $\omega_h t = \theta$; equation 5.10*a* becomes

$$g_m = g_{m(max.)}\left[\frac{\hat{E}_h'}{E_{b0}}(\cos\theta - 1)+1\right] \quad . \quad . \quad 5.10b$$

and the angle α corresponding to cut-off bias $-E_{b0}$ is obtained from 5.10*b* by putting $g_m = 0$

$$\frac{\hat{E}_h'}{E_{b0}}(\cos\alpha - 1) = -1$$

or

$$\alpha = \cos^{-1}\left(1 - \frac{E_{b0}}{\hat{E}_h'}\right) \quad . \quad . \quad . \quad 5.11.$$

The coefficient of $\cos\theta$ in 5.10*b* is given by Fourier analysis over the intervals 0 to α and $(2\pi - \alpha)$ to 2π, or twice the value obtained from 0 to α. (See Appendix 2A.)

Coefficient of $\cos\theta$

$$= \frac{2}{\pi}\int_0^{\alpha} g_{m(max.)}\left[\frac{\hat{E}_h'}{E_{b0}}(\cos\theta - 1)+1\right]\cos\theta\, d\theta$$

$$= \frac{2g_{m(max.)}}{\pi}\left[\sin\alpha\left(1 - \frac{\hat{E}_h'}{E_{b0}}\right)+\frac{\hat{E}_h'}{E_{b0}}\left(\frac{\alpha}{2}+\frac{\sin 2\alpha}{4}\right)\right].$$

But from 5.11

$$\frac{\hat{E}_h'}{E_{b0}} = \frac{1}{1 - \cos\alpha}.$$

$\therefore$ Coefficient of $\cos\theta$

$$= \frac{2g_{m(max.)}}{\pi}\left[\frac{-\sin\alpha\cos\alpha}{1-\cos\alpha}+\frac{1}{1-\cos\alpha}\left(\frac{\alpha}{2}+\frac{\sin 2\alpha}{4}\right)\right]$$

$$= \frac{g_{m(max.)}}{\pi}\left[\frac{\alpha - \sin\alpha\cos\alpha}{1-\cos\alpha}\right] \quad . \quad . \quad . \quad . \quad 5.12a$$

$$= 2g_c$$

for as shown above in expression 5.9*b* $g_c = \frac{1}{2}$ the coefficient of $\cos\theta$ in the g_m expression.

The ratio of $\frac{g_c}{g_{m(max.)}}$ for different values of $\frac{\hat{E}_h}{E_{b0}}$ is plotted as curve 1 in Fig. 5.11*a*, and it is to be noted that a maximum g_c of 0·268 $g_{m(max.)}$ is obtained when $\hat{E}_h = 0{\cdot}65.E_{b0}$; i.e., when the valve is partially cut off by the oscillator voltage. There are disadvantages to operation into cut-off since oscillator harmonics, which can combine with undesired signals to cause interference whistles, are produced. If we limit $\hat{E}_h$ to $\frac{1}{2}E_{b0}$, no cut-off occurs, g_c is only reduced by about 8%, and for this particular (linear) g_m curve oscillator harmonic interference is absent, since the com-

ponent of cos $n\theta$, where n is greater than 1, in the Fourier analysis is zero over the intervals 0 to π and π to 2π. The practical form of $g_m E_g$ curve is seldom linear so that interference whistles are possible even when $\hat{E}_h < \frac{E_{b0}}{2}$ but they are considerably increased as soon as $\hat{E}_h$ exceeds $\frac{E_{b0}}{2}$.

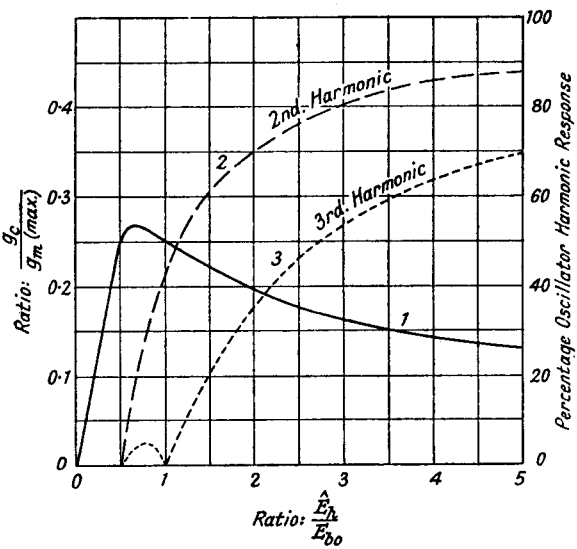


FIG. 5.11*a*.—Conversion Conductance and Oscillator Harmonic Response Curves for the Normal Type of Frequency Changer with Linear Mutual Conductance Characteristic.

We can calculate the ratio of oscillator harmonic to fundamental conversion conductance by finding the coefficient of cos $n\theta$ in 5.10*b* over the interval 0 to α in the same manner as for the fundamental. Thus for the second harmonic conversion conductance

$$g_c(2h) = \frac{g_{m(max.)}}{\pi}\int_0^{\alpha}\left[\frac{\hat{E}_h}{E_{b0}}(\cos\theta - 1)+1\right]\cos 2\theta\, d\theta.$$

$$= \frac{g_{m(max.)}}{\pi}\int_0^{\alpha}\left[\frac{\hat{E}_h}{E_{b0}}\left(\frac{\cos\theta+\cos 3\theta}{2} - \cos 2\theta\right)+\cos 2\theta\right]d\theta$$

$$= \frac{g_{m(max.)}}{\pi}\left[\frac{\hat{E}_h}{E_{b0}}\left(\frac{\sin\alpha}{2}+\frac{\sin 3\alpha}{6} - \frac{\sin 2\alpha}{2}\right)+\frac{\sin 2\alpha}{2}\right]$$

$$= \frac{g_{m(max.)}}{2\pi}\left[\frac{\hat{E}_h}{E_{b0}}\left(\sin\alpha+\frac{\sin 3\alpha}{3} - \sin 2\alpha\right)+\sin 2\alpha\right] \quad . \quad 5.12b$$

where α is as given in 5.11.

The ratio of $\frac{g_c(2h)}{g_c(h)} \times 100\%$, which we will call the percentage second harmonic response, is shown in curve 2 of Fig. 5.11*a*. It is zero until $\hat{E}_h$ exceeds $\frac{E_{b0}}{2}$ and steadily rises as $\hat{E}_h$ is increased.

Similarly it can be shown that the conversion conductance for the third harmonic of the oscillator is

$$g_c(3h) = \frac{g_{m(max.)}}{\pi}\left[\frac{\hat{E}_h}{E_{b0}}\left(\frac{\sin 4\alpha}{8}+\frac{\sin 2\alpha}{4}-\frac{\sin 3\alpha}{3}\right)+\frac{\sin 3\alpha}{3}\right]. \quad 5.12c.$$

The percentage third harmonic response $\frac{g_c(3h)}{g_c(h)} \times 100\%$ is the dotted curve 3 of Fig. 5.11*a*. It has zero value until $\hat{E}_h$ exceeds $\frac{E_{b0}}{2}$, rises to a maximum of 6·0% at $\hat{E}_h = 0{\cdot}8.E_{b0}$, falls to zero at $\hat{E}_h = E_{b0}$ and thereafter rises steadily. The need for preventing too high a value of oscillator voltage is made very clear from the curves.

The above analysis is applicable to all types of $g_m E_g$ curves, but when a separate electrode is used for the oscillator voltage the g_m curve for the oscillator grid must be considered.

Taking expression 5.2*b* as representative of the hexode valve we obtain after differentiating with respect to E_{g1}

$$\frac{dI_a}{dE_{g1}} = g_m = a_1(b_0+b_1E_{g3})$$

and plotting this against E_{g3}, the bias on the oscillator grid, we have a straight line similar to that in Fig. 5.10. Proceeding along the lines set out above, we obtain a result for conversion conductance which is identical with that for common electrode application of the oscillator. Thus the curves in Fig. 5.11*a* are also applicable to the hexode valve, where $g_{m(max.)}$ equals a_1b_0, the value of g_m when the oscillator grid bias is zero.

We can estimate the maximum conversion conductance obtainable from a linear I_aE_g characteristic, such as that for a non-variable mu valve or a diode frequency changer, by the same method. Since I_a is linearly proportional to E_g, the g_mE_g curve is a straight line parallel to the E_g axis, cutting the g_m axis at $g_{m(max.)}$. At the cut-off bias voltage, $-E_{b0}$, it falls sharply to zero. The curve of mutual conductance against time for a given oscillator peak voltage exceeding $-E_{b0}$ is therefore rectangular with zero g_m between α and $2\pi - \alpha$, where $\alpha = \cos^{-1}\left(1 - \frac{E_{b0}}{\hat{E}_h}\right)$. Conversion

conductance is given by half the coefficient of $\cos\theta$ in $g_m = g_{m(max.)}$ over twice the interval 0 to α

$$\therefore\ g_c = \frac{1}{2}\left[\frac{2}{\pi}\int_0^{\alpha} g_{m(max.)} \cos\theta\, d\theta\right]$$

$$= \frac{g_{m(max.)} \sin\alpha}{\pi} \quad . \quad . \quad . \quad . \quad 5.13a$$

$$= \frac{g_{m(max.)}}{\pi}\sqrt{\frac{2E_{b0}}{\hat{E}_h} - \frac{E_{b0}^2}{\hat{E}_h^2}} \quad . \quad . \quad 5.13b.$$

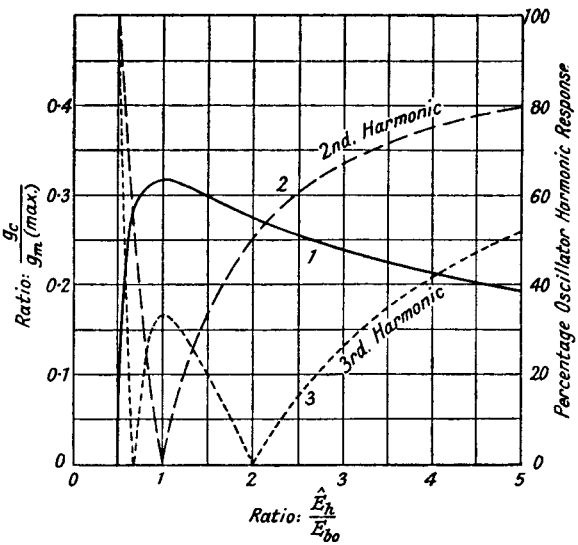


FIG. 5.11*b*.—Conversion Conductance and Oscillator Harmonic Response Curves for a Frequency Changer with Constant Mutual Conductance Characteristic.

This result is plotted in Fig. 5.11*b* as curve 1, and we see that when $\hat{E}_h < \frac{E_{b0}}{2}$ there is no cut-off and hence no frequency changing. This is to be expected since the output wave shape is an exact reproduction of the original input. Actually when $\hat{E}_h = \frac{E_{b0}}{2}$, the addition of the signal voltage causes the cut-off point to be passed and some small g_c value is obtained. When $\hat{E}_h = E_{b0}$, $g_c = \frac{g_{m(max.)}}{\pi}$ and this is its maximum value.

The percentage second and third harmonic responses are found from the coefficients of $\cos 2\theta$ and $\cos 3\theta$ in the g_m — time curve.

Thus $$g_c\left(\frac{2h}{h}\right) = 100 \cos \alpha\,\% = \left(1 - \frac{E_{b0}}{\hat{E}_h}\right) 100\,\%.$$

and $$g_c\left(\frac{3h}{h}\right) = \frac{100 \sin 3\alpha}{3 \sin \alpha}\,\%.$$

The curves 2 and 3 show the variation of harmonic response for different ratios of $\frac{\hat{E}_h}{E_{b0}}$. Second harmonic is zero at $\frac{\hat{E}_h}{E_{b0}} = 1$, whilst third harmonic response is zero at $\frac{\hat{E}_h}{E_{b0}} = 0{\cdot}66$ and 2. For large values of oscillator voltage the linear $I_a E_g$ characteristic has lower oscillator harmonic responses than the parabolic characteristic operated into cut-off, and this is largely due to a higher fundamental conversion conductance.

5.6. Measurements on Frequency Changers.

5.6.1. Introduction. The most important measurements which are required are those of conversion conductance, oscillator harmonic response, and signal handling capacity for a given modulation percentage and distortion of the modulation envelope.

5.6.2. Conversion Conductance. Conversion conductance may be measured by direct or indirect means. The first involves the use of low frequency input voltages for the signal and oscillator, whilst the second method uses frequencies in the range over which it is intended to operate. The former has the advantage of requiring only simple apparatus, but the results may not show close agreement with measurements made under short wave operating conditions, because of transit time of electrons and increased grid admittance. However, for the medium and long wave bands agreement between the two methods is generally good and variations rarely exceed 5%.

5.6.3. Indirect Measurements of Conversion Conductance. The voltages for one indirect method are obtained from the 50 c.p.s. mains supply, and Fig. 5.12 shows a diagram of connections for a hexode valve. The diagram is similar for a pentode except that oscillator and signal inputs may be applied to the same electrode. Suitable bias and screen voltages are applied, and the signal voltage is fixed at some convenient value such as 0·1 to 0·5 peak volts. A value of 0·5 volts should not be exceeded or it may be found that the measured g_c is not independent of the signal voltage. The change-over switch (S) in the signal circuit allows the phase relationship between signal and oscillator to be reversed. The oscillator

voltage is variable and is obtained from the same transformer as the signal voltage. Operation of the change-over switch produces a change of mean current in the D.C. milliammeter, and this change is proportional to the conversion conductance. The relationship is expressed by

$$g_c = \frac{I_{a1} - I_{a2}}{2\hat{E}_s} \quad . \quad . \quad . \quad . \quad 5.14$$

where I_{a1} = the mean current for the in-phase condition of $\hat{E}_s$ and $\hat{E}_h$
and I_{a2} = ,, ,, ,, ,, ,, out-of-phase ,, ,, ,, ,, ,,

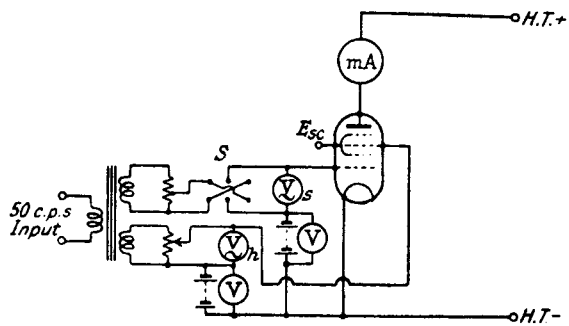


FIG. 5.12.—A Circuit for the Measurement of Conversion Conductance of a Hexode Valve by means of the 50 c.p.s. Mains Supply.

The proof of the above expression may be derived from 5.2*b*, thus :

$$I_a = (a_0 + a_1 E_{g1})(b_0 + b_1 E_{g3}).$$

Replacing E_{g1} by $(\hat{E}_s \cos\theta - E_{b1})$ and E_{g3} by $(\hat{E}_h \cos\theta - E_{b3})$ we have for the in-phase current

$$(a_0 + a_1(\hat{E}_s \cos\theta - E_{b1}))(b_0 + b_1(\hat{E}_h \cos\theta - E_{b3}))$$

the D.C. component of which is

$$I_{a1} = a_0 b_0 - a_1 b_0 E_{b1} - a_0 b_1 E_{b3} + \frac{a_1 b_1}{2}\hat{E}_s\hat{E}_h.$$

For the out-of-phase condition $E_{g1} = -\hat{E}_s \cos\theta - E_{b1}$, and the D.C. component is

$$I_{a2} = a_0 b_0 - a_1 b_0 E_{b1} - a_0 b_1 E_{b3} - \frac{a_1 b_1}{2}\hat{E}_s\hat{E}_h$$

$$I_{a1} - I_{a2} = a_1 b_1 \hat{E}_s \hat{E}_h.$$

But from expression 5.3*b* (Section 5.1.3)

$$g_c = \frac{a_1 b_1}{2}\hat{E}_h = \frac{I_{a1} - I_{a2}}{2\hat{E}_s}.$$

The above result assumes that there is no distortion of the hexode $I_a E_g$ characteristics, and since in practice it is rare to obtain

linear characteristics there are possibilities of error in the measurement. The error is, however, not usually very serious, particularly if the signal voltage is not made large.

A second indirect method [21] employs one A.C. voltage only—no signal voltage being used—and the circuit is shown in Fig. 5.13. A filter in the anode circuit attenuates all but the fundamental A.C. voltage. The input fundamental, which must be completely free

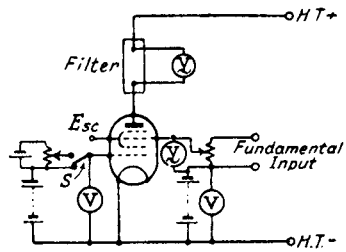


FIG. 5.13.—A Circuit for the Measurement of Conversion Conductance of a Hexode Valve by means of the Change of Fundamental Output.

from distortion, acts as the oscillator voltage, and the input signal is simulated by varying the signal grid bias a predetermined amount such as 0·5 volts. The change in fundamental output is noted and conversion conductance is given by,

$$g_c = \frac{\text{output fundamental voltage change}}{2 \times \text{signal grid bias change} \times Z_0 \text{ (anode circuit filter)}}.$$

This may be proved as follows:

$$I_a = (a_0 + a_1 E_{g1})(b_0 + b_1 E_{g3})$$

where $$E_{g3} = \hat{E}_h \cos\theta - E_{b3}$$

$$\therefore\ I_a = a_0 b_0 + a_1 b_0 E_{g1} + a_0 b_1 (\hat{E}_h \cos\theta - E_{b3}) + a_1 b_1 E_{g1} (\hat{E}_h \cos\theta - E_{b3})$$

and the fundamental A.C. component in the above is

$$I_f = (a_0 b_1 + a_1 b_1 E_{g1}) \hat{E}_h \cos\theta.$$

If E_{g1} is decreased by a value of $+E_s$ the change in fundamental current component is

$$\Delta I_f = a_1 b_1 E_s \hat{E}_h \cos\theta$$

but

$$g_c = \frac{a_1 b_1 \hat{E}_h}{2}$$

$$= \frac{\Delta I_f}{2E_s} = \frac{\Delta I_f Z_0}{2E_s Z_0} = \frac{\Delta E_f}{2E_s Z_0} = \frac{\Delta E_f}{Z_0} \quad . \quad . \quad 5.15$$

if $$E_s = 0{\cdot}5 \text{ volts}.$$

Both indirect methods can be made to give satisfactory results compared with direct measurements, but for frequency changers operating over short wave ranges it is often preferable to use a direct method which takes into account coupling between signal and oscillator circuits, and electron transit time effects.

5.6.4. Direct Measurement of Conversion Conductance. For direct measurement of conversion conductance a signal frequency is chosen at approximately the centre of the desired range of tuning frequencies. Suitable values are 700 kc/s for medium and long wave bands and 10 Mc/s for the short wave. A schematic diagram is given in Fig. 5.14. The signal voltage, supplied by a signal generator, must be limited to a value not exceeding 0·5 peak volts, though for high gain frequency changers, with a high anode impedance, it may be necessary to reduce to 0·1 peak volts if anode voltage distortion is to be prevented. The oscillator voltage is obtained from an oscillator followed by an amplifier, the gain of

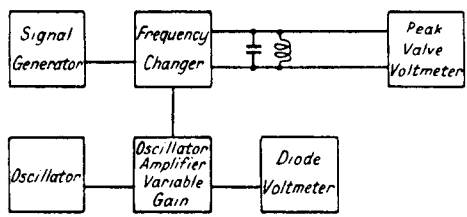


FIG. 5.14.—Schematic Diagram for the R.F. Measurement of Conversion Conductance.

which is changed by bias variation. In the anode circuit of the amplifier is a tuned circuit, which rejects oscillator harmonics. This is essential for harmonic response measurement. Large oscillator voltages are obtained by capacitance-resistance coupling from this tuned circuit to the appropriate oscillator electrode in the frequency changer. Voltages not exceeding about 15 peak volts are obtained from a pick-up coil wound on the earthed end of the tuned circuit. A diode voltmeter may conveniently be used for measuring the oscillator volts across the pick-up coil, and it may also be calibrated to give the voltage across the tuned circuit by noting the diode current for different voltages measured by a slide-back or peak voltmeter across the tuned circuit.

The anode circuit of the frequency changer consists of a parallel tuned circuit resonant at the I.F., and a slide-back or peak voltmeter is used to measure the output voltage. This type of voltmeter has two advantages; its range of measurement is large—the output

may reach a peak value of 100 volts—and its grid input conductance can be made very low so that there is little damping of the output circuit. To calculate conversion conductance we must measure the resonant or dynamic resistance of the tuned output circuit, and this can be accomplished as follows. The bias on the signal grid of the frequency changer is increased to a value (about — 8 volts) at which the anode slope resistance can be expected to be much higher than the tuned-circuit impedance. The input signal is adjusted to give a certain output voltage, such as 1 volt peak (a low value is chosen to prevent the possibility of non-linearity between input and output voltages). The tuned circuit is next paralleled by a known non-inductive resistance and the input signal increased to give the same output voltage, viz., 1 volt. If $\hat{E}_s'$ and $\hat{E}_s''$ are the input peak signal voltages with and without the non-inductive resistance R, we have

$$\hat{E}_0 = 1 = g_c\hat{E}_s'R_D = g_c\hat{E}_s''\frac{R_DR}{R_D + R}$$

where $\hat{E}_0$ = the output peak voltage
g_c = conversion conductance
R_D = the dynamic resistance of the tuned circuit

$$\therefore\ R_D = R\left[\frac{\hat{E}_s''}{\hat{E}_s'} - 1\right].$$

If we can assume that $R_a \gg R_D$ we can now calculate conversion conductance for

$$g_c = \frac{\hat{E}_0}{\hat{E}_sR_D} \quad . \quad . \quad . \quad . \quad 5.16.$$

5.6.5. Measurement of Oscillator Harmonic Response. Interference whistles from interaction between undesired signals and harmonics of the oscillator are normally more serious than undesired signal harmonic interference, and the ratio of conversion conductance for the second and third oscillator harmonics to that for the fundamental is a good indication of the interference capability of the frequency changer.

The apparatus is the same as that used in the previous section, but the signal frequency is adjusted to give the I.F. by interaction with the particular oscillator harmonic under consideration. For example, if desired signal and I.F. frequencies are 700 and 465 kc/s respectively, the signal frequencies for measuring second and third harmonic responses are 1,865 (2 × 1,165 — 465) and 3,030 kc/s (3 × 1,165 — 465). It is essential that the oscillator voltage source be free from harmonics, and this may be checked with

oscillator amplifier at full gain by placing a potentiometer across the output and reducing the oscillator voltage applied to the frequency changer to about one-fifth of its optimum value. The oscillator harmonic response under these conditions should be not greater than 1% of the fundamental. The percentage oscillator harmonic response $\left(\frac{g_c \text{ harmonic}}{g_c \text{ fundamental}} \times 100\%\right)$ is plotted in Fig. 5.15 against signal grid bias under normal conditions for the hexode type (full line), and the pentode with common electrode application. Increase of oscillator voltage increases both second and third harmonic response maxima and may change the position of third harmonic minima. The harmonic response for the hexode has generally lower maxima than the pentode and is less variable with

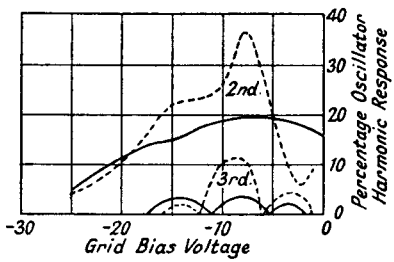


Fig. 5.15.—Typical Oscillator Harmonic Response Curves for a Pentode and Hexode Valve.
(Dotted Line—Pentode. Full Line—Hexode.)

bias change, because oscillator harmonics are caused by curvature of the $I_a E_{g_h}$ characteristic, which is less affected by signal grid-bias variation than the $I_a E_g$ characteristic of the pentode.

5.6.6. Signal Handling Capacity. The maximum modulated input signal and output I.F. voltage, which can be handled by a frequency changer, is defined as the carrier peak voltage modulated K% which gives 5% total harmonic distortion of the audio frequency envelope of the carrier. The modulation percentage is usually fixed at 80% in order to keep distortion low in the apparatus other than the frequency changer. The apparatus is substantially similar to that described in Section 7.11, except for the inclusion of the oscillator voltage at the appropriate point.

Typical curves for a pentode with cathode application of the oscillator and for a hexode are illustrated in Fig. 5.16. The input and output carrier voltage curves for the hexode are lower than those for the pentode because of the semi-screened grid $I_a E_a$ char-

acteristics of the former. Conversion conductance obtained by calculation from the input and output voltages is always higher at a large negative bias than conversion conductance measured with a small signal voltage. This is due to the fact that the input signal is large and the variable mu characteristic of the signal grid

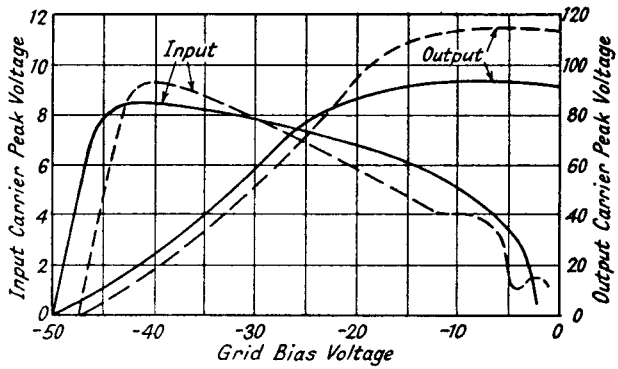


FIG. 5.16.—Typical Signal Handling Capacity Curves for Pentode and Hexode Frequency Changers.
(5% A.F. Envelope Distortion and 80% Modulation of Carrier.)
(Dotted Line—Pentode. Ful, Line—Hexode.)

causes the output voltage to increase at a greater rate than the input. The importance of these curves is discussed in Section 12.4.2 on automatic gain control.

5.7. The Properties Required of a Frequency Changer Valve.[31]

5.7.1. Introduction. Having discussed the salient features of frequency changing, we can now specify the essential requirements of a frequency changer valve. It should have

(1) A low value of anode and total current, and high slope resistance.
(2) A high value of conversion conductance, which is maintained in the short wave ranges.
(3) Low oscillator harmonic response.
(4) Minimum cross-modulation.
(5) Minimum coupling between oscillator and signal circuits.
(6) Minimum variation of signal grid-cathode capacitance for bias variations on the signal grid.
(7) Low signal grid input admittance.
(8) Small oscillator frequency drift.
(9) Minimum microphonic effects.

5.7.2. Anode and Total Current, Slope Resistance. Low values of anode and total currents aid economy of operation (this

is only an absolute essential in battery receivers), but their importance, in conjunction with g_c, is their influence on shot noise. The equivalent shot-noise voltage at the grid of the frequency changer is approximately $\frac{K\sqrt{I_a}}{g_c}.\sqrt{\frac{2.\Delta f}{10{,}000}}\mu$V; where K is a factor dependent on the type of valve and varying between 0·5 and 1·5, and $2\Delta f$ is the pass-band width of the I.F. amplifier. The total current does not appear directly in the formula, but is actually contained in K, which is increased as the total current is increased. The importance of minimum I_a and I_T is thus apparent.

Since for the triode valve, anode and total currents are equal, signal-to-noise ratio is comparatively high, but this type of valve is quite unsuitable for frequency changing because of its low slope resistance, R_a, which results in low amplification and heavy damping of the primary of the I.F. transformer with consequent loss of selectivity. The tetrode or pentode valve is superior to the hexode or heptode because it has a lower ratio of total cathode to anode current, about 1·3 to 1 as compared with from 2 to 3 to 1 for the heptode and hexode. In certain beam tetrodes with aligned grids, screen current is reduced from the usual value of 25% of the anode current to 10% or less, and signal-to-noise ratio for these valves approaches that of the triode. Taking the signal-to-noise ratio of the latter as a basis of comparison, the beam tetrode with aligned grids, the normal tetrode or pentode, and the heptode or hexode have signal-to-noise ratios of about 0·8, 0·5 and 0·25 (respectively) [38] of that for a triode. If the difficulties introduced by coupling between signal and oscillator circuits can be overcome, such as by using an untuned signal circuit, a tetrode with grid-cathode oscillator application is preferable to a hexode or heptode when there is no preceding R.F. stage, because of its greater signal-to-noise ratio. The effects of coupling on the short wave range can also be reduced by using a high I.F. and preset signal-tuned circuits, as occurs in some types of band-spread receivers. In the multi-electrode valve, any secondary emission currents to the anode add their quota of noise so that a high anode resistance R_a, which indicates small secondary emission from the screen or any other possible emitting surface (apart from the cathode) to the anode, is desirable; at the same time it increases the conversion stage gain, which leads to increased signal-to-noise ratio.

5.7.3. Conversion Conductance. In order to obtain the highest signal-to-noise ratio the conversion conductance should have the highest possible value. As this tends to produce high

harmonic response, a compromise value of conversion conductance must be chosen, and generally it is about 85% of the maximum value.

5.7.4. Oscillator Harmonic Response. For small oscillator harmonic response the curve connecting g_c and oscillator voltage should be straight, and oscillator voltage limited to the maximum value which can be obtained on the straight part of this curve.

5.7.5. Cross-Modulation. Cross-modulation is reduced by high R.F. selectivity before the frequency changer and by having minimum curvature of the $g_c E_g$ characteristic of the signal grid. Signal harmonic response is also reduced by this method.

5.7.6. Signal and Oscillator Circuit Interaction. In an ideal frequency changer, variation in signal circuit tuning should have no effect on oscillator frequency and amplitude. Coupling by interelectrode capacitance and the common electron stream occurs in practice, and the signal circuit tuning always has some influence. Its effect is most pronounced at the higher short wave frequencies where the frequency ratio between oscillator and signal is smallest.

5.7.7. Signal Grid-Cathode Capacitance Variation. A capacitance variation of about 2 $\mu\mu$F occurs when the signal grid bias is varied and is due to variation of the distance between the grid and the virtual cathode produced by space charge. It is a function of the total current passing through the control grid and the greatest mistuning effect occurs for small signal-tuning capacitances, i.e., at the high-frequency end of any given range. The effect is much less pronounced in the heptode than in the hexode valve.

5.7.8. Low Signal Grid Input Admittance. The conductance component of the signal grid input admittance of a hexode frequency changer generally increases as the signal frequency increases due to the inductance of the cathode lead and feedback through the grid-cathode capacitance, and also due to transit time of the electrons. A heptode valve may show a negative input conductance, which decreases as the signal grid bias is increased, becoming zero at a particular bias voltage and finally asymptotic to a positive value equal to the losses in the valve-holder, etc., as measured when the valve-heater is open-circuited (the valve is cold). Both these effects are discussed in Section 5.8. Anode-grid capacitance coupling has little influence on input conductance because the value of capacitance is small and the anode circuit is an I.F. transformer offering a very low impedance to the signal frequencies.

5.7.9. Oscillator Frequency Drift. Oscillator frequency variations due to the frequency changer valve itself are covered by

5.7.6 and are more fully discussed in Section 5.8.1. Those due to the oscillator and its component parts are dealt with in Sections 6.6 and 7.

5.7.10. Microphony. Microphony is largely affected by electrode structure and supports. The position of the valve with respect to the loud-speaker and the use of a shock-absorbing valve mounting can determine the extent of microphonic noises. Sound-wave coupling may also occur between the loudspeaker and rotor plates of the oscillator tuning capacitance, and to reduce this the complete tuning capacitance chassis is often mounted on rubber. Increasing the thickness and spacing of the rotor plates of the oscillator tuning capacitance also reduces the tendency to sound coupling.

5.8. Special Considerations in Short Wave Frequency Changing.[25]

5.8.1. Introduction. Short and ultra short wave frequency changing is chiefly complicated by the increased effect of capacitive and electronic coupling, the reduced frequency ratio between signal and oscillator and increased signal grid input admittance, which becomes comparable with the resonant conductance $\frac{1}{R_D}$ of the signal-tuned circuit.

Capacitance coupling has the same effect in hexode or heptode valves. It normally precludes the use of common electrode application of signal and oscillator frequencies, though for special reasons (such as increased signal-to-noise ratio) a tetrode [38] may be used as a frequency changer with common coupling for ultra high frequency conversion (40 Mc/s). The signal circuit is then either untuned or preset tuned, and the oscillator voltage made as low as possible consistent with satisfactory frequency changing; the oscillator frequency can be less than that of the signal. Inter-electrode capacitance coupling has a two-way effect which can be observed when the valve-heater is open-circuited (no emission). The signal circuit reflects through the signal grid-oscillator grid capacitance an impedance into the oscillator-tuned circuit, and both tuning and amplitude of the oscillator voltage are affected. In the reverse direction this coupling induces in the signal circuit an oscillator voltage component which is generally in phase with the initial oscillator voltage. Conversion conductance is consequently increased.

Frequency and amplitude variations of the oscillator due to

signal circuit tuning changes are illustrated in Fig. 5.17. Assuming that $f_s < f_h$, tuning the signal circuit towards the oscillator frequency, i.e., reducing the signal circuit capacitance, causes the oscillator frequency to fall because the signal circuit appears as a capacitance and resistance in parallel (Section 4.2.2). The capacitance component reaches a maximum at a signal tuning setting of C_3, and then falls to zero when signal and oscillator circuits are tuned to the same frequency. Below this value of C_2 the signal circuit is inductive and the oscillator frequency increases, reaches a maximum at C_1, where $C_2 - C_1 \simeq C_3 - C_2$, and then falls as the signal capacitance is further decreased. Amplitude is a minimum at the signal tuning capacitance setting of C_2.

The reverse direction effect of capacitance coupling causes an oscillator voltage component in the signal circuit, which adds to

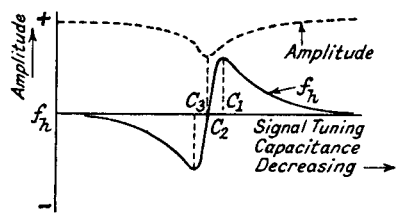


FIG. 5.17.—Curves showing the Effect of the Signal Tuning Circuit Capacitance on the Oscillator Frequency and Amplitude with Capacitance Coupling between the two Electrodes.

the true oscillator voltage when the signal circuit is capacitive, i.e., $f_s < f_h$, and opposes when the signal circuit is inductive, i.e., $f_s > f_h$. For grid-cathode application of the oscillator voltage the reverse is true, i.e., the oscillator voltage component in the signal circuit opposes the true oscillator voltage when the signal circuit is capacitive ($f_s < f_h$). A voltage in the grid circuit in phase with that in the cathode circuit means that the net voltage from grid to cathode is reduced. The magnitude of the oscillator component voltage depends on the value of intermediate and oscillator frequencies, the interelectrode capacitance, and to a less extent the Q of the signal circuit. As an example let us assume the following signal circuit constants $L_s = 2\ \mu\text{H}$, $C_s = 125\ \mu\mu\text{F}$, $f_s = 10$ Mc/s, $Q = 100$, $f_1 = 465$ kc/s, $C_{sig.\text{-}osc.\ elec.} = 0{\cdot}2\ \mu\mu\text{F}$. Using expression 4.8*a* (Section 4.2.2), the impedance of the signal circuit at the oscillator frequency 10·465 Mc/s is

$$Z = \frac{R_D}{1 + jQ\left(\frac{\omega_h}{\omega_s} - \frac{\omega_s}{\omega_h}\right)} = \frac{\omega_s L_s Q}{1 + 9{\cdot}1j} = \frac{12{,}560}{1 + 9{\cdot}1j}.$$

The reactance of the signal to oscillator interelectrode capacitance is

$$jX_{sig.\ osc.} = \frac{-j10^{12}}{6{\cdot}28 \times 10{\cdot}465 \times 10^6 \times 0{\cdot}2} = -76{,}000j$$

and the ratio of oscillator voltage across the signal-tuned circuit to oscillator electrode voltage is

$$\frac{12{,}560}{-76{,}000j(1+9{\cdot}1j)+12{,}560} = \frac{12{,}560}{712{,}000} = 0{\cdot}01762,$$

which is negligible. The undesirability of grid-cathode oscillator application under normal conditions can be demonstrated by assuming a grid-cathode capacitance of 3 $\mu\mu$F when the voltage ratio becomes 0·213.

Electron coupling and transit time give different results in the hexode valve from those in the heptode and they will be considered separately.

5.8.2. The Hexode as a Short Wave Frequency Changer. Since the signal grid of the hexode is next to the cathode, increase of operating frequency causes increased input admittance due to cathode lead inductance and electron transit time (Sections 2.8.3 and 2.8.6 show that input conductance is proportional to the square of the signal frequency.) The inclusion of a series resistance R_k in the cathode circuit decreases input capacitance and conductance and their variation with signal grid bias change. A resistance of from 20 to 50Ω may so reduce input conductance at high frequencies that the loss of amplification due to negative feedback (the equivalent conversion conductance is reduced to $\frac{g_c}{1+g_cZ_k}$) is more than offset, and overall amplification may actually be increased, as described in Section 4.10.3. Coupling from the signal to the oscillator circuit due to the common electron stream is not very important in the hexode, but in the reverse direction (from oscillator to signal) signal circuit performance is adversely affected by the repulsion of electrons from the neighbourhood of the oscillator grid back to the signal grid region. Some of these electrons (e.g., those repelled at the time when the oscillator voltage has its maximum negative value) may gain sufficient velocity to be collected by the signal grid. The effect occurs at all operating frequencies, but the occasions, when conditions favourable to the collection of electrons by the signal grid are realized, are multiplied as the oscillator frequency is increased. This electron collection causes grid current to flow in the signal circuit, and to prevent it the negative grid bias must be increased—about $-2{\cdot}5$ volts may be required at 20 Mc/s as

compared with − 1·0 volt at 1 Mc/s. This increase in minimum negative bias means a lower conversion conductance in the short and ultra short wave ranges. The collection of electrons by the signal grid is proportional [34] to the square of the amplitude of the oscillator voltage, and to the square of the distance between the first screen G_2 and the oscillator grid G_3, and inversely to the D.C. voltage on G_2. Hence excessive oscillator voltage must be avoided, and the distance between G_2 and G_3 should be as small as possible. Raising the screen voltage on G_2 has disadvantages since G_2 and G_4 are connected together, and increasing E_{g4} reduces the anode slope resistance, thus causing damping of the I.F. transformer.

5.8.3. The Heptode as a Short Wave Frequency Changer. A frequency changer having an oscillator in the same electron stream suffers from a number of disadvantages. A high oscillator mutual conductance is difficult to obtain, so that oscillation at the low frequency end of a short wave range requires tight coupling, which tends to produce a large oscillation amplitude and possibly squeggering at the high-frequency end. Another defect is that mentioned in Section 5.3.2, viz., the oscillator anode current is affected by bias on the signal grid and therefore application of A.G.C. bias varies oscillator amplitude and frequency. The second defect is largely eliminated by the constructions shown in Fig. 5.8*a* and 5.8*b*, or by using " electron coupling " [15] for the oscillator with a cathode coupling coil as in Fig. 5.18. The current in the cathode

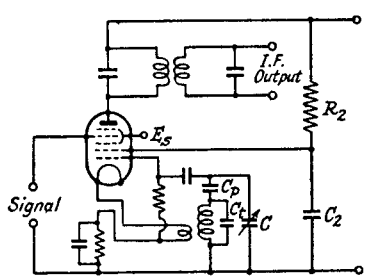


FIG. 5.18.—A Heptode Frequency Changer with Cathode Coupled Oscillator.

coil is independent of signal grid bias. Both defects are largely overcome by employing a separate oscillator valve.

The signal grid input admittance of the heptode valve may be positive or negative. It is often negative at low signal grid biases, i.e., it appears as a negative capacitance and negative resistance in parallel. This is due to the fact that there is no reservoir of electrons in the space between the two screens, G_3 and G_5 (Fig. 5.7), and the

space current tends to be constant and independent of signal grid bias variation (the shape of the I_aE_g characteristics of the signal grid are similar to those of the oscillator grid of the hexode as shown in Fig. 2.10). Let us consider an A.C. voltage applied to G_4; as this is increasing (the signal grid voltage is becoming less negative) electrons in the G_3G_4 space are accelerated and their velocity increased. Since the electron space current is the product of electron density and velocity, it follows that constant space current entails reduced electron density when velocity is increased. The reduced electron density means reduced negative charge in the neighbourhood of the signal grid, which is equivalent to reduced positive charge on the grid, i.e., the charge on the grid is 180° out of phase with the signal voltage producing it. If we assume that the signal voltage is

$$E_s = \hat{E} \sin \omega t,$$

the charge on the signal grid is

$$Q_g = -\hat{Q} \sin \omega t.$$

The current in the grid circuit due to this charge is the differential of Q_g with respect to time so that

$$I_g = \frac{dQ_g}{dt} = -\hat{I} \cos \omega t = \hat{I} \sin (\omega t - 90).$$

The phase relationships between E_s, Q_g and I_g are shown in Fig. 5.19*a*. I_g lags behind E_s by 90° and the input admittance

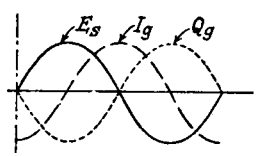


FIG. 5.19*a*.—The Phase Relationship of the Signal Voltage and the Grid Current due to Electron Motion.

appears as a negative capacitance, $-C_{g4}$ as shown by the vector diagram of Fig. 5.19*b*. Electron transit time causes a lag in the charge so that

$$Q_g = -\hat{Q} \sin (\omega t - \phi)$$

where ϕ = angle of lag due to the time of travel of electrons between G_3 and G_4 (the signal grid)

and $$I_g = \hat{I} \sin [\omega t - (90 + \phi)].$$

This modifies the vector diagram to that of Fig. 5.19*c*, which means an input admittance with a negative conductance ($-G_{g4}$) as well

as a negative capacitance component ($-C_{g4}$). The actual value of input admittance is determined by the voltage on the oscillator grid ; it is generally negative for low negative voltages on G_1, becoming smaller, and finally positive to a value equal to losses in the valve base and holder, as this bias is increased. Under normal conditions the signal grid input admittance varies considerably over a complete cycle of oscillator voltage and measurement gives an average value of input admittance. For television reception, the signal circuit

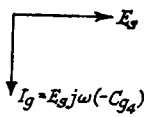


FIG. 5.19b.—The Vector Diagram for the Signal Voltage and Grid Current.

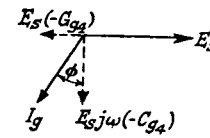


FIG. 5.19c —The Vector Diagram for the Signal Voltage and Grid Current showing the Lag due to Electron Transit Time.

connected to the frequency changer may require damping in order to prevent the negative input conductance from reducing to too great an extent the width of the pass-band. The average value of negative input conductance is generally decreased as the oscillator voltage amplitude is increased, because the latter increases the D.C. self-bias on the oscillator grid and also carries the frequency changer current into cut-off during some part of its negative half-cycle. Whilst the valve is cut off, the instantaneous input conductance is positive. The improvement to be obtained in this manner is limited and it has the disadvantage of increasing oscillator harmonic response.

It is interesting to note that the above description of the cause of negative input admittance can also be used as a basis for explaining positive input admittance when the signal grid is next to the cathode, as in the hexode. In this instance increasing signal voltage means increased density of electrons in the neighbourhood of the grid (the cathode forms a reservoir of electrons which can be drawn upon as required) and therefore increased positive charge on the grid, i.e., the signal voltage and induced charge due to the motion of the electrons are in phase and

$$E_s = \hat{E} \sin \omega t$$
$$Q_g = \hat{Q} \sin \omega t$$
$$\therefore\ I_g = \hat{I} \cos \omega t$$

so that I_g leads upon E_s by 90°, or the input admittance is equivalent to a positive capacitance. Similarly transit time introduces a lag

in I_g and a positive conductance component in the admittance. If the heater voltage is reduced so that current saturation is approached, the input admittance of the hexode signal grid can become negative, because the cathode electron reservoir disappears.

Electron coupling from the signal to oscillator grid has a tendency to start grid current at increasing negative biases on the oscillator grid as the operating frequency is increased, but electron repulsion from G_4 to G_1 is much less than in the hexode since the signal voltage is usually much less than the oscillator voltage, and in any case grid current is required on G_1 for self-bias purposes. Hence this form of electron coupling is unimportant. In the reverse direction, oscillator to signal electrode, a very undesirable form of coupling exists. It can be shown [33] by varying the signal circuit tuning; as this frequency, initially lower than that of the oscillator, approaches the latter, the anode current of the frequency changer decreases to a minimum, from which it rises to its normal value when $f_s = f_h$. From this point it increases to a maximum and then falls back, becoming asymptotic to its normal value as f_s is further increased. This can be explained if the oscillator voltage component induced in the signal circuit is out-of-phase with the true oscillator voltage for $f_s < f_h$ and in-phase for $f_s > f_h$. This is the reverse of what happens with capacitive coupling between the two circuits (Section 5.8.1, Fig. 5.17). Electron coupling must therefore be equivalent to a negative capacitance from G_1 to G_4, and it may be explained by considering the charge induced on the signal grid G_4 by electron movement due to the oscillator grid voltage. Increasing voltage (less negative) on the latter increases total current and the electron density in the G_3G_4 space. The increased negative space-charge induces an increased positive charge on G_4, i.e., the charge on G_4 is in-phase with the voltage on G_1. This is the reverse of positive capacitance coupling which, for increasing positive voltage on G_1, induces an increasing negative charge on G_4. The equivalent negative capacitance coupling, which has a value from -2 to $-0{\cdot}5$ $\mu\mu$F, can be neutralized by the addition of capacitance externally between G_1 and G_4. Neutralization is, however, only complete at one particular set of operating conditions, and change of the D.C. voltages on the electrodes, or oscillator voltage amplitude varies the electron coupling negative capacitance value.

The coupling current flowing from G_1 to G_4 lags behind the voltage applied to G_1 and the vector relationship is as shown in Fig. 5.19*b* when $E_s = E_{g1}$ and $I_g = I_{g1g4}$. Transit time of the electrons causes

a lag in the current vector and at high frequencies the vector diagram is similar to that of Fig. 5.19c ; the coupling admittance includes a negative conductance component and neutralization involves the use of a series RC circuit between G_1 and G_4. Another factor complicates the problem ; the negative conductance component increases as the operating frequency increases so that electron coupling is dependent on the oscillator frequency as well as on the electrode voltages. Only partial neutralization can therefore be achieved when operating over a range of frequencies.

Owing to the negative capacitance coupling the oscillator frequency is affected by signal tuning, but in the reverse direction to that due to positive capacitance coupling. A curve of oscillator frequency variation against signal-tuning capacitance shows the change in f_h to be positive when $f_s < f_h$ and negative when $f_s > f_h$; the curve is the reverse of Fig. 5.17.

5.9. Image Signal Suppression Circuits.[11, 26]

5.9.1. Introduction. Section 5.1.3. indicates the necessity for discrimination against the image signal (f_h+f_1) in the R.F. circuits preceding the frequency changer, and a special filter is sometimes

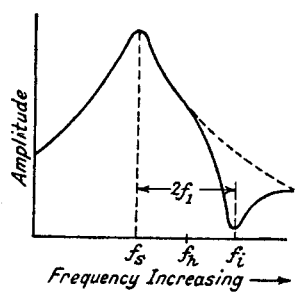


FIG. 5.20.—The Required Selectivity Curve for the R.F. Circuits preceding the Frequency Changer.

included to attenuate particularly this frequency. The effect of the filter is to give a selectivity curve with a pronounced dip at the image frequency (Fig. 5.20). There are two important methods of achieving image suppression, (1) by the use of series or parallel circuits tuned to the image signal and (2) by feedback of the image frequency component from a later stage into the input so as to neutralize the input component.

5.9.2. Series and Parallel Suppression Circuits. The simplest form of series circuit is obtained by tapping the output circuit down the coil of the parallel tuned signal circuit as shown

in Fig. 5.21. The top part of the coil and the tuning capacitor form the series resonant circuit rejecting the image frequency. The particular case shown is that of the aerial circuit and for the output voltage we have, neglecting resistance components

$$E_{AB} = I\left[j\omega L_2 + \frac{1}{j\omega C}\right] + Ij\omega M$$

and for E_{AB} to be zero

$$f = \frac{1}{2\pi\sqrt{(L_2+M)C}} \quad . \quad . \quad . \quad 5.17a.$$

If the input generator is a tetrode valve instead of an aerial, the constant current generator conception (Section 2.7) gives L_1 in

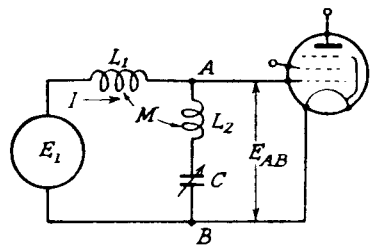


FIG. 5.21.—The Equivalent Circuit for the Tapped Coil Image Suppressor.

parallel with L_2C, and as L_2C are a short circuit at their resonant frequency there can be no current in L_1, so that

$$f = \frac{1}{2\pi\sqrt{L_2C}} \quad . \quad . \quad . \quad . \quad 5.17b.$$

is the condition for E_{AB} to equal 0.

Generally $M \ll L_2$ and 5.17*a* and 5.17*b* are almost identical.

This method suffers from two disadvantages; it reduces the amplitude of the desired signal, and it can only give maximum image suppression at one particular signal tuning frequency since variation of C over the tuning range makes it impossible to maintain the rejection frequency equal to the image frequency. We will illustrate this point for the aerial generator case. Let us assume that currents I_s and I_i flow in the circuit at the signal and image frequencies respectively. If tapping down is not employed, the image discrimination is given by the ratio of signal to image voltages across C and this is

$$\frac{E_{Cs}}{E_{Ci}} = \frac{I_s}{j\omega_s C} \cdot \frac{j\omega_i C}{I_i} = \frac{I_s}{I_i} \cdot \frac{f_i}{f_s} \quad . \quad . \quad . \quad 5.18$$

where $f_s = \dfrac{1}{2\pi\sqrt{(L_1+L_2+2M)C}}$, $f_i = f_s + 2f_1$, and f_1 = intermediate

frequency. Across the points AB the signal to image discrimination is $\frac{E_{ABs}}{E_{ABi}}$,

but $$E_{ABs} = I_s\left[R_{2s} + j\left(\omega_s(L_2+M) - \frac{1}{\omega_s C}\right)\right] \quad . \quad 5.19a$$

where R_{2s} = resistance of the L_2 part of the coil at f_s

$$E_{ABs} = I_s R_{2s}\left[1 + \frac{j\omega_s(L_2+M)}{R_{2s}}\left(1 - \frac{1}{\omega_s^2(L_2+M)C}\right)\right]$$

$$= I_s R_{2s}\left[1 + jQ_{2s}\left(1 - \frac{f_2^2}{f_s^2}\right)\right] \quad . \quad . \quad . \quad . \quad 5.19b$$

where $f_2 = \frac{1}{2\pi\sqrt{(L_2+M)C}} = \frac{1}{2\pi\sqrt{(L_1+L_2+2M)C}} \cdot \sqrt{\frac{L_1+L_2+2M}{L_2+M}}$

$= kf_s$

where $k = \sqrt{\frac{L_1+L_2+2M}{L_2+M}}$.

If we assume the total resistance R_1+R_2 to be proportionally divided between L_1 and L_2

$$Q_{2s} = \frac{\omega_s(L_2+M)}{R_{2s}} = \frac{\omega_s(L_2+L_1+2M)}{R_{1s}+R_{2s}} = Q_s.$$

$$\therefore\ E_{ABs} = I_s R_{2s}\left[1 + jQ_s\left(1 - \frac{f_2^2}{f_s^2}\right)\right]$$

and $$E_{ABi} = I_i R_{2i}\left[1 + jQ_i\left(1 - \frac{f_2^2}{f_i^2}\right)\right]$$

$$\therefore\ \frac{E_{ABs}}{E_{ABi}} = \frac{I_s R_{2s}\left[1 + jQ_s\left(1 - \frac{f_2^2}{f_s^2}\right)\right]}{I_i R_{2i}\left[1 + jQ_i\left(1 - \frac{f_2^2}{f_i^2}\right)\right]} \quad . \quad . \quad . \quad 5.20.$$

The improvement in rejection due to tapping down is

$$20\ \log_{10}\frac{|E_{ABs}|}{|E_{ABi}|} \cdot \frac{|E_{Ci}|}{|E_{Cs}|}\ \text{db.}$$

$$= 10\ \log_{10}\left[\frac{R_{2s} \cdot f_s}{R_{2i} f_i}\right]^2 \cdot \left[\frac{1 + Q_s^2\left(1 - \frac{f_2^2}{f_s^2}\right)^2}{1 + Q_i^2\left(1 - \frac{f_2^2}{f_i^2}\right)^2}\right] \quad . \quad 5.21.$$

If there is only a small variation in Q from f_s to f_i $\frac{R_{2s}}{R_{2i}} = \frac{f_s}{f_i}$ and 5.21 reduces to

$$10\ \log_{10}\frac{f_s^4 + Q^2(f_s^2 - f_2^2)^2}{f_i^4 + Q^2(f_i^2 - f_2^2)^2} \quad . \quad . \quad . \quad 5.22.$$

Curves in Fig. 5.22 show the increased image attenuation achieved by the tapped coil in comparison with the full coil over the medium wave-band of a receiver. The I.F. is 465 kc/s, the coil inductance 156 μH and $Q = 100$. Curve 1 gives $f_2 = f_i = 1{,}530$ kc/s at $f_s = 600$ kc/s, and curve 2 gives $f_2 = f_i = 1{,}930$ kc/s at $f_s = 1{,}000$ kc/s. Curve 1 is more satisfactory because attenuation is greatest for image frequencies either in the medium wave range or close to it. Curve 2 actually has less discrimination for these same image frequencies and gives greatest attenuation on frequencies well outside the medium wave range. If the optimum suppression

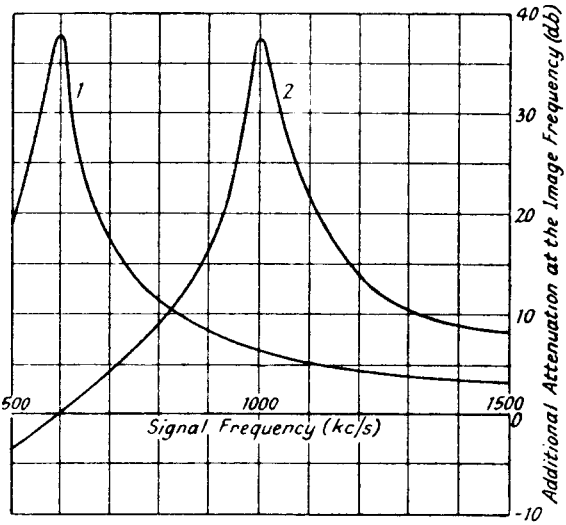


FIG. 5.22.—Image Frequency Rejection Curves for the Tapped Coil Suppressor.

point is chosen at too high a tuning frequency, the tapped coil may give less discrimination against the image than the full coil as the tuning frequency approaches the low-frequency end of the range. This occurs when

$$\frac{f_s^4+Q^2(f_s^2-f_2^2)^2}{f_i^4+Q^2(f_i^2-f_2^2)^2} < 1$$

or

$$(f_i^2+f_s^2)\frac{1+Q^2}{Q^2} > 2f_2^2 \quad . \quad . \quad . \quad 5.23a.$$

By noting that $f_i = f_s+2f_1$, $f_2 = kf_s$ and $\frac{1+Q^2}{Q^2} \simeq 1$, expression 5.23*a* reduces to

$$f_s < \frac{2f_1}{\sqrt{2k^2-1}-1} \quad . \quad . \quad . \quad 5.23b.$$

If an R.F. amplifier precedes the frequency changer, a better average image rejection curve is obtained by tapping down the aerial input and anode coils of this valve by differing amounts so as to stagger the frequencies of optimum rejection.

Tapping down the coil has the disadvantage of reducing the gain at the signal frequency and the loss is

$$20 \log_{10} \frac{|E_{Cs}|}{|E_{ABs}|}$$

Since $\omega_s(L_1+L_2+2M) = \dfrac{1}{\omega_s C}$, expression 5.19*a* may be rewritten

$$E_{ABs} = I_s[R_{2s}+j\omega_s(L_1+M)]$$

also

$$E_{Cs} = \frac{I_s\omega_s(L_1+L_2+2M)}{j}.$$

Thus

$$20 \log_{10} \frac{|E_{Cs}|}{|E_{ABs}|} = 20 \log_{10} \frac{\omega_s(L_1+L_2+2M)}{\sqrt{R_{2s}^2+\omega_s^2(L_1+M)^2}} \quad . \quad 5.24a$$

$$= 20 \log_{10} \frac{L_1+L_2+2M}{L_1+M}$$

if $R_{2s} \ll \omega_s(L_1+M)$,

$$= 20 \log_{10} \frac{1}{1-\dfrac{L_2+M}{L_1+L_2+2M}}$$

$$= 20 \log_{10} \frac{1}{1-\dfrac{1}{k^2}} \quad . \quad . \quad . \quad . \quad 5.24b.$$

The loss for curves 1 and 2 (Fig. 5.22) is 1·4 and 2·72 dbs. respectively.

5.9.3. Image Suppression by Neutralizing Feedback Voltage.[26] Image suppression by feeding back part of the output voltage into the input, or vice versa, necessitates some form of selective circuit (accepting the desired and rejecting the image) between input and output. Let us suppose that equal amplitudes of desired and image frequencies are applied to the input of the band-pass filter of Fig. 5.23*a*.

In the second tuned circuit of the filter the signal voltage is increased and the image voltage decreased owing to the selectivity characteristic of the first tuned circuit. By coupling to this circuit a coil connected to the input, signal and image voltages may be injected into the second tuned circuit direct from the input. By correct adjustment of the amplitude and phase of the image voltage complete cancellation may be obtained. Some reduction of the

signal voltage must also occur, but this may be made quite small with coils of normal Q value. Generally no attempt is made to secure accurate antiphase conditions, and maximum image suppression is obtained by adjustment of the number of turns of the coupling coil. A single turn, wound on the earthed end of the second tuned coil, is often adequate.

By assuming the resistance components to be zero, and the coupling reactance between the two tuned circuits to be small in

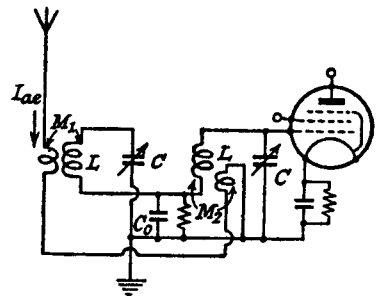


FIG. 5.23*a*.—Image Suppression by a Neutralizing Voltage applied from the Aerial to the Second Tuned Circuit.

comparison with the reactances in the tuned circuits at the image frequency, we have for the voltage transferred to the second circuit across the coupling capacitance C_0

$$E_{C_0} = \frac{I_{ae} j\omega_i M_1 \dfrac{1}{j\omega_i C_0}}{j\left[\omega_i L - \left(\dfrac{1}{\omega_i C} + \dfrac{1}{\omega_i C_0}\right)\right]} \quad . \quad . \quad . \quad 5.25$$

where I_{ae} = current in the aerial circuit
and M_1 = coupling from aerial to first tuned circuit.

The neutralizing voltage transferred to the second circuit by direct coupling from the second aerial coil is $I_{ae} j\omega_i M_2$ and for cancellation

$$I_{ae} j\omega_i M_2 = \frac{I_{ae} \dfrac{M_1}{C_0}}{j\left(\omega_i L - \dfrac{C+C_0}{\omega_i C C_0}\right)}$$

$$j\omega_i M_2 = \frac{\dfrac{M_1}{C_0}}{j\omega_i L\left(1 - \dfrac{C+C_0}{\omega_i^2 C C_0 L}\right)}$$

but
$$w_s^2 = \frac{C+C_0}{LCC_0} \quad . \quad . \quad . \quad . \quad 5.26$$

$$\therefore \omega_i^2 M_2 L\left(1 - \frac{\omega_s^2}{\omega_i^2}\right) = \frac{M_1}{C_0}$$

$$(\omega_i^2 - \omega_s^2) = \frac{M_1}{M_2 L C_0}.$$

Combining the above with 5.26

$$\frac{f_i^2}{f_s^2} - 1 = \frac{M_1 C}{M_2(C+C_0)} \quad . \quad . \quad . \quad 5.27.$$

Only one point of maximum suppression is obtained, but a second can be produced by tapping the output from a part of the second coil as described in Section 5.9.2.

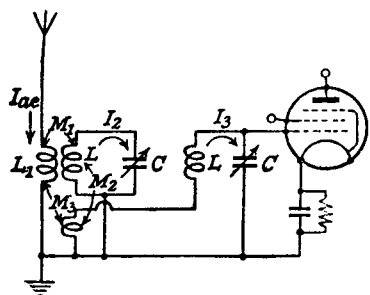


FIG. 5.23*b*.—An Alternative Circuit to Fig. 5.23*a* for Image Suppression by a Neutralizing Voltage.

An alternative circuit is that of Fig. 5.23*b*. The current and voltage equations are

$$E_{ae} = I_{ae}Z_1 + I_2 j\omega M_1 + I_3 j\omega M_3$$
$$0 = I_{ae} j\omega M_1 + I_2 Z_2 + I_3 j\omega M_2$$
$$0 = I_{ae} j\omega M_3 + I_2 j\omega M_2 + I_3 Z_3$$

where Z_1, Z_2 and Z_3 are the series impedances of each circuit, e.g., $Z_1 = R_{a0} + R_1 + j(X_{a0} + \omega L_1)$, where $R_{a0} + jX_{a0}$ is the aerial terminal impedance. If the resistance components are negligible $I_3 = 0$ at the optimum rejection frequency, thus

$$I_2 Z_2 = -I_{ae} j\omega_i M_1$$
$$I_2 j\omega_i M_2 = -I_{ae} j\omega_i M_3$$

from which
$$\frac{Z_2}{j\omega_i M_2} = \frac{M_1}{M_3} \quad . \quad . \quad . \quad . \quad 5.28.$$

But $$Z_2 = j\left(\omega_i L - \frac{1}{\omega_i C}\right) = j\omega_i L\left(1 - \frac{f_s^2}{f_i^2}\right)$$

where $$f_s = \frac{1}{2\pi\sqrt{LC}},$$ the desired frequency.

Replacing Z_2 in 5.28

$$\left[1 - \frac{f_s^2}{f_i^2}\right] = \frac{M_1 M_2}{M_3 L} \quad . \quad . \quad . \quad 5.29.$$

Again only one point of optimum suppression is obtained since the R.H.S. of the equation is a constant. The actual optimum suppression frequency may be set by sliding the coupling coil in the last circuit so as to vary M_3 and M_2 in opposite directions. Generally the aerial coil is wound on the earthed end of the first tuning coil,

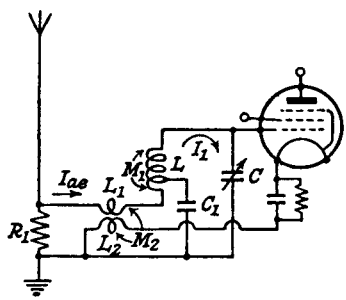


FIG. 5.23c.—Image Suppression by a Neutralizing Voltage Applied from the Aerial to the Cathode Circuit of the First R.F. Valve.

and the coupling coil in the last tuned circuit is arranged to slide over it.

A circuit permitting two points of optimum suppression in the wave range is shown in Fig. 5.23c. The neutralizing voltage is inserted in the cathode-earth lead of the first R.F. valve by a coil coupled to the aerial circuit, which is broadly tuned to the centre of the wave range to be accepted. By neglecting the resistance components, the rejection frequency may be calculated as follows: the equation for the tuned circuit is

$$(I_{ae} - I_1)\frac{1}{j\omega_i C_1} = I_1\left(j\omega_i L + \frac{1}{j\omega_i C}\right) + I_{ae} j\omega_i M_1$$

$$I_1 j\omega_i L\left(1 - \frac{f_s^2}{f_i^2}\right) = -I_{ae} j\left(\omega_i M_1 + \frac{1}{\omega_i C_1}\right)$$

where $$f_s = \frac{1}{2\pi\sqrt{\frac{LCC_1}{C + C_1}}}$$

C_1 is a comparatively large capacitance (about 0·005 μF) and its reactance is low enough to allow the impedance on the aerial side in parallel with it to be neglected, i.e., the tuned circuit frequency is determined by LC and C_1.

Voltage between grid and cathode $= E_C - E_{L_2}$

$$= \frac{I_1}{j\omega_i C} - I_{ae} j\omega_i M_2$$

$$= -I_{ae}\left[\frac{\left(\omega_i M_1 + \frac{1}{\omega_i C_1}\right)}{\omega_i L\left(1 - \frac{f_s^2}{f_i^2}\right)j\omega_i C} + j\omega_i M_2\right]$$

$$= 0$$

when $$\omega_i^2 LC\left(1 - \frac{f_s^2}{f_i^2}\right) = \frac{M_1}{M_2} + \frac{1}{\omega_i^2 M_2 C_1}.$$

But $$LC = \frac{C_1 + C}{\omega_s^2 C_1}$$

$$\therefore \left[\frac{f_i^2}{f_s^2} - 1\right]\frac{C_1 + C}{C_1} = \frac{M_1}{M_2} + \frac{f_3^2}{f_i^2} \quad . \quad . \quad 5.30$$

where $$f_3 = \frac{1}{2\pi\sqrt{M_2 C_1}}.$$

Two points of optimum suppression may be achieved over the wave range because the R.H.S. of equation 5.30 contains two independently variable terms. The first term is of importance at high frequencies. The circuit can give a fairly constant suppression over the wave range and has the advantage of introducing little damping from the aerial into the tuned circuit. The grid-cathode capacitance should be as low as possible as it modifies considerably the image suppression characteristics.

In all image suppression circuits the maximum reduction varies between 20 and 30 dbs.

5.9.4. Image Suppression on the Short Wave Range. The need for image signal suppression on the short wave range is even greater than on the long and medium wave ranges because the signal-tuned circuits are less selective, and the image-to-real signal-frequency ratio is less. For example, at a signal frequency of 1 Mc/s and an I.F. of 465 kc/s, the image signal is reduced by 43 dbs. (see expression 4.8*a*) for a single R.F. tuned circuit of $Q = 100$ (an average value), whereas for a signal frequency of 15 Mc/s and $Q = 50$ (an average value) the image signal is only reduced by 15.8 dbs. Methods of image signal rejection outlined in 5.9.2 and

5.9.3. are not successful for the short wave band because the frequency ratio between the image and real signal is so much smaller. The rejection of the image signal (10·930 Mc/s) for a signal frequency of 10 Mc/s is equivalent to the rejection of an image spaced 93 kc/s from a signal at 1 Mc/s. There have been no special developments for image suppression when a receiver is tuned over the whole of the short wave range, but two methods have been successfully applied with band-spread receivers having preset signal tuning.

One method [37] is illustrated in Fig. 5.24*a*. The preset signal circuit consists of a transformer, the primary and secondary of which are tuned to the centre of the band to be received. The aerial circuit is tuned by the series capacitance C_1 and the inductance L_1, and the latter is coupled to the secondary coil L_2 by combined

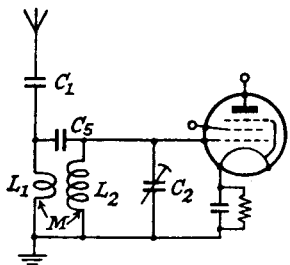


FIG. 5.24*a*.—Image Suppression on the Short Wave Band by Combined Mutual Inductance and Series Capacitance Coupling.

positive mutual inductance and series capacitance coupling (C_5 in Fig. 5 24*a*). The capacitance C_2 tunes the secondary circuit. Section 3.4.10 shows that the equivalent shunt coupling for this form of coupling is

$$Z_C = -\left[\frac{R_2\omega(L_1-M)+R_1\omega(L_2-M)}{B_1}+\frac{\omega^2(L_1-M)(L_2-M)(R_1+R_2)}{B_1^2}\right] + j\left[\omega M - \frac{\omega^2(L_1-M)(L_2-M)}{B_1}\right]$$

and the rejection frequency is obtained when the reactive component of Z_C is 0, i.e.,

when $$\omega M - \frac{\omega^2(L_1-M)(L_2-M)}{B_1} = 0.$$

Replacing B_1 by $\frac{1}{\omega C_5} - \omega(L_1+L_2-2M)$

$$\omega M\left[\frac{1}{\omega C_5} - \omega(L_1+L_2-2M)\right] = \omega^2(L_1-M)(L_2-M)$$

or
$$\omega = \frac{1}{\sqrt{\frac{C_5(L_1L_2 - M^2)}{M}}} \quad . \quad . \quad . \quad 5.31a$$
$$= \frac{1}{\sqrt{MC_5\left(\frac{1-k^2}{k^2}\right)}}. \quad . \quad . \quad . \quad 5.31b$$

where
$$k = \frac{M}{\sqrt{L_1L_2}}.$$

The method of adjusting the image rejection is as follows: With C_5 disconnected, C_1 and C_2 are tuned to give maximum signal frequency output at the centre of the band-spread range, the modulated input signal voltage being obtained from a signal generator through a standard dummy aerial. C_5 is next inserted between L_1 and L_2, so connected to give positive mutual inductance

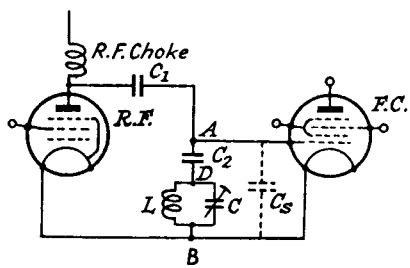


FIG. 5.24b.—Image Suppression on the Short Wave Band by a Series Parallel Suppression Circuit.

(see Section 3.4.2), and the signal generator is tuned to the image frequency (f_s+2f_1). C_5 is adjusted to give minimum audio output from the receiver. An average value for the extra image rejection is about 12 dbs.

Another circuit [40] for obtaining image suppression is shown in Fig. 5.24b. C_1 is the normal coupling capacitance (0·0001 μF) from the anode of the R.F. valve to the frequency changer. A capacitance C_2 is inserted between the top of the preset tuned-signal circuit and the grid of the frequency changer. It forms a series resonant circuit, at some frequency lower than the signal frequency, with the inductive reactance of the LC circuit, which itself is tuned to the signal frequency. In this instance it means that the image frequency must be less than the signal frequency, and the oscillator, contrary to normal practice, has a frequency of $f_s - f_1$. This causes no particular complication as the signal circuits are preset and no ganging problems therefore arise. The use of an inductance in place of C_2 would enable the more normal oscillator condition

of $f_h = f_s + f_1$ to be realized, but satisfactory rejection is difficult to obtain owing to resistance in the inductance, and stray capacitance from the grid of the frequency changer to earth. The series impedance of the LC circuit is (expression 4.4)

$$Z_{DB} = \frac{R}{(1-\omega^2LC)^2+(\omega CR)^2} + j\frac{\omega L(1-\omega^2LC) - \omega CR^2}{(1-\omega^2LC)^2+(\omega CR)^2}$$

$$Z_{DB} \simeq \frac{R}{(1-\omega^2LC)^2} + j\frac{\omega L}{(1-\omega^2LC)}$$

since R is usually much less than ωL. For series resonance at the image frequency

$$\frac{\omega_i L}{1-\omega_i^2 LC} = \frac{1}{\omega_i C_2}$$

$$\omega_i = \frac{1}{\sqrt{L(C+C_2)}} = \frac{\omega_s}{\sqrt{1+\frac{C_2}{C}}} \quad . \quad . \quad . \quad 5.32$$

and then $$Z_{AB} = \frac{R}{(1-\omega_i^2LC)^2}$$

$$= \frac{R}{\left[1-\frac{\omega_i^2}{\omega_s^2}\right]^2}$$

where $$\omega_s^2 = \frac{1}{LC}.$$

In the absence of the capacitance C_2

$$Z_{AB(c_2=0)} = Z_{DB} = \frac{\frac{L}{CR}}{1+jQ_s\left(\frac{\omega_i}{\omega_s} - \frac{\omega_s}{\omega_i}\right)} \qquad \text{[Expression 4.8}a\text{]}$$

so that the increase in image rejection is

$$20\log_{10}\frac{Z_{DB}}{Z_{AB}} = 20\log_{10}\frac{L}{CR^2}\frac{\left[1-\left(\frac{f_i}{f_s}\right)^2\right]^2}{\sqrt{1+Q_s^2\left(\frac{f_i}{f_s}-\frac{f_s}{f_i}\right)^2}} \text{ db.}$$

But $$\frac{L}{CR^2} = \frac{\omega_s L}{R}\cdot\frac{1}{\omega_s CR} = Q_s^2$$

and increased image rejection $= 20\log_{10}\dfrac{Q_s^2\left[1-\left(\frac{f_i}{f_s}\right)^2\right]^2}{\sqrt{1+Q_s^2\left[\frac{f_i}{f_s}-\frac{f_s}{f_i}\right]^2}}$. 5.33.

For $f_s = 15$ Mc/s, $Q = 50$ and $f_1 = 465$ kc/s

$$\text{increased image rejection} = 20 \log_{10} \frac{50^2\left[1 - \left(\frac{14{\cdot}07}{15}\right)^2\right]^2}{\sqrt{1+50^2\left[\frac{14{\cdot}07}{15} - \frac{15}{14{\cdot}07}\right]^2}}$$

$$= 15{\cdot}34 \text{ dbs.}$$

The ratio of $\frac{C_2}{C}$ for this signal frequency is given from expression 5.32.

$$\frac{C_2}{C} = \left(\frac{f_s}{f_1}\right)^2 - 1 = 0{\cdot}135$$

or $C_2 = 10{\cdot}8\ \mu\mu\text{F}$ if $C = 80\ \mu\mu\text{F}$. (An average value for band-spread signal tuning.)

The above represents the maximum image suppression for a signal tuned to the centre of the band-spread range and the average rejection over the whole range would be about 13 dbs. Higher values of image suppression are obtained at lower frequencies due to the increased ratio of $\frac{f_s}{f_i}$; thus at 6 Mc/s for the same Q the maximum extra image suppression is increased to 21·6 dbs. Expression 5.32 shows that the ratio $\frac{C_2}{C}$ must be changed when the range is varied, and it has to be increased as the central signal frequency is reduced. The required increase in C_2 is less when L is switched and C is fixed than when C is switched. Usually the first condition applies for reasons discussed in Section 4.10.2.

Stray capacitance C_s across the points AB should be reduced to the smallest possible value as it affects adversely the image rejection performance, chiefly by reducing the impedance to the signal frequency across the points AB. Thus, denoting the dynamic impedance of the LC circuit at the signal frequency by R_D,

$$Z_{AB}(C_s = 0) = R_D + \frac{1}{j\omega C_2}$$

$$Z_{AB}(C_s \neq 0) = \frac{\left[R_D + \frac{1}{j\omega C_2}\right]\frac{1}{j\omega C_s}}{R_D + \frac{1}{j\omega C_2} + \frac{1}{j\omega C_s}}$$

$$\simeq \left(R_D + \frac{1}{j\omega C_2}\right)\frac{C_2}{C_2 + C_s}$$

when $\frac{1}{\omega C_2} + \frac{1}{\omega C_s} \gg R_D$

$$\therefore\ Z_{AB} \simeq \frac{C_2}{C_2 + C_s} Z_{AB}\cdot(C_s = 0). \quad . \quad . \quad 5.34.$$

5.10. Push-Pull Frequency Changing. Push-pull frequency changing possesses certain advantages over single valve frequency changing.

(1) The signal or oscillator frequency currents in the anode circuit can be cancelled. This is of most advantage when the signal or oscillator frequency approaches that of the I.F.

(2) The interaction between the signal and oscillator circuits due to electrode capacitance and electron coupling [39] may be considerably reduced.

(3) All even oscillator harmonic responses may be cancelled.

If the oscillator grids of the two valves are connected in parallel and the anode circuits are in push-pull, matched valves give no resultant oscillator current in the anode circuit. Even oscillator harmonic responses are therefore cancelled ; this may be proved by assuming the I_aE_g relationship of each valve to be represented by

$$I_a = (a_0+a_1E_{g1}+a_2E_{g1}^2+\ \ldots)(b_0+b_1E_{g3}+b_2E_{g3}^2+\ \ldots)$$

if

$$E_{g1} = \hat{E}_s \cos \omega_s t$$

$$E_{g3} = \hat{E}_h \cos w_h t$$

(the bias voltage for simplification is assumed to be zero)

the coefficient of I.F. current in the first valve $= \dfrac{\hat{E}_h\hat{E}_s a_1 b_1}{2}$

,, ,, ,, ,, ,, ,, ,, second ,, $= -\dfrac{\hat{E}_h\hat{E}_s a_1 b_1}{2}$.

Since the valves are in push-pull with regard to their anode circuits we must subtract the two currents, hence the total I.F. current $= \hat{E}_s\hat{E}_h a_1 b_1$.

The coefficient of the 2nd harmonic response current in the first valve

$$= \frac{\hat{E}_h^2\hat{E}_s a_1 b_2^2}{2}.$$

The coefficient of the 2nd harmonic response current in the second valve

$$= \frac{(-\hat{E}_h)^2\hat{E}_s a_1 b_2^2}{2}$$

$$= \frac{\hat{E}_h^2\hat{E}_s a_1 b_2^2}{2}.$$

Subtraction of these terms gives zero coefficient, and this will be found to occur for all even harmonic responses. If the valves are not matched, $a_1' \neq a_1''$ and $b_2' \neq b_2''$ and some oscillator harmonic response is found.

BIBLIOGRAPHY

1. A New System of Short Wave Amplification. E. H. Armstrong, *Proc. I.R.E.*, Feb. 1921, p. 3.
2. The Superheterodyne, Its Origin, Development and Some Recent Improvements. E. H. Armstrong, *Proc. I.R.E.*, Oct. 1924, p. 539.
3. On the Origin of the Superheterodyne Method. W. Schottky, *Proc. I.R.E.*, Oct. 1926, p. 695.
4. Recent Developments in Superheterodyne Reception. G. L. Beers and W. L. Carlson, *Proc. I.R.E.*, March 1929, p. 501.
5. Considerations in Superheterodyne Design. E. G. Watts, *Proc. I.R.E.*, April 1930, p. 690.
6. Undesired Responses in Superheterodynes. R. H. Langley, *Electronics*, May 1931, p. 618.
7. A Single Dial Control Short Wave Converter for Working a Broadcast Receiver as a Short Wave Superheterodyne. H. A. Chinn, *Q.S.T.*, June 1931, p. 9.
8. Single Valve Frequency Changers. W. T. Cocking, *Wireless World*, July 29th (p. 74), Aug. 5th (p. 110), 1932.
9. The Screen-Grid Valve as Frequency-Changer in the Superhet. E. L. C. White, *Wireless Engineer*, Nov. 1932, p. 618.
10. The Hexode Tube. H. A. Wheeler, *Radio Engineering*, March (p. 19) and April (p. 12), 1933.
11. Second Channel Suppression. R. I. Kinross, *Wireless World*, June 9th, 1933, p. 416.
12. A Note on Interference Tones in Superheterodyne Receivers. W. F. Floyd, *Proc. Phys. Soc.*, July 1st, 1933, p. 610.
13. The Pentagrid Converter. C. L. Lyons, *Wireless Engineer*, July 1933, p. 364.
14. Superheterodyne Receivers. The Advantage of a High Intermediate Frequency. *World Radio*, Jan. 19th, 1934, p. 87.
15. Suppression of Interlocking in First Detector Circuits. P. W. Klipsh, *Proc. I.R.E.*, June 1934, p. 699.
16. On Conversion Detectors. M. J. O. Strutt, *Proc. I.R.E.*, Aug. 1934, p. 981.
17. Second Channel and Harmonic Reception in Superheterodynes. G. W. O. Howe, *Wireless Engineer*, Sept. 1934, p. 461.
18. Developing Single Span Tuning. W. T. Cocking, *Wireless World*, March 23rd (p. 196), Nov. 16th (p. 391), Nov. 23rd (p. 418), 1934.
19. Heptode Frequency Changers. R. J. Wey, *Wireless Engineer*, Dec. 1934, p. 642.
20. Mixing Valves. M. J. O. Strutt, *Wireless Engineer*, Feb. 1935, p. 59.
21. The Operation of Superheterodyne First Detector Valves. J. Stewart, *Journal I.E.E.*, Feb. 1935, p. 227.
22. Frequency Transformation with Reference to Single Span. F. M. Colebrook, *Wireless World*, Feb. 15th, 1935, p. 174.
23. An Improved Short Wave Frequency Changer. E. J. Alway, *Wireless World*, March 1st, 1935, p. 213.
24. Whistling Notes in Superheterodyne Receivers. M. J. O. Strutt, *Wireless Engineer*, April 1935, p. 194.

25. The Application of Superheterodyne Frequency Conversion Systems to Multirange Receivers. W. A. Harris, *Proc. I.R.E.*, April 1935, p. 279.
26. Image Suppression in Superheterodyne Receivers. H. A. Wheeler, *Proc. I.R.E.*, June 1935, p. 569.
27. Superheterodyne Whistles. M. G. Scroggie, *Wireless World*, Sept. 13th, 1935, p. 302.
28. Interfering Responses in Superheterodynes. H. K. Morgan, *Proc. I.R.E.*, Oct. 1935, p. 1164.
29. A New Tube for use in Superheterodyne Frequency Conversion. C. F. Nesslage, E. W. Herold., W. A. Harris, *Proc. I.R.E.*, Feb. 1936, p. 207.
30. Diode Frequency Changers. M. J. O. Strutt, *Wireless Engineer*, Feb. 1936, p. 73.
31. Frequency Changers in All Wave Receivers. M. J. O. Strutt, *Wireless Engineer*, April 1937, p. 184.
32. Why the Triode Hexode ? J. A. Szabadi, *Wireless World*, May 7th (p. 446) and May 14th (p. 472), 1937.
33. Frequency Changing with the Octode. E. Lukacs, *Wireless World*, March 17th, 1938, p. 238.
34. Electron Transit Time Effects in Multigrid Valves. M. J. O. Strutt, *Wireless Engineer*, June 1938, p. 315.
35. A New Converter Tube for All Wave Reception. E. W. Herold, W. A. Harris, T. J. Henry, *R.C.A. Review*, July 1938, p. 67.
36. A New Converter Valve. J. L. H. Jonker and A. J. W. M. Van Overbeek, *Wireless Engineer*, Aug. 1938, p. 423.
37. Pye 906. International A.C. Superhet. *Supplement to Wireless and Electrical Trader*, Oct. 7th, 1939.
38. Superheterodyne Converter Systems. E. W. Herold, *R.C.A. Review*, Jan. 1940, p. 325.
39. A New Ultra Short Wave Frequency Changer. J. A. Sargrove, *Wireless World*, May 1941, p. 124.
40. Making the Most of Short Waves. L. A. Moxon. *Wireless World*, June 1941, p. 148.
41. The Diode as Rectifier and Frequency Changer. D. A. Bell, *Wireless Engineer*, Oct. 1941, p. 395.
42. *Moderne Mehrgitter-Electronenröhren.* M. J. O. Strutt. Text-book.

CHAPTER 6

OSCILLATORS FOR SUPERHETERODYNE RECEPTION

6.1. Introduction.[1] Any reversible system which is capable of storing and releasing energy, i.e., changing its form from potential to kinetic and back again, can be made to oscillate. If the release and storage can be achieved without loss of energy the system will continue to oscillate when once set in motion. This condition cannot be realized in practice because energy is always dissipated when any interchange takes place, and the oscillation amplitude decays according to an exponential law unless the loss is made good. The simplest form of oscillator is the pendulum, which stores energy when rising and releases it when falling. The potential energy is a maximum at the top of the stroke, when the pendulum is momentarily stationary, and storage is complete. Kinetic energy is maximum at the centre of the stroke, and the potential energy is then zero. The pendulum has as an electrical counterpart, the LC circuit, for energy storage is possible in the capacitance and energy release in the inductance. A fully charged capacitance corresponds to maximum potential energy in the pendulum at the top of its stroke, whilst maximum current in the inductance corresponds to maximum kinetic energy in the pendulum at the centre of its stroke. The energy stored in the capacitance is $\frac{1}{2}CE^2$, and is a maximum when the voltage is maximum and current through the system is zero. The energy released in the inductance is $\frac{1}{2}LI^2$ and is zero when I is 0, i.e., when $\frac{1}{2}CE^2$ is maximum. It is maximum one-quarter of a period later (90°) when E is zero. The analogy with the pendulum is therefore exact, for the potential energy and kinetic energy maxima are separated by one-quarter of the oscillation period.

In the pendulum, energy is lost mainly in friction at the bearings, and this loss is made up by energy transferred through a suitable impulsing mechanism to the pendulum from the unidirectional driving source, a tensioned spring, a suspended weight, or an electrical power supply. This impulsing must be made in the right direction and at the correct instant so as to restore the decaying amplitude to its original value. Correct timing or phasing of the impulse is essential if the natural oscillation of the pendulum is to remain

unaltered, and it should be applied when the pendulum is at the centre of its stroke. An impulse given at any other position increases or decreases the frequency of oscillation. If the impulse is given on the down stroke in the direction of pendulum motion, the frequency is increased, whereas it is decreased if given in the opposite direction.

Energy loss in the tuned circuit occurs due to the resistance component of the inductance and of the capacitance. The former is generally far more important than the latter, which can be reduced to a very low value by the use of suitable dielectrics. Make-up energy to compensate for these losses can be derived from a D.C. source, and the impulsing action may be supplied mechanically by a buzzer or electrically by a valve. In the case of a buzzer the impulse is of short duration and at intervals separated by perhaps many cycles of the *LC* circuit oscillation frequency. The result is a train of oscillations of decaying amplitude built up to the initial amplitude at regular time intervals, when the buzzer contacts close and connect it to the D.C. power source. Conditions are different in the case of a valve, which is able to sustain oscillations by virtue of its amplifying action. An essential condition is that energy is fed back from the valve output circuit to the input circuit. This input energy is amplified in the valve and helps to make up the losses in the output circuit, which is usually directly or indirectly the *LC* circuit. The decay of an oscillation once started in the *LC* circuit is prolonged by the energy feedback and amplification, and if sufficient feedback occurs oscillations can be prolonged indefinitely. The make-up energy may be supplied to the *LC* circuit during a part or the whole of the oscillation cycle. Short period impulsing (Class C operation) over about a third of the cycle is the most efficient method of operation, and is consequently widely used for transmitting oscillators. For a superheterodyne receiver oscillator, Class C impulsing is not normally employed and Class B operation is more usual. The impulse is applied over about three-quarters of the cycle and this has the advantage of causing less harmonic production. Low harmonic content is essential to prevent whistle interference in the frequency changer (Sections 5.4.3 and 5.4.4), and oscillator efficiency is of little consequence.

An essential component of the valve oscillator is the feedback impedance between the anode and grid circuits, which allows the valve to apply the make-up energy in the correct phase relationship. If the phase angle of the feedback voltage is not correct the oscillation frequency is different from the natural frequency of the

LC circuit and the frequency is decreased or increased in the same manner as that of the pendulum.

In a valve-driven circuit oscillation occurs as soon as the ratio of the energy feedback to the output energy exceeds the energy gain of the valve, and oscillation amplitude builds up to larger and larger values until balance is restored by a reduction in the energy gain of the valve. The efficiency of the valve as an energy amplifier is only independent of oscillation amplitude for small amplitudes, and usually the energy gain progressively falls when the amplitude is increased beyond a certain value. The limit of oscillation amplitude is determined by cut-off of anode current at one end, and either by grid current damping or saturation of anode current at the other. Grid current is usually the more important limiting factor. The non-linear action of the valve is essential to the stabilisation of amplitude and it is clear that harmonics of the oscillation frequency cannot be avoided. The amplitudes of the individual harmonics in comparison to the fundamental can, however, be made very small by careful design and operation.

6.2. Types of Valve Oscillators and the Conditions for Self-Oscillation.

6.2.1. Introduction. There are five chief types of valve-maintained oscillators. In the tuned anode oscillator (shown in Fig. 6.2*a*), the *LC* circuit is connected to the valve anode, and a coil coupled to the inductance branch is used to return energy to the grid circuit for amplification by the valve. The tuned grid oscillator is similar to the tuned-anode oscillator of Fig. 6.2*a* except that the positions of the *LC* circuit and feedback impedance are reversed, i.e., the former is connected between grid and earth.

The simplest form of back-coupled oscillator is the Hartley (Fig. 6.4). The tuned *LC* circuit, connected between anode and grid, is its own feedback impedance. The position of the tapping point of the cathode on the coil governs the amount of feedback, and is such that the latter is in the right direction to increase the energy supply from the anode of the valve to the *LC* circuit.

The Colpitts oscillator of Fig. 6.5 is a variant on the Hartley, the only difference being that the capacitance branch is split to provide a tapping for the cathode instead of the inductance branch. The *LC* circuit is again its own feedback impedance.

In the last form of oscillator (Meissner) shown in Fig. 6.1 the *LC* circuit is not inserted in either grid or anode circuit. Coils L_a and L_g, coupled to the tuned circuit, supply energy to the latter.

We will now proceed to an examination of the conditions necessary to start oscillations in all the above types of oscillator except the Meissner. One assumption is made, viz., that the valve is

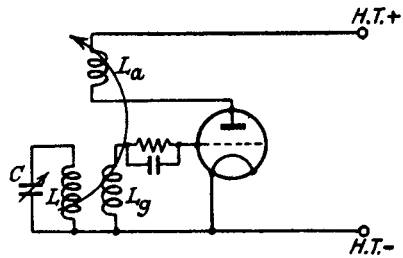


FIG. 6.1.—The Meissner Oscillator.

functioning as a linear amplifier. In actual fact, once oscillation has started, constant amplitude can only be maintained if the valve has non-linear characteristics.[10, 13] Nevertheless, the linear theory is most useful in giving an understanding of the principles underlying valve oscillators, and is helpful in indicating the effect of the valve on the frequency of oscillation.

6.2.2. The Tuned Anode Oscillator. The actual and equivalent circuits for the tuned anode oscillator are shown in Figs. 6.2*a* and 6.2*b*. Making the assumptions of linear [21] operation and high

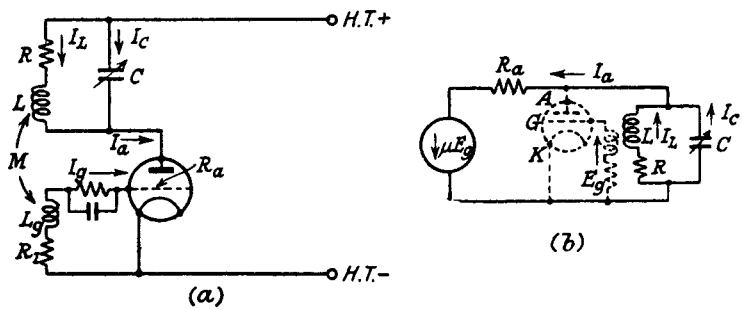


FIG. 6.2*a* and 2*b*.—The Tuned Anode Oscillator and its Equivalent Circuit.

grid input impedance, the current-voltage equation for the anode circuit is

$$\mu E_g = I_a R_a + I_L(R + j\omega L) \quad . \quad . \quad . \quad 6.1a$$

where $-E_g$ = voltage induced in the grid circuit, the negative sign must be included since the vector direction with reference to the cathode is opposite to that for the generated anode voltage, i.e.,

when μE_g is increasing in a positive direction, E_g is increasing in a negative direction. Hence the voltage equation for the grid circuit is

$$-E_g = j\omega M I_L \quad . \quad . \quad . \quad . \quad 6.2.$$

Combining 6.1*a* and 6.2.

$$-\mu j\omega M I_L = I_a R_a + I_L(R+j\omega L) \quad . \quad . \quad 6.1b$$

but

$$I_a = I_L + I_C \quad . \quad . \quad . \quad . \quad 6.3$$

and

$$I_L(R+j\omega L) = \frac{I_C}{j\omega C} \quad . \quad . \quad . \quad . \quad 6.4.$$

From 6.3 and 6.4

$$I_a = I_L[1+(R+j\omega L)j\omega C] \quad . \quad . \quad 6.5$$

replacing I_a in 6.1*b* by 6.5.

$$-\mu j\omega M I_L = I_L[R_a[(R+j\omega L).j\omega C+1)]+R+j\omega L]$$
$$I_L[R+R_a - \omega^2 LCR_a + j(\mu\omega M + \omega L + RR_a\omega C)] = 0 \quad 6.6.$$

Since $I_L = 0$ is obviously not a possible solution, the term in the bracket must be zero. Equating real and imaginary terms separately to zero we have

$$\omega = \frac{1}{\sqrt{LC}}\sqrt{1+\frac{R}{R_a}} = \omega_0\sqrt{1+\frac{R}{R_a}} \quad . \quad . \quad 6.7$$

where $\omega_0 = \dfrac{1}{\sqrt{LC}}$ the natural pulsance of the LC circuit

and

$$M = -\left[\frac{L+CRR_a}{\mu}\right] = -\left[\frac{L}{\mu}+\frac{CR}{g_m}\right] \quad . \quad . \quad 6.8.$$

The negative sign for M is to be expected since the grid voltage E_g with respect to the cathode must be in antiphase to μE_g also referred to the cathode. The physical significance of $-M$ is treated in Section 3.4.2, and the grid coil L_g is shown reversed in Fig. 6.2*a* to emphasise this feature. Two points of interest arise from expression 6.7 ; the valve slope resistance affects the oscillation frequency, making it higher than the natural frequency of the LC circuit, and the difference in frequency is least with smallest coil resistance and largest valve resistance. The increase of oscillation frequency means that the impedance of the LC tuned circuit is capacitive, and this is clearly shown from the vector diagram of Fig. 6.3, in which all voltage and current vectors are referred to the cathode. The generated voltage vector μE_g, shown as horizontal to the right, is taken as the datum. It has two components (see expression 6.1*a*), that represented by $I_a R_a$ across the valve resistance and that, $E_0 = I_L(R+j\omega L) = \dfrac{I_C}{j\omega C}$, across the tuned circuit. Since

the oscillation frequency exceeds the natural frequency of the tuned circuit, I_a is capacitive with respect to μE_g and leads it as shown. Of the two current components of I_a, I_C leads E_0 by 90° (assuming the capacitance C to have no resistance component) and I_L lags behind E_0 by less than 90° due to the coil resistance R. The grid voltage E_g lags behind I_L by 90° and is at 180° to μE_g; i.e., for continuous oscillation I_L must always lag behind μE_g by 90°. This is not possible if I_a is inductive, for E_0 then leads on μE_g, and as the angle between E_0 and I_L cannot exceed 90°, I_L can never be at 90° to μE_g. The effect of an increase of R_a is shown by the dashed vectors of Fig. 6.3, I_a is reduced to I_a', but the vector

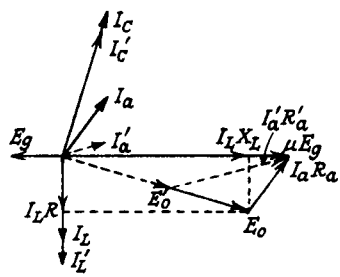


FIG. 6.3.—The Vector Diagram for the Tuned Anode Oscillator.

I_aR_a is increased to $I_a'R_a'$ because I_a' is reduced to a less extent than R_a' is increased. Since I_L', and therefore E_0', must remain at the same phase angle relative to μE_g it follows that $I_a'R_a'$ and therefore I_a' must have a smaller phase angle with respect to μE_g. I_a' must have a smaller angle of lead on E_0', which means that the LC circuit impedance is less capacitive and the oscillation frequency closer to the natural frequency of the tuned circuit thus confirming expression 6.7. The same process can also show that decrease of R reduces the difference between the oscillation and natural frequencies.

Let us now consider the effect of a finite grid input impedance due mainly to grid current. If I_g is the current in the grid circuit and R_g represents the grid input impedance (assumed to be resistive), the equations become

$$\mu E_g = -\mu I_g R_g = I_a R_a + I_L(R + j\omega L) + I_g j\omega M \quad . \quad 6.9$$

$$I_a = I_L + I_C \quad . \quad . \quad . \quad . \quad 6.10$$

$$I_L(R + j\omega L) + I_g j\omega M = \frac{I_C}{j\omega C} . \quad . \quad . \quad 6.11.$$

From 6.10 and 6.11

$$I_a = I_L[1+j\omega C(R+j\omega L)] - I_g\omega^2 MC \quad . \quad . \quad 6.12.$$

But $I_L j\omega M = I_g(R_g+j\omega L_g)$

$$I_g = \frac{I_L j\omega M}{R_g+j\omega L_g} \quad . \quad . \quad . \quad . \quad 6.13.$$

Replacing I_a and I_g in 6.9 by 6.12 and 6.13.

$$\frac{-\mu j I_L R_g \omega M}{R_g+j\omega L_g}$$

$$= I_L\left[R_a(1+j\omega C(R+j\omega L))+R+j\omega L+\frac{j\omega M}{R_g+j\omega L_g}(j\omega M-\omega^2 MCR_a)\right]$$

$$I_L[\mu j\omega M R_g+[R_a+j\omega C(R+j\omega L)R_a+R+j\omega L](R_g+j\omega L_g) + j\omega M(j\omega M - \omega^2 MCR_a)] = 0.$$

Equating the real term to zero gives the value of ω

$$R_g(R_a - \omega^2 LCR_a+R) - \omega^2 LL_g - \omega^2 L_g CRR_a - \omega^2 M^2 = 0 \quad . \quad 6.14.$$

Hence

$$\omega = \sqrt{\frac{R_g(R+R_a)}{LCR_aR_g+LL_g+L_gCRR_a+M^2}}$$

$$= \omega_0 \sqrt{\frac{1+\frac{R}{R_a}}{1+\frac{L_g}{CR_aR_g}+\frac{L_gR}{LR_g}+\frac{M^2}{LCR_aR_g}}} \quad 6.15.$$

It will be seen that expression 6.15 reduces to 6.7 when R_g is infinite. The frequency of oscillation is least affected when R_g is as large as possible, i.e., grid current should be small; L_g should also be as small as possible. The larger the value of C the less effect has R_g, so that we should expect the departure from the natural frequency to be less at the lower frequencies of a given tuning range.

6.2.3. The Tuned Grid Oscillator. If the effect of grid current is neglected the following equations are obtained for the tuned grid oscillator.

$$\mu E_g = I_a(R_a+j\omega L_a)+Ij\omega M \quad . \quad . \quad . \quad 6.16$$

where I = the circulating current in the tuned grid circuit

$$-E_g = \frac{I}{j\omega C} \quad . \quad . \quad . \quad . \quad 6.17$$

$$I_a j\omega M + I\left[R+j\left(\omega L - \frac{1}{\omega C}\right)\right] = 0 \quad . \quad . \quad 6.18.$$

Combining 6.16 and 6.17

$$I = \frac{I_a(R_a+j\omega L_a)}{-j\left(\omega M+\frac{\mu}{\omega C}\right)} \quad . \quad . \quad . \quad 6.19.$$

Replacing this value of I in 6.18

$$I_a\left[j\omega M+\frac{\left[R+j\left(\omega L-\frac{1}{\omega C}\right)\right](R_a+j\omega L_a)}{-j\left[\omega M+\frac{\mu}{\omega C}\right]}\right]=0$$

or

$$\omega^2M^2+\mu\frac{M}{C}+R_aR-\omega L_a\left(\omega L-\frac{1}{\omega C}\right)$$
$$+j\left[R_a\left(\omega L-\frac{1}{\omega C}\right)+R\omega L_a\right]=0 \quad . \quad 6.20.$$

Equating the imaginary term to zero gives the oscillation frequency as

$$\omega = \frac{1}{\sqrt{LC\left(1+\frac{L_aR}{LR_a}\right)}} \quad . \quad . \quad . \quad 6.21a$$

$$= \frac{\omega_0}{\sqrt{1+\frac{L_aR}{LR_a}}} \quad . \quad . \quad . \quad . \quad 6.21b.$$

This is less than ω_0 and the LC circuit is thus an inductive impedance at the oscillation frequency, exactly the reverse of the tuned anode oscillator. Least frequency variation is obtained when R_a is large, and R and L_a small.

Equating the real term to zero and replacing ω by expression 6.21a gives a quadratic in M

$$\frac{M^2}{LC(1+\alpha)}+\frac{\mu M}{C}+RR_a+\frac{L_a}{C}-\frac{L_a}{C(1+\alpha)}=0$$

where

$$\alpha=\frac{L_aR}{LR_a}$$

and

$$M=\frac{-\mu L(1+\alpha)\pm\sqrt{[\mu L(1+\alpha)]^2-4[L_aL\alpha+RR_aLC(1+\alpha)]}}{2} \quad . \quad 6.22a.$$

The minimum value of M required to start oscillation is that with the positive sign before the root.

$$\text{Thus } M = \frac{-\mu L(1+\alpha)+\mu L(1+\alpha)\left[1-\dfrac{4[L_aL\alpha+RR_aLC(1+\alpha)]}{[\mu L(1+\alpha)]^2}\right]^{\frac{1}{2}}}{2}$$

expanding the last term by the Binomial theorem.

$$M \simeq \frac{-[L_aL\alpha+RR_aLC(1+\alpha)]}{\mu L(1+\alpha)}$$

$$\simeq -\left[\frac{L_a\alpha}{\mu(1+\alpha)}+\frac{CR}{g_m}\right] \quad . \quad . \quad . \quad 6.22b.$$

Expression 6.22*b* is similar in form to 6.8 and M is again negative. The value of α is unlikely ever to be greater than 1 and L_a is usually much less than L, so that a smaller value of M is required to initiate oscillation for the tuned-grid oscillator than for the tuned-anode oscillator.

6.2.4. The Hartley Oscillator. The fundamental equations for the Hartley oscillator (Fig. 6.4), neglecting grid current, are as follows:

$$\mu E_g = I_aR_a+I_1(R_1+j\omega L_1) - I_2j\omega M.$$

In this equation and in all others that follow, M denotes the

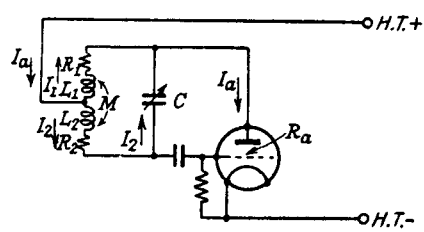


FIG. 6.4.—The Hartley Oscillator.

numerical value of mutual inductance and the negative sign, which the circuit clearly indicates is necessary, is placed outside M.

$$-E_g = I_2(R_2+j\omega L_2) - I_1j\omega M$$

$$I_a = I_1+I_2.$$

Hence

$$-\mu[I_2(R_2+j\omega L_2) - I_1j\omega M] = (I_1+I_2)R_a+I_1(R_1+j\omega L_1) - I_2j\omega M.$$

$$\text{But} \quad I_1(R_1+j\omega L_1) - I_2j\omega M = I_2(R_2+j\omega L_2) - I_1j\omega M+\frac{I_2}{j\omega C}$$

$$\therefore \; I_2[R_a+R_1+j\omega(L_1-\mu M)]\left[R_2+j\omega(L_2+M)+\frac{1}{j\omega C}\right]$$

$$+[R_a+\mu R_2+j\omega(\mu L_2-M)][R_1+j\omega(L_1+M)] = 0 \quad 6.23.$$

Equating the imaginary term to zero we have

$$R_2\omega(L_1-\mu M)+(R_a+R_1)\left[\omega(L_2+M)-\frac{1}{\omega C}\right]$$
$$+R_1\omega(\mu L_2-M)+(R_a+\mu R_2)\omega(L_1+M)=0$$

$$\omega^2[R_a(L_1+L_2+2M)+R_2(\mu+1)L_1+R_1(\mu+1)L_2]=\frac{R_a+R_1}{C}.$$

Generally $R_a(L_1+L_2+2M)\gg R_2(\mu+1)L_1$ and $R_1(\mu+1)L_2$ so that

$$\omega\simeq\sqrt{\frac{1+\frac{R_1}{R_a}}{(L_1+L_2+2M)C}}=\omega_0\sqrt{1+\frac{R_1}{R_a}} \quad . \quad . \quad 6.24$$

where ω_0 = the natural frequency of the LC circuit. The oscillation frequency, as in the case of the tuned anode oscillator, is higher than ω_0, but the frequency difference is small when R_a is large and R_1 small.

The condition for oscillation is obtained by equating the real part of expression 6.23 to zero.

$$(R_a+R_1)R_2-\omega^2(L_1-\mu M)(L_2+M)+\frac{L_1-\mu M}{C}$$
$$+(R_a+\mu R_2)R_1-\omega^2(\mu L_2-M)(L_1+M)=0$$

$$R_a(R_1+R_2)+R_1R_2(1+\mu)+\frac{L_1-\mu M}{C}-\omega^2(L_1L_2-M^2)(\mu+1)=0.$$

Neglecting $R_1R_2(1+\mu)$ and replacing ω^2 by 6.24

$$\frac{(L_1L_2-M^2)(\mu+1)\left(1+\frac{R_1}{R_a}\right)}{(L_1+L_2+2M)C}=\frac{L_1-\mu M}{C}+R_a(R_1+R_2).$$

If $\frac{R_1}{R_a}\ll 1$.

$$(L_1L_2-M^2)(\mu+1)=(L_1-\mu M)(L_1+L_2+2M)$$
$$+CR_a(R_1+R_2)(L_1+L_2+2M)$$

which reduces to

$$\mu=\frac{L_1+M}{L_2+M}+\frac{CR_a(R_1+R_2)(L_1+L_2+2M)}{(L_1+M)(L_2+M)}. \quad . \quad 6.25a$$

for continuous oscillation.

By changing the ratio of $\frac{L_1+M}{L_2+M}$, i.e., the cathode tapping point on the coil, it is possible to adjust the oscillating conditions. A larger inductance between the grid and cathode makes for easier

oscillation. The frequency requirement of high R_a points to the need for a high g_m for expression 6.25*a* approaches

$$\frac{\mu}{R_a} = g_m = \frac{C(R_1+R_2)(L_1+L_2+2M)}{(L_1+M)(L_2+M)} \quad . \quad . \quad 6.25b.$$

6.2.5. The Colpitts Oscillator. The conditions for start of oscillation may be determined in the same manner as for the previous

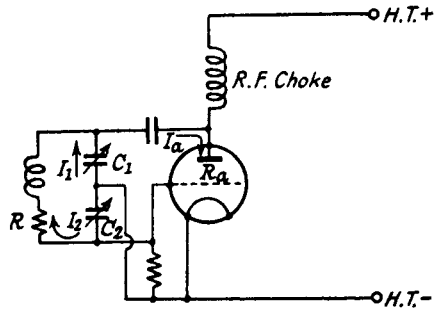


FIG. 6.5.—The Colpitts Oscillator.

circuits. Referring to Fig. 6.5, we find the following current and voltage relationships (neglecting grid current):

$$I_a = I_1+I_2$$

$$\mu E_g = I_aR_a+\frac{I_1}{j\omega C_1} = I_2R_a+I_1\left(R_a+\frac{1}{j\omega C_1}\right)$$

$$-E_g = \frac{I_2}{j\omega C_2}.$$

$$\frac{I_1}{j\omega C_1} = I_2\left(R+j\omega L+\frac{1}{j\omega C_2}\right)$$

$$-\frac{\mu I_2}{j\omega C_2} = I_2\left[\left(R+j\omega L+\frac{1}{j\omega C_2}\right)\left(R_a+\frac{1}{j\omega C_1}\right)j\omega C_1+R_a\right].$$

$$I_2[\mu+jR_a\omega C_2-\omega^2R_aRC_1C_2+j\omega C_2R-j\omega^3LR_aC_1C_2 - \omega^2LC_2+j\omega C_1R_a+1] = 0 \quad 6.26.$$

Equating the imaginary term to zero

$$R_a\omega C_2+\omega C_2R-\omega^3LR_aC_1C_2+\omega C_1R_a = 0.$$

$$\omega^2 = \frac{1}{LC_1}+\frac{1}{LC_2}+\frac{R}{R_aLC_1}$$

$$= \frac{1}{\frac{LC_1C_2}{C_1+C_2}}\left(1+\frac{RC_2}{R_a(C_1+C_2)}\right). \quad . \quad . \quad 6.27a$$

or

$$\omega = \omega_0\sqrt{1+\frac{R}{R_a}\left(\frac{C_2}{C_1+C_2}\right)}. \quad . \quad . \quad 6.27b$$

where ω_0 is the natural frequency of the LC circuit.

The condition for start of oscillation is obtained by equating the real term to zero.

I.e., $$\mu+1 = \omega^2(LC_2+RR_aC_1C_2)$$

Replacing ω by the value in 6.27*a*

$$\mu+1 = \left[\frac{C_1+C_2}{C_1}+\frac{R_aR(C_1+C_2)}{L}\right]\left[1+\frac{RC_2}{R_a(C_1+C_2)}\right]$$

If $$\frac{RC_2}{R_a(C_1+C_2)} \ll 1$$

$$\mu+1 = 1+\frac{C_2}{C_1}+\frac{R_aR(C_1+C_2)}{L}$$

or for oscillation to be sustained

$$\mu = \frac{C_2}{C_1}+\frac{R_aR(C_1+C_2)}{L} \quad . \quad . \quad . \quad 6.28a.$$

The ease with which a Colpitts circuit may be made to oscillate is very clearly indicated by expression 6.28*a*. In practice $C_2 \simeq C_1$, so that for normal values of circuit components even a very low μ valve can be made to sustain oscillation.

We also note that for small frequency variation due to the valve, R should be small and R_a large. When $\frac{R_aR}{L}(C_1+C_2) \gg \frac{C_2}{C_1}$ the valve parameter determining oscillation is g_m for expression 6.28*a* becomes

$$\frac{\mu}{R_a} = g_m = \frac{R(C_1+C_2)}{L} \quad . \quad . \quad . \quad 6.28b.$$

Summarizing the results of the examination assuming linear operation we find that for minimum frequency variation due to the valve, R_a and R_g (the grid circuit input resistance) should be as large as possible, whilst R, I_g, and the inductance of the feedback coil should be as small as possible. The valve makes the oscillation frequency greater than the resonant frequency of the LC circuit for the tuned anode, Hartley and Colpitts oscillators and less for the tuned-grid.

Ease of oscillation is secured in all cases by a low value of R and high value of g_m and oscillation conditions for the Hartley and Colpitts circuits are particularly easy to fulfil.

A lower value of M or μ is required as the tuning capacitance is decreased, i.e. frequency is increased.

6.3. The Conditions to be fulfilled by a Superheterodyne Receiver Oscillator. An oscillator supplying the oscillator voltage to a frequency changer must fulfil the following requirements if satisfactory operation is to be achieved:

1. The conditions for self-oscillation must be easily realizable. An oscillator requiring a feedback coil comparable in inductance to the *LC* circuit inductance is obviously unsatisfactory. Excessive anode current in the normal oscillating condition cannot be tolerated.

2. Oscillation must be maintained over the required frequency range without "blind spots", frequencies at which oscillation ceases.

3. The variation in amplitude over the frequency range should be as small as possible. The necessity for this is clearly indicated in Section 5.5.

4. The oscillator must not have a number of degrees of freedom, and must not be liable to parasitic or squegger oscillations.

5. Supply voltage variations should have minimum effect on the oscillation frequency.

6. The frequency should be independent of bias or supply voltage variations on the frequency changer.

7. Harmonic frequency voltages generated by the oscillator should be small.

8. Temperature and humidity variations should have minimum effect on the oscillation frequency.

Each of the four types of oscillator fulfils some of the requirements, and generally by careful design those not normally fulfilled can be approached.

The tuned anode oscillator, for example, has generally less harmonic content, a larger amplitude, and is more stable to supply voltage fluctuations than the tuned-grid, whilst the latter has a more constant amplitude as the oscillator frequency is varied over the tuning range.[2] The tuned anode oscillator has the added advantage of being less affected by bias variations on the frequency changer valve, when the oscillator grid of the frequency changer and grid of the oscillator valve are connected together.

The Hartley and Colpitts circuits, on the other hand, can be made to oscillate without difficulty even at high frequencies in the short wave range.

We will now consider requirement 3 in detail.

6.4. The Maintenance of Constant Output over the Frequency Range. In Section 5.5 it is shown that the oscillator voltage at the frequency changer must not be allowed to exceed a given value, otherwise harmonic interference whistles are produced by the frequency changer. Conversion conductance is rapidly reduced if the oscillator voltage falls much below the optimum

value. Hence it is essential to maintain the oscillator output voltage close to the required value over the frequency range.

The tendency in most oscillators is for the amplitude to increase as the frequency increases. The conditions for start of oscillation given in Section 6.2 indicate that this is to be expected, since decrease in the tuning capacitance C reduces the value of M or μ required to start oscillation. It is not practicable to reduce the value of M or μ as C is decreased. The decrease in C is offset to some extent by an increase of coil resistance R, but this is not sufficient to prevent increasing amplitude. The only alternative is to insert an impedance which increases the damping of the tuned circuit or reduces the feedback as the frequency rises. A normal method of achieving this is to include a resistance between the grid of the valve and the self-bias resistance and capacitance. This is

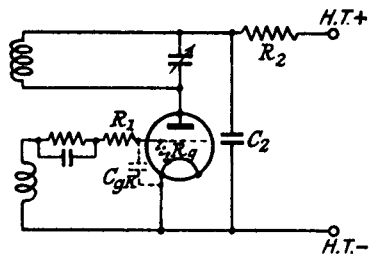


FIG. 6.6.—The Tuned Anode Oscillator with Series Grid Resistance as an Amplitude Stabilizer.

shown in Fig. 6.6 for the tuned-anode oscillator. The resistance R_1 forms with the grid input impedance a potentiometer which reduces the proportion of voltage transferred to the grid as the frequency increases. Its chief effect as an amplitude stabilizer is, however, in combination with the grid-earth capacitance for it produces across the grid coil a parallel damping resistance, which decreases as the frequency increases. For example, if $R_1 = 1{,}000\Omega$ and the grid earth capacitance $C_1 = 10\ \mu\mu$F, the admittance of these two in series is

$$Y = \frac{1}{R_1 + \dfrac{1}{j\omega C_1}} = \frac{j\omega C_1(1 - j\omega C_1 R_1)}{1 + \omega^2 C_1^2 R_1^2}$$

$$= \frac{R_1\omega^2 C_1^2}{1 + \omega^2 C_1^2 R_1^2} + \frac{j\omega C_1}{1 + \omega^2 C_1^2 R_1^2}$$

$$= +Gj B.$$

The damping resistance across the grid coil is therefore

$$R_g = \frac{1}{G} = \frac{1+\omega^2 C_1^2 R_1^2}{R_1 \omega^2 C_1^2}$$
$$= 0{\cdot}84 \text{ M}\Omega \text{ at } 550 \text{ kc/s}$$
$$= 0{\cdot}113 \text{ M}\Omega \text{ at } 1{,}500 \text{ kc/s}.$$

A typical curve showing the amplitude correction obtained by this method is given in Fig. 6.7. A value of 500 to 1,000Ω is suitable for long and medium waves and 50Ω for the short waves. R_2 and C_2 are R.F. decoupling components in the anode circuit and have values from 20,000 to 50,000Ω and 0·1 μF, respectively. Limitation of amplitude variation in the tuned-grid oscillator can also be

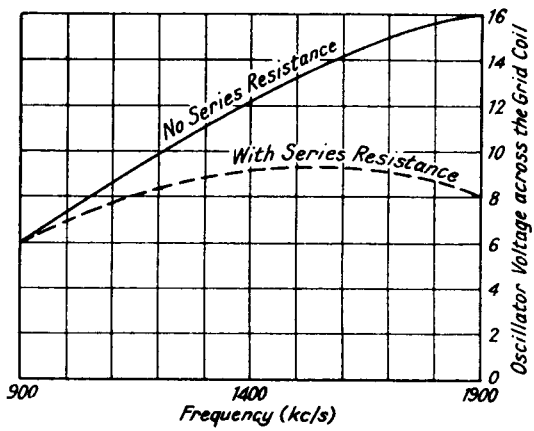


FIG. 6.7.—Curves showing the Stabilizing Effect of the Series Grid Resistance in the Tuned Anode Oscillator.

achieved by inserting a resistance between the anode and feedback coil, and the values given above are suitable.

Negative feedback, by inserting a small coil L_k in the cathode lead, can be applied to limit amplitude variation when the oscillator valve is separate from the frequency changer. The increasing impedance of L_k reduces the effective mutual conductance of the valve to $g_m' = \frac{g_m}{1+jg_m\omega L_k}$ as the frequency increases, and so reduces oscillator output.

Amplitude variation in a tuned anode oscillator may be reduced by using resistance-capacitance coupling to the LC circuit as in Fig. 6.8. The anode resistance R_a reduces the gain variation due to the increase in the tuned circuit impedance $\left(\frac{L}{CR}\right)$ as the oscillation

frequency increases. This resistance has the disadvantage of damping the LC circuit and increasing the possibility of harmonic production in the oscillator itself. It also reduces oscillator frequency stability with respect to changes of H.T. voltage, because R_0 is

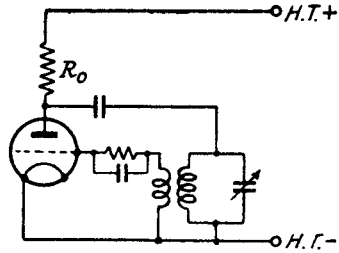


FIG. 6.8.—Amplitude Stabilization of the Tuned Anode Oscillator by a Shunt Resistance.

effectively in parallel with the tuned circuit; it increases R in expression 6.7 with the result that $\frac{R}{R_a}$ is increased and variations of R_a produced by changes of H.T. voltage have greater effect.

Amplitude stability may be increased for the tuned-anode oscillator by using a diode connected to the tuned circuit to supply the grid bias. A suitable circuit is shown in Fig. 6.9. This method considerably reduces the maximum amplitude, and a long grid-base

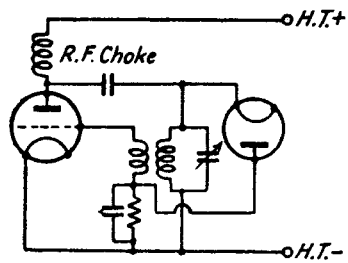


FIG. 6.9.—Amplitude Stabilization of the Tuned Anode Oscillator by Grid Bias derived from the Tuned Circuit with a Diode.

valve may be necessary to obtain sufficient voltage for the frequency changer.

6.5. Frequency Stability.[3, 8] One of the most important requirements in an oscillator is that its frequency should remain constant. Any variation of frequency from the correct setting causes the I.F. carrier to be displaced from the centre of the pass-band of the I.F. amplifier. For an amplitude modulated wave, frequency distortion, producing high-pitched shrill reproduction, is

the chief result, but if the off-centering is excessive the equivalent of single side-band transmission is obtained and amplitude distortion occurs. Rapid fluctuations of H.T. voltage, due to hum, interference, etc., are more serious in the frequency modulation receiver, since they cause frequency modulation of the oscillator rather than amplitude modulation. The detector system in the amplitude modulation receiver has no response to a frequency modulated I.F. carrier. In the case of frequency modulated reception, slow variation of oscillator frequency is serious because it limits the permissible frequency deviation of the carrier and causes amplitude distortion at high modulation levels.

Frequency variation may be separated into long- and short-period effects. Long-period changes, i.e., slow drift of oscillator frequency, are generally the most troublesome and produce greatest frequency variation ; heat and humidity are the chief causes by varying the inductance and capacitance of the tuning circuit. Increase of temperature normally increases inductance and capacitance and lowers the oscillation frequency. Heating of the valve also adds its quota by producing a change in the parameters, μ, g_m and R_a, and in the electrode spacings. Short-period changes are chiefly due to H.T. supply voltage fluctuation from hum, mains interference or feedback from the audio frequency stages, but an important source is also the frequency changer. Frequency variations from the latter are mainly a result of A.G.C. and coupling (electronic and capacitive—Sections 5.8.2 and 5.8.3) to the signal circuit. For this reason A.G.C. is often not applied to a frequency changer operating on the short and ultra short wave ranges.

Though long- and short-term fluctuations form a convenient classification it is preferable to analyse under the headings of the components, causing the frequency variations, e.g., the valve, the *LC* circuit and its associated components, such as the range switch, coupling capacitors, etc., and frequency changer.

6.6. Frequency Variations due to the Valve.[21]

6.6.1. Introduction. Frequency variations due to the valve may be ascribed to four causes :

1. The valve is not functioning as a pure negative resistance, but has a reactive component, the value of which is affected by power supply voltages. This reactance is in parallel with the *LC* circuit and directly affects the resonant frequency.

2. Harmonics are present. They are essential to stabilized

oscillation, but the smaller the harmonic content the less is the frequency variation.

3. Interelectrode capacitance change. This capacitance is made up of two components, one due to the electrode dimensions and spacing, and the other to the space-charge distribution of the electrons. Change of the former due to heating of the electrode structure is a long-period effect, whilst the latter may have a long-period effect due to heating of the valve changing the emission properties, and a short-period effect due to H.T. supply voltage variations. Greatest space-charge capacitance is usually in the neighbourhood of a negative electrode, and hence grid-cathode space-charge capacitance is greater than anode-cathode space-charge capacitance. A tuned grid oscillator therefore shows greater frequency variation than a tuned anode for given fluctuations of supply voltage.

4. Variation in valve slope resistance R_a resulting from power supply fluctuations. The frequency formulae developed in preceding sections indicate that changes of R_a affect the oscillation frequency.

6.6.2. Valve Reactance. The valve has the effect of reducing the resistance component of the LC circuit to zero so that the slightest random disturbance sets up oscillation. At the same time it introduces a reactive component which modifies the equation for start of self-oscillation to

$$R+j\left(\omega L-\frac{1}{\omega C}\right)-(R_1+jX_1)=0.$$

If $X_1=0$ we see that the oscillation frequency is $f=\dfrac{1}{2\pi\sqrt{LC}}$ and is independent of the valve. The frequency f is $\dfrac{1}{2\pi\sqrt{\left(L-\dfrac{X_1}{\omega}\right)C}}$ or $\dfrac{1}{2\pi\sqrt{L\left(C+\dfrac{1}{\omega X_1}\right)}}$ if X_1 is respectively inductive or capacitive.

From the above it would appear possible to neutralize the frequency variation due to X_1 by inserting a suitable correcting reactance in series with the grid or anode lead. An improvement in frequency stability can be brought about by the inclusion of a neutralizing reactance but, since the latter must be varied when the LC circuit frequency is varied, it is not feasible for correcting oscillators covering a large frequency range. Furthermore, owing to non-

linear operation of the valve, reactance neutralization cannot prevent, but can only reduce frequency variation due to the valve. The effect of valve reactance can be reduced by using very loose coupling between the valve and tuned circuit, and one way is to tap the valve connection into a part only of the tuning coil, rather than across the whole coil.

6.6.3. Harmonics. In oscillators there must be an energy as well as power balance. The latter entails neutralization of LC circuit resistance whilst the former requires equality of inductive and capacitive energy. Harmonic currents, generated by the valve, flow more easily through the capacitance than the inductance branch of the LC circuit so that the capacitive energy tends to exceed the inductive energy. To restore energy balance the current through the inductive branch must be increased. This may be realized by a reduction of fundamental oscillation frequency. Experiment has tended to show the opposite effect, viz., that increase of harmonics result in an increase of frequency, and it is assumed that another effect is occurring simultaneously in the opposite direction. Nevertheless, better frequency stability is registered when the harmonic content is low. A valve having a high R_a, a coil resistance as low as possible, a low L/C ratio, and minimum grid current all contribute to low harmonic content.

6.6.4. Interelectrode Capacitance Variation. Variation of interelectrode capacitance due to space-charge effects is greatest between grid and cathode. There is usually an increase in grid-cathode capacitance of about 2 $\mu\mu$F from the cold to hot condition (filament on and off), but the variation under operating conditions is much less than this, being of the order of 0·04 $\mu\mu$F. The magnitude of the variation is dependent on the anode current, grid voltage, heater voltage, and mutual conductance. A valve with a high anode current and mutual conductance generally produces large capacitance change. The space-charge capacitance change is not linearly proportional to grid voltage so that there is a mean as well as periodic change of capacitance. Any change in the oscillation amplitude therefore affects the interelectrode capacitance. This may be minimized by using a large value of tuning capacitance in the LC circuit, by employing a tuned anode in preference to a tuned grid oscillator, and by tapping the valve into a part of the tuning coil. Interelectrode changes between anode and cathode are usually less than 0·01 $\mu\mu$F, and their effect on the oscillator frequency may be reduced by the insertion of a resistance, R_3 in Fig. 6.10, between the anode and tuned circuit. This has the advantage of reducing

the effect of changes of R_a, but it is not very suitable for the short wave ranges because of stray capacitance across the resistance.

6.6.5. Valve Internal Resistance. The linear theory of oscillation shows that frequency is dependent on R_a, which should be as high as possible consistent with a high g_m value. One method of reducing the influence of changes of R_a is to use loose coupling between the valve and tuned circuit, i.e., tapping across a part instead of the whole tuning inductance. The series resistance mentioned in 6.6.4 helps in this direction, but the use of two valves, as in the Franklin oscillator of Fig. 6.12, is a better solution.

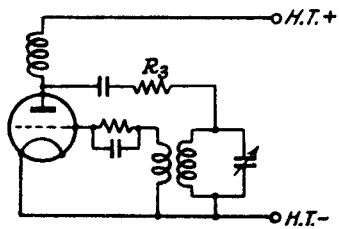


FIG. 6.10.—Reduction of Valve Effects on Frequency Stability by a Series Resistance.

6.6.6. Miscellaneous Effects causing Frequency Variation. Rapid fluctuations of the H.T. voltage may be prevented from influencing frequency by a resistance-capacitance decoupling circuit, R_2C_2 in Fig. 6.6. The decoupling resistance is also useful in reducing long-period variations in H.T. voltage, because it acts as one arm of a potentiometer with the oscillator valve D.C. resistance as the other, so making the D.C. anode voltage change less than the H.T. voltage change.

Frequency variation due to valve temperature change is a comparatively short-period effect occurring in the first 5 or 10 minutes after switching on, for the valve internal temperature settles down quite rapidly to a value not greatly affected by normal external temperature changes.

6.6.7. Special Methods of Reducing Frequency Variations due to the Valve. Various methods may be used to reduce the effect of the valve on frequency. One is by negative feedback ; an impedance, consisting of a resistance, or a resistance and inductance in series, is inserted in the cathode circuit of the valve. The series RL circuit is preferable since it is a better amplitude stabilizer than the resistance alone, across which there is stray capacitance tending to make negative feedback less effective at the higher oscillating frequencies. A triode oscillator with the tuned circuit connected

between grid and H.T. negative, the cathode tapped partway up the tuning inductance, and the anode connected to H.T. positive is an example of negative feedback control.

The electron coupled [4] oscillator is another alternative. A screened-grid or pentode valve is used with the grid and cathode as the active oscillating elements. Negative feedback compensation is

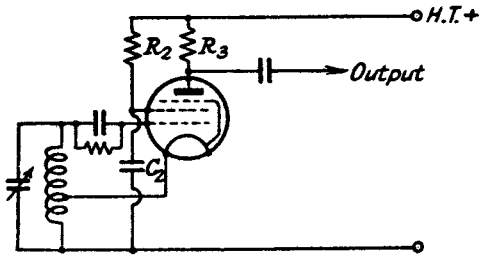


FIG. 6.11.—The Electron Coupled Oscillator.

realized because the cathode is connected to a tapping point on the *LC* circuit. The diagram of connections is shown in Fig. 6.11, and the normal anode merely serves as a means of developing the oscillator voltage for application to the frequency changer. The screen carries no R.F. voltages (capacitance C_2 has a high value—0·1 μF), and therefore reduces the capacitance coupling between the oscillator proper and the anode. Two advantages are gained from this type; the voltage applied to the frequency changer is not obtained directly from the oscillating circuit so that variations

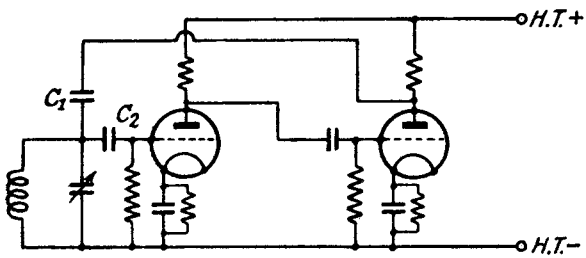


FIG. 6.12.—The Franklin Oscillator.

in frequency changer conditions have almost no effect on oscillator frequency. The second advantage is that the frequency variation due to increase or decrease of screen voltage is in the opposite direction to that for increase or decrease of anode voltage. Hence by suitably adjusting R_2 and R_3 the oscillation frequency can be made almost independent of H.T. supply voltage changes.

A very successful method of reducing valve effect is the Franklin

circuit [15] of Fig. 6.12. The capacitances C_1 and C_2 are very small (1 to 2 $\mu\mu$F) so that the LC circuit is only very loosely coupled to the valves. Since two valves are required the method is not very suitable for use in broadcast receivers.

6.7. Frequency Variations due to the LC circuit and its Associated Components.

6.7.1. Introduction. The chief cause of variation of the inductance, capacitance and resistance of the LC circuit and its associated components is temperature change, though humidity can have a more serious effect if care is not taken in moisture-proofing the coils or fixed capacitors. By the use of high-grade waxes or varnishes, humidity variations can generally be reduced to a second order effect. Increase of humidity has much the same effect as increase of temperature, viz., it increases the capacitance (in the case of a coil, its self-capacitance) and reduces the oscillator frequency. The increase in capacitance is due to the displacement of air in the insulating material by water vapour, which has a higher dielectric constant.

Variations in the values of inductance and capacitance of the LC circuit are largely responsible for the total frequency variation. Associated components, such as the padding and self-bias capacitances, the range switch and wiring, also add their quota to the frequency change but are much less important. The trimmer capacitance can cause very large changes of frequency when the tuning capacitance is small.

6.7.2. Inductance Variations. Changes of temperature produce a variation in the inductance, resistance and self-capacitance of a coil. The last two are not, however, important in comparison with the change in inductance. Taking the formula for a single-layer solenoid (the shape most likely to be used on the medium and short wave ranges) given by expression 4.12

$$L = \frac{r^2N^2}{9r+10l}$$

where l = length of winding
r = radius of winding
N = total turns in the coil

we see that increase of length decreases L, whereas increase of radius increases L. Change of radius has the greater effect and increasing temperature therefore increases L. If we assume that the coefficients

of radial and axial expansion, α_r and α_l respectively are equal, the inductance for an increase in temperature of $t°$ C. becomes

$$L+\Delta L = \frac{r^2(1+\alpha t)^2N^2}{9r(1+\alpha t)+10l(1+\alpha t)} \quad . \quad . \quad 6.29a.$$

$$L+\Delta L = \frac{r^2N^2}{9r+10l}(1+\alpha t)$$

i.e., the increase in inductance is directly proportional to the change in temperature multiplied by the coefficient of expansion of the conductor. Expressing the result in a more convenient form, the coefficient of inductance increase equals the coefficient of expansion of the conductor, or

$$\frac{\Delta L}{L.t} = \alpha ;$$

the value of α for copper is 16×10^{-6} parts per degree Centigrade. But

$$f = \frac{1}{2\pi\sqrt{LC}}$$

and

$$f - \Delta f = \frac{1}{2\pi\sqrt{(L+\Delta L)C}}$$

where Δf is the decrease in oscillator frequency due to the increase ΔL in L

$$\frac{f-\Delta f}{f} = \sqrt{\frac{L}{L+\Delta L}}$$

$$1 - \frac{\Delta f}{f} = \left(1+\frac{\Delta L}{L}\right)^{-\frac{1}{2}} \quad . \quad . \quad . \quad 6.30a.$$

Expanding by the Binomial Theorem and neglecting $\left(\frac{\Delta L}{L}\right)^2$ and all higher powers—this is justified because $\frac{\Delta L}{L}$ is so small—expression 6.30*a* may be rewritten

$$\frac{\Delta f}{f} = \frac{1}{2}\left(\frac{\Delta L}{L}\right) \quad . \quad . \quad . \quad . \quad 6.30b.$$

Thus the ratio change in frequency due to the unhindered expansion of a copper coil is 8×10^{-6} parts per degree centigrade, or for the normal temperature rise of 30° C. the ratio change of frequency is 240 in 10^6. Frequency variations, due to coil expansion, of 240 and 2,400 c.p.s. at oscillator frequencies of 1 and 10 Mc/s respectively, would be expected under these conditions.

Since increase of radius has the opposite effect to increase of length, it is possible by suitably choosing α_r and α_l to make inductance independent of temperature. Re-writing 6.29*a* with α_r and α_l

$$L+\Delta L = \frac{r^2(1+\alpha_r t)^2 N^2}{9r(1+\alpha_r t)+10l(1+\alpha_l t)} \quad . \quad . \quad . \quad 6.29b.$$

Differentiating the above with respect to temperature t and equating the result to 0 gives the condition for no change of inductance

$$\frac{d(L+\Delta L)}{dt} = \frac{r^2N^2[(9r(1+\alpha_r t)+10l(1+\alpha_l t))(1+\alpha_r t)2\alpha_r - (1+\alpha_r t)^2(9r\alpha_r+10l\alpha_l)]}{[9r(1+\alpha_r t)+10l(1+\alpha_l t)]^2}$$

when $t = 0$

$$\frac{d(L+\Delta L)}{dt} = 9r\alpha_r+10l(2\alpha_r - \alpha_l)$$

$$= 0$$

when $$\frac{\alpha_l}{\alpha_r} = 2+\frac{9r}{10l} \quad . \quad . \quad . \quad . \quad . \quad 6.31.$$

By satisfying the above expression for the ratio of axial to radial expansion, it is possible to make L independent of temperature. The practical application involves winding the coil turns loosely, and fixing the ends firmly to the coil former so that radial expansion is determined by the conductor, and axial by the coil former. Suitable coil-former materials, having coefficients of expansion greater than twice that for copper, are not easy to obtain, but experiments have shown the validity of the theoretical argument. An alternative is to reduce the radial and axial expansion, either by shrinking the coil on to a former having a lower coefficient of expansion, e.g., ceramic material having coefficient of 7×10^{-6} parts per degree centigrade, or by using copper-plated low-coefficient material, such as invar, for the coil. The latter is not an economic proposition for receivers, and the best alternative is the ceramic coil former. Synthetic resin or bakelite coil formers are not to be recommended because they tend to show a non-reversible ageing, which results in a permanent change in inductance after several cycles of temperature change. They are also more affected by humidity variation than ceramic material.

6.7.3. Capacitance Variations. Temperature variation has two effects in capacitors; it changes the permittivity of the dielectric and also changes the disposition of the plates. In fixed capacitors, mica is the most suitable dielectric and capacitance change is largely a result of pressure variation between the plates. This effect can be considerably reduced by replacing the metal

plates by metal (silver) sprayed on to the mica dielectric. Both sides of the mica are sprayed and each side forms an active plate; adjacent sides are connected together so that spacing variations due to changes of pressure become of little importance. This type of silvered mica fixed capacitor has a very small temperature coefficient and its effect on oscillator frequency variation is very small.

Fixed capacitors (maximum value about 3 $\mu\mu$F) with a bimetallic plate can be constructed to give a decrease of capacitance with increase of temperature, and they may be used to compensate for the increase in capacitance and inductance of the main elements of the tuning system.

Variable capacitors often have a high temperature coefficient and their effect on oscillator frequency is generally much greater than that of the coil. For this reason tuning by inductance variation with fixed capacitors, which can be constructed with low temperature coefficients, is to be preferred, and is employed in some short wave band-spread receivers. By observing the following points the temperature coefficient can be reduced. The dielectric should be air, as the use of a solid dielectric, such as mica, tends to distort the plates and increase the high-frequency resistance component. The insulating end supports have an important effect, particularly when the capacitance is minimum, and ceramic material is best as it has a low permittivity change with temperature. Synthetic plastic materials are not usually satisfactory and have non-cyclical variation. The rotor shaft and stator frame should be of the same material, so that expansion allows the rotor plates to remain centrally disposed with respect to the stator. Initial accurate alignment of a rotor plate midway between two stator plates reduces capacitance change due to differential expansion, because a given change of spacing has much greater effect when the rotor-to-stator plate spacings are unequal. Increase of stator-rotor plate spacing also reduces capacitance change for a given change of spacing. Good mechanical rigidity and construction are essentials for small capacitance variation with temperature.

Particular care is necessary in the choice of the trimmer capacitor, which can have a very large influence on the oscillator frequency as the temperature varies. Mica dielectric pressure-type capacitors are generally unsatisfactory and show an appreciable increase in capacitance with increase of temperature. Air dielectric capacitors with ceramic insulation should be employed.

6.7.4. Frequency Variations due to Associated Components. By the use of silvered mica construction, frequency

drift due to fixed capacitors associated with the tuned circuit can be reduced to very small proportions. That due to the wave-change switch and wiring [14] can also be made a second-order effect by observing certain precautions. The switch should be of rigid mechanical construction with ceramic insulation, and the wiring rubber-covered or enamelled. Especial care is necessary to ensure that leads are securely fixed and not in tension. Preliminary cyclical heating is an aid to frequency stability.

Stray capacitance between the tuning inductance and the feedback coil must be reduced to a minimum as this generally has a high temperature coefficient. It is preferable not to wind the feedback coil on top of the tuning inductance but to have an air space between them.

Examples of the capacitance-temperature coefficients of typical insulating materials given below show the desirability of using ceramic [20] rather than other forms of insulation.

Varnished cambric	+2,100	parts	in	10^6	per	degree	Centigrade
Synthetic resin	+1,600	,,	,,	,,	,,	,,	,,
Enamel	+ 470	,,	,,	,,	,,	,,	,,
Ceramic	+ 100	,,	,,	,,	,,	,,	,,

6.7.5. Compensation for Temperature-Frequency Change. Certain types of capacitors can be constructed to give a negative temperature coefficient, i.e., capacitance decreases with increase of temperature, and they may be used to compensate for the positive temperature coefficient of the normal tuning inductance and capacitance. Compensation is only complete, however, at one particular frequency, and the temperature of the corrector must follow that of the component it is intended to compensate. Two corrector capacitors may be used, one, with a heater winding generally supplied from the A.C. mains transformer, to compensate for the rapid initial frequency change due to the valve warming-up, and the other placed so as to follow the temperature of the tuning system and to correct for change in L and C. The corrector capacitor may be connected in parallel with the tuning inductance or capacitance—the bimetallic construction described in Section 6.7.3 has a capacitance change of the order of 0·5 $\mu\mu$F for 30° C. variation—or in parallel with the padding capacitance. Compensation does not obviate the necessity for aiming at the smallest possible frequency variation of the uncompensated oscillator.

6.8. Frequency Variations due to the Frequency Changer. [9] The input admittance at the oscillator grid of the frequency changer

includes three possible variables, grid current, space-charge capaci tance, and space-charge and capacitance coupling to the signal grid. The effect of grid current variation due to bias changes on the frequency-changer valve is least important when the oscillator and frequency-changer grids are connected together, as in the triode-hexode. It is more important when an R.C. coupling is employed between the two grids, as for suppressor grid application in the pentode (Fig. 5.4), but the frequency variation can be reduced to small proportions by using a high resistance (1 MΩ) from suppressor grid to earth.

Frequency variation due to capacitive and space-charge effects is largely beyond the receiver designer's control, because it is most dependent on valve construction. All the effects can be reduced by making the coupling between the oscillator-tuned circuit and the frequency changer as loose as possible consistent with adequate oscillator voltage. The greatest load effect is produced when the oscillator grid is directly coupled to the tuned circuit as for the tuned grid oscillator connection. By the use of a tuned anode oscillator a reduction in frequency variation of about 5 to 1 may be obtained for a given change of frequency changer bias.

A reduction of load-frequency variation is achieved by using a low impedance LC circuit, i.e., a high capacitance and low inductance, for the addition of a resistance and reactance in parallel with any LC circuit has least effect when the L/C ratio is small.

Space-charge coupling between signal and oscillator grids is very much less in hexode than in heptode type frequency changers. The use of a neutralizing capacitance about 1 $\mu\mu$F may be used in the latter case as described in Section 5.8.3.

Stray capacitive coupling between the signal and oscillator grids must be avoided if the tuning of the signal circuit is not appreciably to affect the oscillator operation on the short wave ranges.

6.9. Precautions Necessary to Preserve Frequency Stability. The following summarizes the steps which may be taken to increase the frequency stability of the oscillator.

1. A valve having a high g_m and high R_a is required.

2. Grid current should be as low as possible, i.e., a high grid self-bias resistance, provided squeggering is not experienced, and feedback coupling only just sufficient to sustain oscillation at the required amplitude are necessary.

3. The coil resistance should be as low as possible.

4. The L/C ratio is required to be low so that variations of stray capacitance and load reflected from the frequency changer

have less effect. Too low a value, however, makes oscillation difficult. Except in band-spread receivers with fixed signal tuning, and push-button operated receivers, the values of L and C are fixed by ganging considerations, and little control is possible.

5. The use of a phase-correcting reactance is not feasible as it is only suitable for fixed-frequency operation.

6. The feedback coupling, if a coil, should have the smallest value of coil inductance with maximum value of mutual inductance coupling to the tuned LC circuit.

7. The tuned circuit should be loosely coupled to the valve, i.e., the valve should be connected across a part of the tuning inductance, and tuned anode is preferable to tuned grid because space-charge capacitance changes are greatest at the grid electrode.

8. The H.T. supply should be adequately decoupled and smoothed. The possibility of audio-frequency feedback from the power output stage may be reduced by a push-pull audio-output connection.

9. Temperature-frequency variations are reduced by mounting inductances and capacitances away from such sources of heat as the mains transformer, the power output and rectifier valves. Low-loss coil formers of ceramic material are to be recommended. Different rates of radial and axial expansion for the coil are a possible means of reducing the temperature effect. High-loss insulators must be avoided as these almost always have high positive temperature coefficients. Rigid mechanical construction of coils, variable capacitors, range switches and wiring is very necessary. Variable capacitance values should be minimum, and inductance tuning with fixed capacitance is desirable.

10. Compensation by a capacitance having a negative temperature coefficient may be provided.

11. Humidity effects demand the use of non-hygroscopic insulating material, waxes and varnish.

12. The optimum oscillator voltage required by the frequency changer should be as low as possible, as this allows loose coupling to and reduces the load reflected from the frequency changer into the oscillator tuned circuit.

13. A frequency correcting device may be connected to the oscillator in order to introduce additional reactance in parallel with the LC circuit as the frequency tends to vary from its correct setting. The magnitude of the reactance is determined by the amount of error, and high initial frequency errors may be reduced to quite small values. Thus an uncorrected oscillator frequency

error of 20 kc/s can be reduced to a final error of only about 200 c.p.s. Methods of automatically correcting frequency variations by means of a variable reactance valve are discussed in Chapter 13.

6.10. Squegger and Parasitic Oscillations. Two undesirable effects, known as squegger and parasitic oscillations, may be met in self-oscillating valve circuits. Squegger oscillation is a regular or irregular interruption of the normal oscillation, and is a result of excessive amplitude. The large positive pulses of grid voltage produce sufficient grid current to charge the capacitor across the self-biasing grid leak to a negative voltage much greater than anode current cut-off. The anode current ceases and oscillations cannot be maintained. The capacitor gradually discharges through the grid leak until the grid voltage is low enough to allow anode current to flow and oscillation to recommence. Grid current again biases the valve beyond cut-off and the cycle is repeated. The squegger effect is equivalent to a 100% modulation of the oscillator by an approximately square wave. The period of the interruption depends on the time constant of the capacitor and grid leak, and it may produce in the output of a receiver an audible note (if the interruption frequency is in the audible range), reduced desired output, excessive whistle interference and noise. It can normally be prevented by using a low value of self-biasing capacitance and grid leak, and feedback coupling only just sufficient to give the required oscillator amplitude at the frequency changer. Values of 100 $\mu\mu$F and 50,000 ohms are typical for the medium and long wave ranges, but for short wave operation the capacitance may be reduced to 50 $\mu\mu$F. A large H.T. decoupling resistance (R_2 in Fig. 6.6) also helps to prevent squegger oscillations, for the oscillator anode voltage is increased by decrease of mean anode current and this moves the cut-off voltage to a higher negative value.

Parasitic [16] oscillations are oscillations at a frequency other than that of the controlling LC circuit. Generally they are at very high frequencies and are due to the inductance of the leads from, and the stray capacitance across, the valve electrodes. These inductances and capacitances form tuned-anode and tuned-grid circuits (Fig. 6.13) and the necessary feedback coupling is provided by the anode-grid capacitance. High-frequency parasitic oscillation is more likely with high μ and g_m valves, but it may also be produced by a negative resistance characteristic in the I_aE_g curve of the oscillator grid of the frequency changer. This negative resistance

effect more often occurs at appreciable positive voltages on the oscillator grid, and it can be reduced by preventing excessive oscillator amplitude. A small resistance (about 20Ω) connected close to grid or anode pin of the valve is helpful in damping the parasitic control circuits. The series grid resistance R_1 of Fig. 6.6 performs this rôle as well as that of amplitude stabilizer.

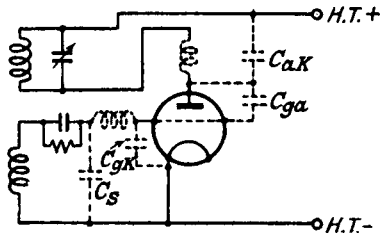


FIG. 6.13.—Parasitic Oscillation due to Lead Inductance and Stray Capacitance.

The modified Colpitts oscillator for short wave operation (Fig. 6.14) uses the anode-cathode, grid-cathode capacitances as the capacitance tap, and parasitic oscillation is possible when the tuning capacitance C is large, for the lead inductances can act as the control inductance, and C_{ak} and C_{gk} as the control capacitance. A remedy is to connect the tuning capacitor as close to the anode and grid pins as possible.

The Hartley oscillator may be troublesome at high frequencies by attempting to act as a modified Colpitts. The introduction of a small R.F. choke or resistance between the centre tap and

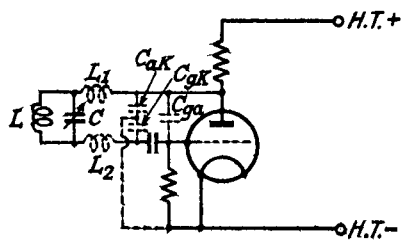


FIG. 6.14.—Parasitic Oscillation in the Modified Colpitts Oscillator.

the H.T. supply is a solution, for the oscillator functions as a Hartley for low frequencies and as a modified Colpitts for high frequencies.

Parasitic oscillation not far removed from the desired frequency is possible when the feedback coil has an inductance comparable

with that of the tuned *LC* circuit. Such a condition is unlikely to arise on the medium and long wave ranges, but it is quite possible on the short wave range. As the oscillator frequency is increased, i.e., the tuning capacitance decreased, a point is reached at which the oscillation frequency jumps to a value approximately given by the resonant frequency of the feedback coil and its stray capacitance. The original functions of the feedback and control *LC* circuit are now reversed, the former acting as the control. The remedy for this type of parasitic oscillation is to use the lowest possible value of feedback coil inductance and to obtain sufficient feedback by tight coupling to the coil of the *LC* circuit.

6.11. Problems in the Design of Short Wave and Ultra Short Wave Oscillators.[19] Oscillators for the short wave and ultra short wave ranges present in a more acute form all the problems encountered on the medium and long wave ranges. In addition they are more prone to squegger and parasitic oscillations and, in the case of the tuned grid and tuned anode circuits, are more difficult to maintain in oscillation over a range of frequencies.

A high g_m valve is an essential as it enables looser coupling to be employed, and this reduces the tendency to squegger and parasitic oscillation. With tuned grid and tuned anode circuits the feedback coil is often interleaved with the main tuning coil in order to obtain high coupling with a small feedback coil.

The ease with which the Hartley and Colpitts circuits can be made to oscillate is a great advantage at high frequencies, and the possibility of feedback coil frequency control is removed. The fact that the Hartley circuit tuning capacitor rotor is not earthed and that the Colpitts requires a split capacitor is no serious disadvantage on band-spread receivers with preset signal tuning. The negative feedback oscillator of Section 6.6.7, with the cathode tapped into the tuning coil, has the advantages of the Hartley circuit and also that one side of the tuning capacitor may be earthed. The modified Colpitts oscillator of Fig. 6.14 is satisfactory at high frequencies and obviates the necessity for a split capacitor, the anode-cathode and grid-cathode interelectrode capacitances performing this function. The tuning capacitor rotor cannot, however, be earthed. The disadvantages of having neither side of the tuning capacitance connected to earth can be overcome by using inductance tuning. On the short wave range this may be accomplished by varying the position of an iron dust core in the tuning coil, and on ultra short waves by inserting a metal plunger of high conductivity material into the coil axis.

Frequency drift of the oscillator due to changes of temperature is more serious as the frequency is increased, partly because the LC circuit components have to be reduced in value, but also because the frequency error Δf, even for a fixed ratio of frequency change $\left(\frac{\Delta f}{f}\right)$, increases as the oscillation frequency rises. Thus for a frequency ratio change of 1 part in 10^3 the frequency drift is 1 kc/s at a frequency of 1 Mc/s, whilst at 45 Mc/s it is 45 kc/s, sufficient to take the I.F. carrier outside the pass range of an I.F. amplifier for amplitude modulated sound signals. Since variable and stray capacitances are the chief offenders, it means that the frequency ratio tends to increase as the oscillation frequency is increased and the tuning capacitance decreased. A low L/C ratio helps to reduce the frequency ratio, but at the same time makes oscillation more difficult to sustain. When the I.F. has a high value, as for the vision signals of a television programme, and the sound signals of a frequency modulated transmission, the use of an oscillator frequency ($f_h = f_s - f_1$) lower than the signal frequency is an advantage in reducing frequency drift. An alternative is to employ a harmonic of the oscillator to combine with the signal ; Section 5.5 shows that a 2nd harmonic response approaching 80% of the fundamental oscillator response is obtainable when a large oscillator voltage is applied to the frequency changer. Provided the increase in possible interference whistle production is no disadvantage, oscillator frequency drift can be reduced to approximately one-half by using the second harmonic of the oscillator. For example, suppose temperature change produces an increase of capacitance by 0·05 $\mu\mu$F ; at 45 Mc/s, the maximum value of tuning capacitance is about 40 $\mu\mu$F if adequate oscillation amplitude is to be realized, and the frequency drift is

$$\Delta f = f \cdot \frac{1}{2} \frac{\Delta C}{C}$$

$$= \frac{45}{2} \times \frac{0{\cdot}05}{40} \times 1{,}000 \text{ kc/s}$$

$$= 28{\cdot}1 \text{ kc/s}.$$

(The proof of this is similar to that developed in Section 6.7.2 for inductance change.)

When the oscillator frequency is halved, the tuning capacitance may be increased to 80 $\mu\mu$F for the same $\frac{L}{CR}$ value (this is a measure

of the oscillation capability of the circuit, and doubling L and C doubles the frequency with little change of $\frac{L}{CR}$). Thus the frequency drift of the oscillator becomes

$$\Delta f = \frac{22{\cdot}5 \times 0{\cdot}05}{2 \times 80} \times 1{,}000$$
$$= 7{\cdot}025 \text{ kc/s},$$

which is one-quarter of that at 45 Mc/s. The frequency drift for the second harmonic is twice this value, i.e., 14·05 kc/s, so that the net frequency drift has been reduced to one-half of its original value for an oscillator of 45 Mc/s fundamental frequency.

The need for short, firmly secured, leads from the valve to the control LC circuit cannot be over-emphasized, and adequate decoupling with non-inductive mica capacitors (0·001 to 0·01 μF) of leads carrying D.C. or mains A.C. voltages is essential. Only a very small inductance can provide undesirable coupling at ultra high frequencies, and decoupling capacitors should therefore be returned to the same point on the chassis. In a shunt-fed circuit (see Fig. 6.10) with a choke between the anode and H.T. positive, particular care must be taken in the construction of the choke, which, if used to cover a number of ranges, should be wound in separate sections, becoming smaller at the anode side. This is to prevent resonance of the choke with its self-capacitance, thereby causing absorption and dead spots at particular frequencies corresponding to parallel resonance of the whole or a section of the choke with self and stray capacitance. When this does occur, it may be possible by adjustment of the turns in the sections nearest the anode to remove the absorption frequency outside the range. When the oscillator is intended for operation at ultra high frequencies only, adjustment of the anode choke, and also of any other decoupling choke in the cathode or heater circuits, may have considerable effect on oscillator amplitude, too large a value of choke reducing the amplitude as well as too small a value.

6.12. Ganging of the Oscillator and Signal Circuits.[5, 11, 12, 18] In order to reduce the number of receiver controls to a minimum, it is usual to couple mechanically the signal and oscillator capacitor rotor plates. Since the oscillator capacitance is required to tune its own circuit to a frequency greater than the signal frequency by a constant amount, a specially shaped rotor and stator are necessary to maintain the correct oscillator frequency. In certain types of receivers the oscillator capacitor is designed to

give the required capacitance law over the medium wave range, but then the same shaped capacitor is not correct for the long and short wave ranges. It is not economical to have a separate capacitor for each range and a compromise is obviously desirable. Such is possible by providing a number of preadjusted components in the oscillator tuned circuit, the term preadjusted or preset being applied to a component, the value of which is under the control of the designer. This component is constant for a given wave range, but is varied when the latter is changed. The oscillator frequency can be given its correct value at the same number of signal frequencies as there are preset components, and by using such a circuit an oscillator capacitor identical with that tuning the signal may be employed.

A circuit having two preset components, the tuning inductance L_h and the padding capacitance C_p, is shown in Fig. 6.15. The

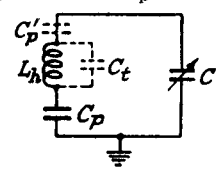


FIG. 6.15.—An Oscillator Ganging Circuit.

padding capacitance may be placed between the high potential ends of L_h and the tuning capacitance C, or it may be placed as shown between L_h and earth. The second position has the advantage of not adding to the capacitance in parallel with the coil, for in the first the capacitance to earth of C_p is in parallel with C_0, the coil self-capacitance (C_t in Fig. 6.15). The oscillator frequency is

$$f_h = \frac{1}{2\pi\sqrt{L_h\left[C_0+\frac{C_pC}{C_p+C}\right]}} \quad . \quad . \quad . \quad 6.32.$$

For any two values, f_{h1} and f_{h2}, this equation may be satisfied simultaneously by suitably selecting L_h and C_p. Suppose that these two oscillator frequencies correspond to two signal frequencies f_{s1} and f_{s2}, for which the tuning capacitance C has values of C_1 and C_2. The following simultaneous equations result

$$\frac{1}{[2\pi f_{h1}]^2} = \frac{1}{[2\pi(f_{s1}+f_1)]^2} = L_h\left[C_0+\frac{C_pC_1}{C_p+C_1}\right] \quad . \quad 6.33a$$

$$\frac{1}{[2\pi f_{h2}]^2} = \frac{1}{[2\pi(f_{s2}+f_1)]^2} = L_h\left[C_0+\frac{C_pC_2}{C_p+C_2}\right] \quad . \quad 6.33b$$

where f_1 = the intermediate frequency.

Solving for C_p gives a quadratic

$$C_p^2\left[\frac{C_1-C_2}{\alpha-1}-(C_0+C_2)\right]-C_p[C_0(C_1+C_2)+C_1C_2]-C_0C_1C_2=0. \quad 6.34$$

where $$\alpha=\left[\frac{f_{h2}}{f_{h1}}\right]^2=\left[\frac{f_{s2}+f_1}{f_{s1}+f_1}\right]^2$$

and

$$C_p=\frac{C_0(C_1+C_2)+C_1C_2+\sqrt{[C_0(C_1+C_2)+C_1C_2]^2+4C_0C_1C_2\left[\frac{C_1-C_2}{\alpha-1}-(C_0+C_2)\right]}}{2\left[\frac{C_1-C_2}{\alpha-1}-(C_0+C_2)\right]} \quad 6.35.$$

The second solution of C_p with a negative sign before the square root is not a practical possibility. The value of L_h is found by substituting expression 6.35 for C_p in either 6.33*a* or 6.33*b*.

Thus $$L_h=\frac{1}{[2\pi f_{h1}]^2\left(C_0+\frac{C_pC_1}{C_p+C_1}\right)} \quad . \quad . \quad 6.36$$

The most suitable values of f_{s1} and f_{s2} are those giving the minimum error variation over the wave range, and to determine them the error curve shape must be known. The best choice is

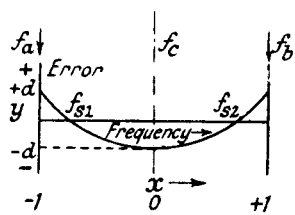


FIG. 6.16.—The Ideal Curve for Two-point Zero Error Ganging.

clearly either that which gives errors equal in magnitude at the ends and centre of the range, or least error at the low frequency end of the range where the signal circuits are normally more selective. Let us consider the first alternative and assume that the error curve is parabolic as in Fig. 6.16, with equal errors at the ends and centre of the range. By designating the error axis as y and the frequency axis as x, the general expression for the parabola is

$$y=ax^2+bx+c \quad . \quad . \quad . \quad . \quad 6.37.$$

Taking f_c, the frequency at the centre of the range, as the origin,

f_a and f_b the end frequencies, as $x = -1$ and $+1$ respectively, and the maximum error as d kc/s, the general expression reduces to

$$\begin{aligned} y &= ax^2 + c \\ &= d \text{ when } x = -1 \text{ and } +1, \\ &= -d \text{ when } x = 0 \end{aligned}$$

so that $a = -2c = 2d$ and

$$y = d(2x^2 - 1) \quad . \quad . \quad . \quad . \quad 6.38.$$

The frequencies of zero error are obtained by equating 6.38 to 0.

Thus $$x = \pm\sqrt{\tfrac{1}{2}} = \pm 0{\cdot}707.$$

The required signal frequency values are therefore

$$f_{s1} = f_c - 0{\cdot}707\,(f_c - f_a)$$

and $$f_{s2} = f_c + 0{\cdot}707\,(f_c - f_a)$$

for $$f_c - f_a = f_b - f_c.$$

Let us now consider the above procedure applied to the design of an oscillator circuit to gang with a signal circuit of 156 μH inductance for an I.F. of 465 kc/s, and covering the medium wave range from 550 to 1,500 kc/s. The capacitance required to tune the signal coil at selected frequencies is given below

f_s(kc/s) .	550	600	700	800	900	1,000
C($\mu\mu$F) .	536·8	451	331·1	253·7	200·5	162·38

f_s(kc/s) .	1,100	1,200	1,300	1,400	1,500
C($\mu\mu$F) .	134·48	112·76	96·1	82·84	72·17

Not all the tuning capacitance is found in the tuning capacitor and some is made up of stray capacitance due to the range switch, wiring, self-capacitance of the coil, and the trimmer capacitance (required for signal circuit ganging) across the coil. An average value for each of these capacitances is 10 $\mu\mu$F, making a total stray capacitance of 40 $\mu\mu$F. Hence the actual value of the tuning capacitance is obtained by subtracting 40 $\mu\mu$F from the values given in the above table. We will make the assumption that the oscillator tuning circuit has similar stray capacitance values, except that there is no trimmer capacitance, and that C_p is located close to L_h and permanently connected to it, i.e., the range switch and wiring is on the tuning capacitance side of C_p. The value of C at a signal frequency of 550 kc/s ($f_h = 1{,}015$) is therefore the value of the tuning capacitance (536·8 — 40) plus the stray capacitance (20 $\mu\mu$F) due to the wiring and range switch, i.e., equals 516·8 $\mu\mu$F. The centre signal frequency of the range is 1,025, so that the signal frequencies of zero error are 1,025 $\pm$ 0·707 (475),

589 and 1,361 kc/s. The oscillator frequencies f_{h1} and f_{h2} for zero error are 1,154 and 1,826 kc/s, the tuning capacitances C_1 and C_2 are 322 and 67·6 $\mu\mu$F; if the oscillator coil self-capacitance C_o is 10 $\mu\mu$F the quadratic for C_p (expression 6.34) is

$$9{\cdot}2\ C_p{}^2 - 2{,}564{\cdot}6\ C_p - 21{,}750 = 0$$

$$C_p = 288\ \mu\mu\text{F}.$$

Hence

$$L_h = 117{\cdot}3\ \mu\text{H}.$$

The actual oscillator frequencies at different signal frequencies may be calculated by replacing C_0, C, C_p and L_h in 6.32 by their appropriate numerical values. An error curve for the I.F. is plotted in Fig. 6.17 and it is seen that the parabolic shape is approached.

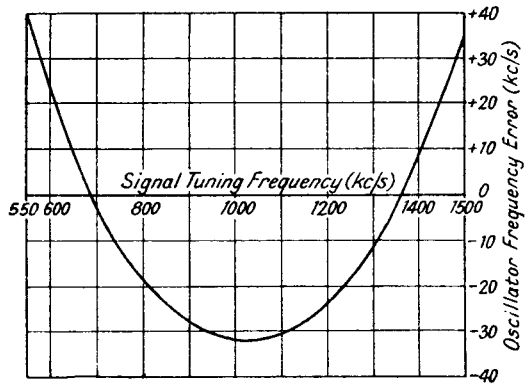


FIG. 6.17.—The Calculated Error Curve for Two-point Ganging.

A greatly improved error curve is obtained by adding another preset component, a trimmer capacitor C_t, in parallel with the coil L_h. It is more usual to place the trimmer C_t directly across L_h rather than across L_h and C_p, since in the first position it can compensate for variations in coil self-capacitance and stray capacitance. In the formulae given below, C_t is the total value of capacitance across L_h, including self and stray capacitance. Zero error can be realized at three signal frequencies f_{s1}, f_{s2} and f_{s3}, and from the following three equations the values of C_p, C_t and L_h can be calculated.

$$\left.\begin{aligned}
\frac{1}{[2\pi f_{h1}]^2} &= \frac{1}{[2\pi(f_{s1}+f_1)]^2} = L_h\left[C_t+\frac{C_pC_1}{C_p+C_1}\right]\\
\frac{1}{[2\pi f_{h2}]^2} &= \frac{1}{[2\pi(f_{s2}+f_1)]^2} = L_h\left[C_t+\frac{C_pC_2}{C_p+C_2}\right]\\
\frac{1}{[2\pi f_{h3}]^2} &= \frac{1}{[2\pi(f_{s3}+f_1)]^2} = L_h\left[C_t+\frac{C_pC_3}{C_p+C_3}\right]
\end{aligned}\right\} \qquad 6.39.$$

Eliminating L_h and C_t we have

$$C_p = \frac{AC_1(C_2 - C_3) + BC_2(C_3 - C_1) + C_3(C_1 - C_2)}{A(C_3 - C_2) + B(C_1 - C_3) + (C_2 - C_1)} \quad . \quad 6.40$$

where $$A = \left[\frac{f_{h3}}{f_{h1}}\right]^2 \text{ and } B = \left[\frac{f_{h3}}{f_{h2}}\right]^2.$$

Solving for C_t

$$C_t = \left[\frac{AC_3C_p}{C_3+C_p} - \frac{C_1C_p}{C_1+C_p}\right]\frac{1}{1-A} \quad . \quad . \quad 6.41$$

$$L_h = \frac{1-A}{[2\pi f_{h3}]^2\left[\frac{C_3C_p}{C_3+C_p} - \frac{C_1C_p}{C_1+C_p}\right]} \quad . \quad . \quad 6.42.$$

Now we must fix the optimum signal frequencies of zero error, and this we can do by assuming the error curve to be a cubic as shown in Fig. 6.18. The general expression for the curve is

$$y = ax^3 + bx^2 + cx + d \quad . \quad . \quad . \quad 6.43a$$

and it is clear that the optimum error curve should be symmetrical

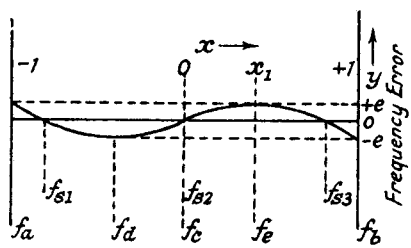


FIG. 6.18.—The Ideal Curve for Three-point Zero Error Ganging.

about the centre frequency f_c with zero error at this frequency. The maximum errors ($y = \pm e$) should all be equal at the four frequencies f_a, f_b, f_d and f_e. By letting $x = 0$, $+1$ and -1 correspond to the centre and end frequencies, f_c, f_b and f_a respectively, we find that $d = 0$ (for $y = 0$, when $x = 0$), $b = 0$ ($y = \pm e$, when $x = \mp 1$) and $(a+c) = -e$. Thus expression 6.43a reduces to

$$y = ax^3 + cx \quad . \quad . \quad . \quad . \quad 6.43b.$$

At f_e there is a point of inflection so

$$\frac{dy}{dx} = 0 = 3ax_1^2 + c$$

and $$x_1 = +\sqrt{\frac{-c}{3a}}.$$

Replacing this value in 6.43*b* and noting that $y = e = -(a+c)$

$$-(a+c) = a\left(\frac{-c}{3a}\right)^{\frac{3}{2}} + c\left(\frac{-c}{3a}\right)^{\frac{1}{2}} \quad . \quad . \quad . \quad 6.44.$$

A cubic equation is obtained by squaring 6.44

$$\frac{4c^3}{27} + ac^2 + 2a^2c + a^3 = 0$$

and one solution of this is $c = \frac{-3a}{4}$.

Inserting this value in 6.43*b* we have

$$y = ax(x^2 - \tfrac{3}{4}) \quad . \quad . \quad . \quad . \quad 6.43c.$$

Zero error is obtained when expression 6.43*c* equals 0, i.e., when $x = 0$ and $\pm\sqrt{\frac{3}{4}}$, and the signal frequencies of zero error are therefore

$$f_{s2} = f_c$$
$$f_{s1} = f_c - \sqrt{\tfrac{3}{4}}(f_c - f_a)$$
$$f_{s3} = f_c + \sqrt{\tfrac{3}{4}}(f_c - f_a).$$

Hence for the medium wave range (550 to 1,500 kc/s) the signal frequencies of zero error are

$$f_{s2} = 1{,}025 \text{ kc/s}, \; f_{s1} = 614 \text{ kc/s}, \; f_{s3} = 1{,}436 \text{ kc/s}.$$

Again assuming a stray and trimmer capacitance in the signal circuit of 40 $\mu\mu$F and a stray capacitance of 20 $\mu\mu$F across the

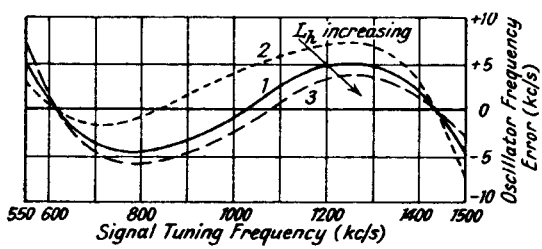


FIG. 6.19.—Calculated Error Curves for Three-point Ganging.

oscillator tuning capacitor, the values of C_1, C_2 and C_3 are 410·5, 134·58 and 58·74 $\mu\mu$F respectively. Replacing these in expressions 6.40, 6.41 and 6.42, and noting that f_{h1}, f_{h2} and f_{h3} are 1,079, 1,490 and 1,901 kc/s respectively, we find that

$$C_p = 601 \; \mu\mu\text{F}, \; C_t = 36{\cdot}5 \; \mu\mu\text{F} \text{ and } L_h = 77{\cdot}4 \; \mu\text{H}.$$

The calculated error curve is curve 1 of Fig. 6.19. A very considerable reduction of maximum ganging error (from 35 to 5 kc/s)

has been achieved by employing three instead of two preset components, and the form of the curve shows the use of the cubic expression for it to be justified.

Inspection of equations 6.39 indicates that variations of C_t affect the error curve at the high frequency end, whilst C_p controls the error curve mainly at the low-frequency end. This therefore suggests the procedure for ganging a receiver. The receiver should first be set to the highest required zero error signal frequency (f_{s3} = 1,436 kc/s), and C_t adjusted to give maximum audio output. The tuning is next changed to f_{s1} = 614 kc/s and C_p adjusted to give maximum output. Returning to 1,436 kc/s, C_t is readjusted if necessary and the procedure is repeated until the best results are obtained.

If we assume that this is the method employed for ganging we can calculate the error curve resulting from manufacturing errors in the value of L_h, and curves 2 and 3 (Fig. 6.19) show the result when L_h is 5% below and 5% above its correct value. High accuracy in calculating these errors is not easy to achieve as the difference between the actual and correct oscillator frequencies is very small (less than 1%) in comparison with the oscillator frequency itself. The values shown should therefore only be taken as indicative of the trend of the error curve against frequency. C_p and C_t have values of 665 and 41·5 $\mu\mu$F for L_h = 73·53 μH, and 546 and 32·9 $\mu\mu$F for L_h = 81·27 μH. In both cases maximum error is increased, and the central zero error frequency is changed to a lower value when L_h is 5% low and a higher value when L_h is 5% high. Curve 2 (L_h 5% low) has the advantage that the error is less at the low-frequency end of the range, where the signal circuits are usually more selective.

The effect of errors in the value of the tuning capacitance C cannot easily be assessed because it is unlikely that the error can be represented by a constant percentage over the wave range, as is possible for L_h.

6.13. Graphical Determination of the Oscillator Tracking Capacitances C_p and C_t.[6, 17] Graphical methods for determining C_p and C_t have been developed, and they are particularly helpful when a rapid calculation of C_p and C_t is required for a number of signal circuits having differing LC values. Their accuracy is, however, limited.

The method is well illustrated by taking the three numerical values C_1, C_2 and C_3 given above. The first step involves the construction of a diagram giving the total capacitance of C_1 and

C_p in series. Vertical lines, AB and DE in Fig. 6.20, are scaled in equal values of capacitance from A and D and are separated by any convenient distance. A and D are in the same horizontal straight line, AB is the C axis and DE the C_p axis. The C axis is

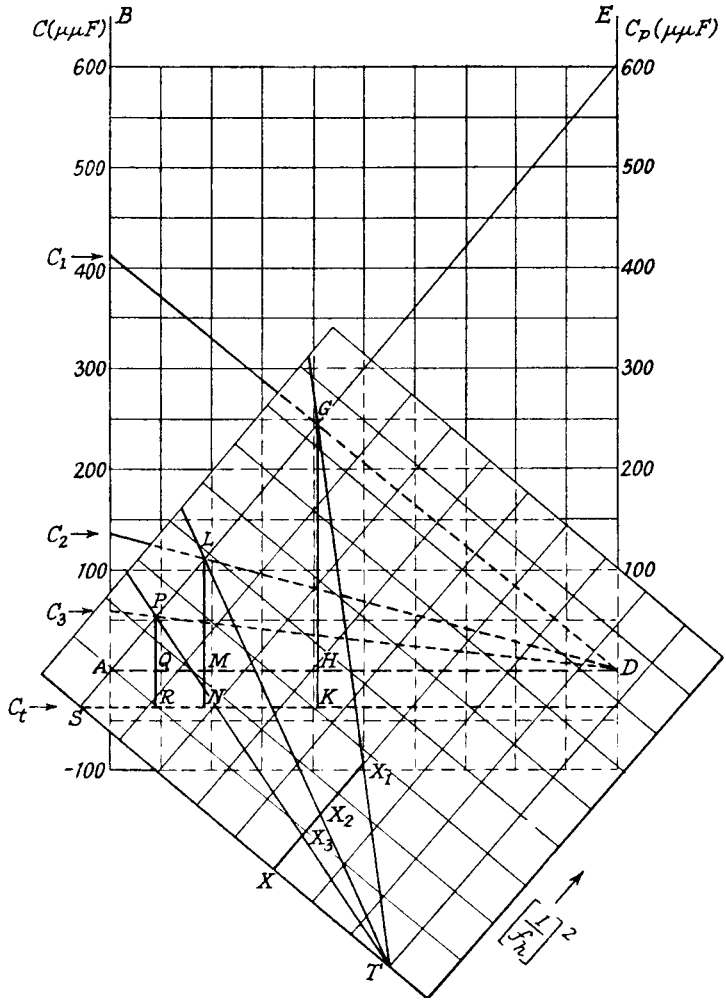


FIG. 6.20.—Graphical Construction for the Determination of the Padding and Trimming Capacitances.

projected below A and the reason for this is shown later. The length of a perpendicular on to the line AD from the intersection of lines drawn from A and D to particular points on the C_p and C scales respectively gives the capacitance for C_p and C connected

in series. Thus if $C = 410{\cdot}5\ \mu\mu F$ and $C_p = 601\ \mu\mu F$, we join D to $410{\cdot}5\ \mu\mu F$ on the C scale and A to $601\ \mu\mu F$ on the C_p scale and find that the capacitance of C and C_p in series is $GH = 244\ \mu\mu F$. The procedure for finding C_p and C_t is as follows. Join D with $C = C_1 = 410{\cdot}5\ \mu\mu F$, $C_2 = 134{\cdot}58$ and $C_3 = 58{\cdot}74\ \mu\mu F$; the line from A to the C_p scale cannot yet be drawn because C_p is unknown. We do, however, know that the total oscillator capacitances, $C_t + \frac{C_1 C_p}{C_1 + C_p}$, etc., must be proportional to $\left[\frac{1}{f_{h1}}\right]^2$, etc. The next step is therefore to draw three lines on transparent linear graph paper, the vertical intercepts XX_1, XX_2 and XX_3, of which from the base line ST are in the ratio of $\left[\frac{1}{f_{h1}}\right]^2 : \left[\frac{1}{f_{h2}}\right]^2 : \left[\frac{1}{f_{h3}}\right]^2$. The transparent sheet is placed over Fig. 6.20 and manœuvred until a " vertical " line on this sheet passes through A and the intersections P, L and G of the $\left[\frac{1}{f_{h3}}\right]^2$, etc., lines with DC_3, DC_2 and DC_1 lines. The point where this line SG cuts the C_p axis gives the required value of C_p, viz., $601\ \mu\mu F$. A line parallel to AD through S cuts the C axis below A, and this point gives the value of C_t, viz., $36{\cdot}5\ \mu\mu F$. The proof is as follows:

Triangles SPR, SLN and SGK are similar so

$$SP : SL : SG = \left[\frac{1}{f_{h3}}\right]^2 : \left[\frac{1}{f_{h2}}\right]^2 : \left[\frac{1}{f_{h1}}\right]^2$$
$$= PR : LN : GK$$
$$= (PQ + QR) : (LM + MN) : (GH + HK)$$

but $$QR = MN = HK = C_t.$$

and $$PQ = \frac{C_3 C_p}{C_3 + C_p}, \quad LM = \frac{C_2 C_p}{C_2 + C_p}, \quad GH = \frac{C_1 C_p}{C_1 + C_p}.$$

$$\therefore \left[\frac{1}{f_{h3}}\right]^2 : \left[\frac{1}{f_{h2}}\right]^2 : \left[\frac{1}{f_{h1}}\right]^2$$
$$= \left[C_t + \frac{C_3 C_p}{C_3 + C_p}\right] : \left[C_t + \frac{C_2 C_p}{C_2 + C_p}\right] : \left[C_t + \frac{C_1 C_p}{C_1 + C_p}\right].$$

The value of C_p can be found with reasonable accuracy, but C_t may be difficult to read exactly. The value of C_t may, however, be checked by substituting the graphical value of C_p in expression 6.41.

No particular difficulties are met in ganging band-spread receivers with signal tuning, and the design proceeds on the lines set out above, the formulae 6.40, 6.41 and 6.42 give the values of the

oscillator tuned circuit components. The values of the tuning capacitance for different signal frequencies are decided by the degree of band-spreading and may be calculated as indicated in Section 4.10.2.

6.14. Approximate Expressions for Ganged Oscillator Circuit Components for Different Intermediate Frequencies.[7] If the assumption is made that the oscillator tuning capacitance is equal to the total tuning signal capacitance, including strays and trimmer, expressions may be derived for the ratios $\frac{C_p}{C_{max.}}$, $\frac{C_t}{C_{max.}}$ and $\frac{L_h}{L_s}$, where $C_{max.}$ is the maximum value of total signal tuning capacitance at the minimum signal frequency, and L_h and L_s are the oscillator and signal inductances respectively. The above can then be given in terms of the ratios of the intermediate and the three zero error signal frequencies to the minimum signal frequency. The expressions defining the signal and oscillator frequencies for zero error are therefore

$$f_{s1} = \frac{1}{2\pi\sqrt{L_s C_1}} \qquad f_{h1} = \frac{1}{2\pi\sqrt{aL_s\left(C_t + \frac{C_p C_1}{C_p + C_1}\right)}}$$

$$f_{s2} = \frac{1}{2\pi\sqrt{L_s C_2}} \qquad f_{h2} = \frac{1}{2\pi\sqrt{aL_s\left(C_t + \frac{C_p C_2}{C_p + C_2}\right)}}$$

$$f_{s3} = \frac{1}{2\pi\sqrt{L_s C_3}} \qquad f_{h3} = \frac{1}{2\pi\sqrt{aL_s\left(C_t + \frac{C_p C_3}{C_p + C_3}\right)}}$$

where $L_h = aL_s$.
Combining the above expressions

$$\left[\frac{f_{s1}}{f_{h1}}\right]^2 = A_1 = \frac{a\left(C_t + \frac{C_p C_1}{C_p + C_1}\right)}{C_1}$$

$$\left[\frac{f_{s2}}{f_{h2}}\right]^2 = A_2 = \frac{a\left(C_t + \frac{C_p C_2}{C_p + C_2}\right)}{C_2}$$

$$\left[\frac{f_{s3}}{f_{h3}}\right]^2 = A_3 = \frac{a\left(C_t + \frac{C_p C_3}{C_p + C_3}\right)}{C_3}$$

Solving for C_p, a and C_t

$$C_p = \frac{A_1C_1^2(C_3 - C_2) + A_2C_2^2(C_1 - C_3) + A_3C_3^2(C_2 - C_1)}{A_1C_1(C_2 - C_3) + A_2C_2(C_3 - C_1) + A_3C_3(C_1 - C_2)} \quad . \quad 6.45a.$$

$$a = \frac{(A_1C_1 - A_2C_2)(C_p + C_1)(C_p + C_2)}{C_p^2(C_1 - C_2)} \quad . \quad . \quad . \quad . \quad 6.46a.$$

$$C_t = \frac{A_1C_1(C_p + C_1) - aC_pC_1}{a(C_p + C_1)} \quad . \quad . \quad . \quad . \quad . \quad . \quad 6.47a.$$

But $f_{min.} = \dfrac{1}{2\pi\sqrt{L_sC_{max.}}}$.

Writing

$$\left[\frac{f_{min.}}{f_{s1}}\right]^2 = B_1 = \frac{C_1}{C_{max.}}$$

$$\left[\frac{f_{min.}}{f_{s2}}\right]^2 = B_2 = \frac{C_2}{C_{max.}}$$

$$\left[\frac{f_{min.}}{f_{s3}}\right]^2 = B_3 = \frac{C_3}{C_{max.}}$$

and dividing the numerator of 6.45a by $(C_{max.})^3$ and the denominator by $(C_{max.})^2$, we have

$$\frac{C_p}{C_{max.}} = \frac{A_1B_1^2(B_3 - B_2) + A_2B_2^2(B_1 - B_3) + A_3B_3^2(B_2 - B_1)}{A_1B_1(B_2 - B_3) + A_2B_2(B_3 - B_1) + A_3B_3(B_1 - B_2)} \quad . \quad 6.45b.$$

But $$A_1 = \left[\frac{f_{s1}}{f_{h1}}\right]^2 = \frac{1}{\left[1 + \dfrac{f_1}{f_{s1}}\right]^2}$$

for $$f_{h1} = f_{s1} + f_1$$

where f_1 = the intermediate frequency.

$$\therefore\ A_1 = \frac{1}{\left[1 + \dfrac{f_{min.}}{f_{s1}} \cdot \dfrac{f_1}{f_{min.}}\right]^2}$$

$$= \frac{1}{(1 + \sqrt{B_1}.D)^2}$$

where $$D = \frac{f_1}{f_{min.}}.$$

Similarly $$A_2 = \frac{1}{[1 + \sqrt{B_2}.D]^2}$$

and $$A_3 = \frac{1}{[1 + \sqrt{B_3}.D]^2}.$$

Replacing A_1, A_2 and A_3 in 6.45*b*

$$\frac{C_p}{C_{max.}} = \frac{\dfrac{B_1^2(B_3-B_2)}{(1+D\sqrt{B_1})^2}+\dfrac{B_2^2(B_1-B_3)}{(1+D\sqrt{B_2})^2}+\dfrac{B_3^2(B_2-B_1)}{(1+D\sqrt{B_3})^2}}{\dfrac{B_1(B_2-B_3)}{(1+D\sqrt{B_1})^2}+\dfrac{B_2(B_3-B_1)}{(1+D\sqrt{B_2})^2}+\dfrac{B_3(B_1-B_2)}{(1+D\sqrt{B_3})^2}} \quad . \quad 6.45c.$$

Hence $\dfrac{C_p}{C_{max.}}$ is defined completely in terms of the ratios of the intermediate, and the three zero error frequencies to the minimum signal-tuning frequency.

Dividing the numerator and denominator in expression 6.46*a* by $(C_{max.})^3$.

$$a = \frac{\left[\dfrac{B_1}{(1+D\sqrt{B_1})^2}-\dfrac{B_2}{(1+D\sqrt{B_2})^2}\right](F+B_1)(F+B_2)}{F^2(B_1-B_2)} \quad . \quad 6.46b$$

where $F = \dfrac{C_p}{C_{max.}}$.

Dividing numerator by $(C_{max.})^2$ and denominator by $(C_{max.})$ in expression 6.47*a*.

$$\frac{C_t}{C_{max.}} = \frac{\dfrac{B_1}{(1+D\sqrt{B_1})^2}(F+B_1)-aFB_1}{a(F+B_1)} \quad . \quad 6.47b.$$

Thus a and $\dfrac{C_t}{C_{max.}}$ can also be defined in terms of the frequency ratios.

Expressions 6.45*c*, 46*b* and 47*b* are plotted in Fig. 6.21 against $\dfrac{f_1}{f_{min.}}$ for

$$\frac{f_{s1}}{f_{min.}} = 1{\cdot}091, \quad \frac{f_{s2}}{f_{min.}} = 1{\cdot}82, \quad \frac{f_{s3}}{f_{min.}} = 2{\cdot}545,$$

which means that for the medium wave range and $f_{min.} = 550$ kc/s, $f_{s1} = 600$ kc/s, $f_{s2} = 1{,}000$ kc/s, $f_{s3} = 1{,}400$ kc/s. These were the original zero error frequencies chosen by the authors of this method. From the curves in Fig. 6.21, the component ratios for an I.F. of 465 kc/s are $\left(\dfrac{f_1}{f_{min.}} = 0{\cdot}845\right)$

$$\frac{C_p}{C_{max.}} = 1{\cdot}067, \quad a = 0{\cdot}535, \quad \frac{C_t}{C_{max.}} = 0{\cdot}027$$

and if the maximum value of the total tuning capacitance is

536·8 $\mu\mu$F, and the signal inductance 156 μH as for the previous example in Section 6.12,

$$C_p = 572\ \mu\mu\text{F},\ L_h = 83{\cdot}5\ \mu\text{H},\ C_t = 14{\cdot}5\ \mu\mu\text{F}.$$

These are not very different from the values calculated in Section 6.12. The difference is partly due to the assumption that the oscillator tuning capacitor has a value equal to the total signal-

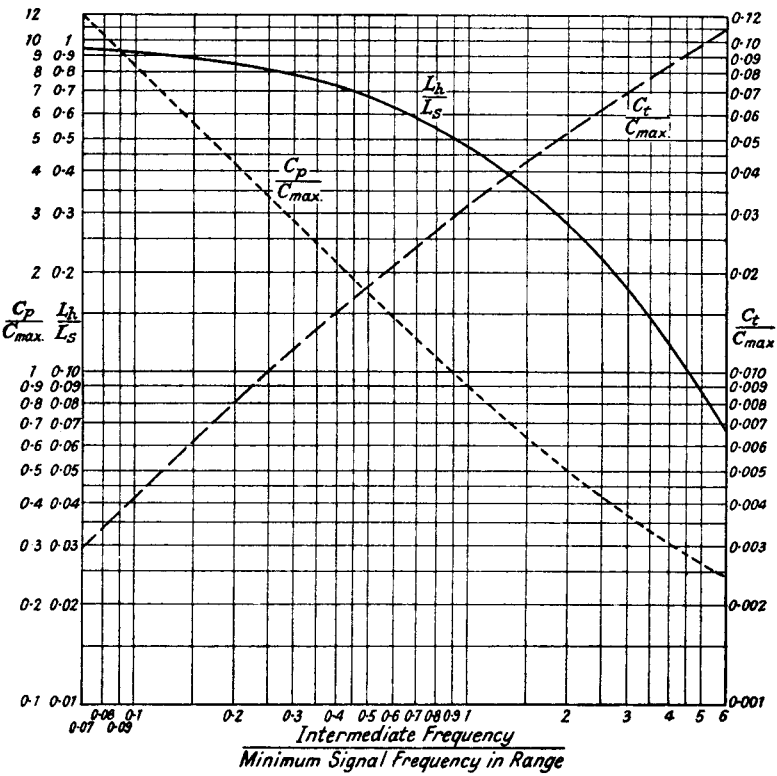


FIG. 6.21.—Curves for Ganging Component Ratios against Intermediate to Minimum Signal Frequency Ratios.

tuning capacitance, and partly to the fact that the zero error signal frequencies are different. If the latter are the same as in Section 6.12, viz., 614, 1,025 and 1,436 kc/s, the component values are

$$C_p = 578\ \mu\mu\text{F};\ L_h = 81{\cdot}2\ \mu\text{H},\ C_t = 13{\cdot}3\ \mu\mu\text{F}.$$

In spite of these differences satisfactory ganging is obtained. The great advantage of the curves is that they may be used to give

oscillator component values for any intermediate frequency in the normal range, and for any maximum value of tuning capacitance having a capacitance range factor not less than 7·43. The ratio range of the extreme zero error signal frequencies is 2·33 : 1, and if a smaller range is desired, as for band-spread reception, a new set of curves is necessary. The curves of Fig. 6.21 are suitable for the long wave (165 to 450 kc/s), the medium wave (550 to 1,500 kc/s) and the short wave band (5·5 to 15 Mc/s).

BIBLIOGRAPHY

1. Frequency Variations in Thermionic Generators. K. E. Edgeworth *Journal I.E.E.*, Vol. 64, 1926, p. 349.
2. On the Variation of Generated Frequency of a Triode Oscillator. K. B. Eller, *Proc. I.R.E.*, Dec. 1928, p. 1,706.
3. Constant Frequency Oscillators. F. B. Llewellyn, *Proc. I.R.E.*, Dec. 1931, p. 2,063.
4. A Recent Development in Vacuum Tube Oscillator Circuits. J. B. Dow. *Proc. I.R.E.*, Dec. 1931, p. 2,095. Correction *ibid.*, Jan. 1932, p. 182,
5. Ganging the Tuning Controls of a Superheterodyne Receiver. A. L. M. Sowerby, *Wireless Engineer*, Feb. 1932, p. 70.
6. The Padding Condenser. B. F. McNamee, *Electronics*, May 1932, p. 160.
7. A Solution of the Superheterodyne Tracking Problem. V. D. Landon and E. A. Seen, *Electronics*, Aug. 1932, p. 250.
8. *Valve Oscillators of Stable Frequency.* F. M. Colebrook, H.M. Stationery Office, Special Report No. 13, 1933.
9. Suppression of Interlocking in First Detector Circuits. P. W. Klipsch, *Proc. I.R.E.*, June 1934, p. 699.
10. The Non-Linear Theory of Electric Oscillations. B. van der Pol, *Proc. I.R.E.*, Sept. 1934, p. 1,051.
11. Oscillator Padding. H. Roder, *Radio Engineering*, March 1935, p. 7.
12. Ganging a Superhet. C. P. Singer, *Wireless Engineer*, June 1936, p. 307.
13. The Non-Linear Theory of the Maintenance of Oscillations. P. Le Corbeiller, *Journal I.E.E.*, Sept. 1936, p. 361.
14. Thermal Drift in Superheterodyne Receivers. J. Miller, *Electronics*, Nov. 1937, p. 24.
15. Tuning Drift. E. L. Gardiner, *Wireless World*, Jan. 6th, 1938, p. 6.
16. Short Wave Oscillator Problems. W. T. Cocking, *Wireless World*, Feb. 9th, 1939, p. 127.
17. Design of the Oscillator Circuit of Superhet Receivers. *Mullard Technical Bulletin*, Feb. 1939, p. 149.
18. Ganging Superheterodyne Receivers. M. Wald, *Wireless Engineer*, March 1940, p. 105, and April 1941, p. 146.
19. Ultra High Frequency Oscillator Frequency Stability Considerations. S. W. Seeley and E. I. Anderson, *R.C.A. Review*, July 1940, p. 77.
20. Ceramic Insulations for High Frequency Work. W. G. Robinson, *Journal I.E.E.*, Nov. 1940, p. 570.
21. *Theory and Design of Valve Oscillators.* H. A. Thomas, Messrs. Chapman and Hall. Text-book.

CHAPTER 7

INTERMEDIATE FREQUENCY AMPLIFICATION

7.1. Introduction. The intermediate frequency amplifier is a special case of a radio frequency amplifier operating at a fixed frequency. The values of L and C forming the tuned circuits are not limited by the need for a variable capacitance range, but may be chosen to give the best performance as regards selectivity and amplification. High selectivity demands high Q values for the coils, and for a given selectivity the required Q value is linearly proportional to frequency. For example, a Q of 50 at 110 kc/s gives the same selectivity as a Q of 211 at 465 kc/s. Stranded wire generally becomes necessary in order to obtain the required selectivity at high intermediate frequencies.

The amplification of an I.F. amplifier stage is largely fixed by the impedance of the I.F. tuned circuits, i.e., $\frac{L}{CR}$, and it is greatest when L is large and C small. There is a limit to the possible reduction of C, for valve and stray capacitances must not form a large proportion of the tuning capacitance; variation of valve capacitance due to A.G.C. and replacement of valves may otherwise cause serious detuning. Furthermore, these stray capacitances have resistive components which reduce the Q value and the impedance of the circuit. A typical value for C is between 100 and 400 $\mu\mu$F. Too small a capacitance value not only means greater liability to variation during service but also greater variation with temperature change, for stray capacitance usually has a high positive temperature coefficient.

Inductance or capacitance tuning may be employed, but the former is preferable because a variable inductance is less affected by temperature change than a variable capacitance. Moreover, fixed capacitors, of the silvered mica type, may be made with very small temperature coefficients (Section 6.7.3). Inductance tuning is usually accomplished by a movable screwed iron-dust core at the centre of the coil. When capacitance tuning is employed, the dielectric should be air as this gives a much lower resistive component than mica and allows greater mechanical stability. Pressure-type mica capacitors are generally unsatisfactory owing to a high positive temperature coefficient and poor mechanical stability.

Single-tuned circuits are not normally used in an I.F. amplifier

since they give a selectivity curve with a comparatively narrow pass-band and poor attenuation outside the pass-range. Coupled circuits give a more satisfactory curve, approximating to the ideal of a flat pass-band with sharper attenuation outside the band. Often they have identical L, C and R values, an important advantage in manufacture.

7.2. Types of Coupled Circuits.[15] Capacitance or inductance coupling between two tuned circuits may be used to give a band-pass selectivity curve. The coupling impedance may be in shunt connection (sometimes called common coupling), e.g., Z_1 in Fig. 7.1, or it may be in series connection (sometimes called top-end coupling), e.g., Z_2 in the same figure. Alternatively, both forms of coupling may be used as illustrated in Fig. 7.1. Combined coupling may occur unintentionally, and an example of this is provided by stray capacitance between two circuits having mutual inductance coupling.

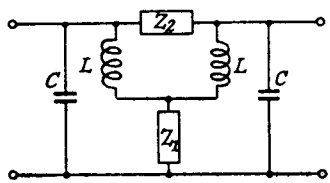


FIG. 7.1.—A Circuit with Combined Shunt and Series Coupling.

The I.F. valve is generally a pentode, so that the input source to the coupled circuits may conveniently be regarded as a generator of constant current $I_a = g_m E_g$. If E_2 is the voltage across the second tuned circuit, the selectivity curve is obtained by plotting the ratio $\frac{E_2}{I_a}$ against frequency. This ratio, $\frac{E_2}{I_a}$, often called the transfer impedance Z_T (note that $g_m E_g Z_T$ gives the secondary output voltage), may be calculated by normal procedure, but before doing so it is always advisable to estimate the frequency limits of the pass-band and any rejection frequencies. A rejection frequency can only occur for mixed inductive and capacitive coupling. If we assume the resistance components to be zero, the pass-band limit frequencies are obtained by making $\frac{E_2}{I_a} = \infty$, and the rejection frequencies by making $\frac{E_2}{I_a} = 0$. Calculation from the transfer impedance formula is laborious and a much simpler method (due to G. W. O. Howe [1, 2]) is possible.

Taking the shunt coupling impedance case, Howe shows that

one pass-band limit frequency is found by series resonance in the outer ring of impedances excluding Z_1 i.e., $\frac{1}{2\pi\sqrt{LC}}$. The second is obtained by separating the two circuits as in Fig. 7.2, the coupling impedance being doubled in the separation. This is logical because on joining up the two impedances $2Z_1$, we have Z_1. The second

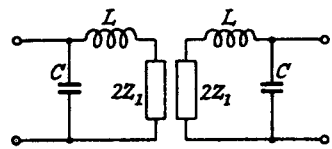


FIG. 7.2.—A Circuit with Separated Shunt Coupled Circuits.

pass-band frequency is given by series resonance of the LC circuit and $2Z_1$. Thus if $Z_1 = j\omega L_1$, the other limit frequency is

$$\frac{1}{2\pi\sqrt{(2L_1+L)C}}$$

or if $Z_1 = \frac{1}{j\omega C_1}$, it is

$$\frac{1}{2\pi\sqrt{\frac{LCC_1}{2C+C_1}}}$$

for the coupling capacitance C_1 must be halved to double Z_1.

It is important at this point to introduce the term,[4] coupling coefficient, k. This is defined, for shunt coupling, as the ratio of the coupling reactance to the sum of the coupling reactance and the reactance of the same form in either circuit. Thus for inductive coupling.

$$k = \frac{j\omega L_1}{j\omega(L+L_1)} = \frac{L_1}{L+L_1},$$

whilst for capacitance

$$k = \frac{\frac{1}{j\omega C_1}}{\frac{1}{j\omega C}+\frac{1}{j\omega C_1}} = \frac{C}{C+C_1}.$$

If the highest frequency is designated f_2, for inductive coupling

$$f_2 = \frac{1}{2\pi\sqrt{LC}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{(2L_1+L)C}}$$

and

$$\frac{f_2^2 - f_1^2}{f_2^2+f_1^2} = \frac{L_1}{L+L_1} = k.$$

If

$$f_m = \frac{1}{2\pi\sqrt{(L_1+L)C}}$$

(note f_m is the primary resonant frequency with the secondary open-circuited or vice versa)

$$f_2 = \frac{f_m}{\sqrt{\frac{L}{L+L_1}}} = \frac{f_m}{\sqrt{1-k}} \text{ and } f_1 = \frac{f_m}{\sqrt{1+k}}.$$

For capacitive coupling

$$f_2 = \frac{1}{2\pi\sqrt{\frac{LCC_1}{2C+C_1}}}, \qquad f_1 = \frac{1}{2\pi\sqrt{LC}}$$

and $$\frac{f_2^2 - f_1^2}{f_2^2+f_1^2} = \frac{C}{C+C_1} = k. \quad \text{If } f_m = \frac{1}{2\pi\sqrt{\frac{LCC_1}{C+C_1}}}.$$

(Again, f_m is the resonant frequency of the primary with the secondary open circuited.)

$$f_2 = \frac{f_m}{\sqrt{\frac{C+C_1}{2C+C_1}}} = f_m\sqrt{1+\frac{C}{C+C_1}} = f_m\sqrt{1+k}$$

and $$f_1 = f_m\sqrt{1-k}.$$

The results for f_2 and f_1 in terms of f_m and k are different from those for inductive coupling. The difference is explained by the fact that capacitive reactance is the inverse of inductive reactance and that C_1 and C are in series, so that the total capacitance is reduced.

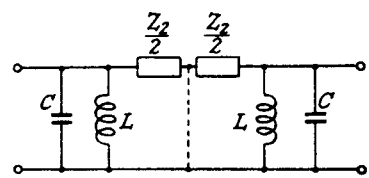


FIG. 7.3.—Separated Series Coupled Circuits.

For series coupling, the coupling impedance may be halved, as shown in Fig. 7.3, and the centre point joined to the junction of L and C. The two half-coupling impedances are in series, and when added make the total coupling impedance Z_2. The pass-band frequencies are given by resonance of the circuits, excluding $\frac{Z_2}{2}$, and then by including it. If $Z_2 = j\omega L_2$

$$f_2 = \frac{1}{2\pi\sqrt{\frac{LL_2C}{2L+L_2}}}; \quad f_1 = \frac{1}{2\pi\sqrt{LC}}$$

$$k = \frac{\text{Coupling susceptance } (L \text{ or } C)}{\text{circuit} + \text{coupling susceptance } (L \text{ or } C)}.$$

The coupling impedance is virtually in parallel with L and C, hence the use of susceptance to define it.

$$\therefore\ k = \frac{\frac{1}{j\omega L_2}}{\frac{1}{j\omega L}+\frac{1}{j\omega L_2}} = \frac{L}{L+L_2}$$

$$\frac{f_2^2 - f_1^2}{f_2^2+f_1^2} = \frac{L}{L+L_2} = k$$

$$f_m = \frac{1}{2\pi\sqrt{\frac{LL_2}{L+L_2}C}}$$

$$f_2 = f_m\sqrt{\frac{2L+L_2}{L+L_2}} = f_m\sqrt{1+k}$$

$$f_1 = f_m\sqrt{1-k}.$$

The coupling inductance is in parallel with the tuning inductance L so that the total inductance is decreased, and f_2 and f_1 in terms of f_m and k are identical with the expressions for shunt capacitance coupling. When $Z_2 = \frac{1}{j\omega C_2}$

$$f_2 = \frac{1}{2\pi\sqrt{LC}};\quad f_1 = \frac{1}{2\pi\sqrt{L(C+2C_2)}}$$

$$k = \frac{j\omega C_2}{j\omega C+j\omega C_2} = \frac{C_2}{C+C_2} = \frac{f_2^2 - f_1^2}{f_2^2+f_1^2}$$

$$f_m = \frac{1}{2\pi\sqrt{L(C+C_2)}}$$

$$f_2 = \frac{f_m}{\sqrt{1-k}};\quad f_1 = \frac{f_m}{\sqrt{1+k}}.$$

These last expressions are identical with those for shunt inductance coupling.

In all the above examples we see that only one of the limit frequencies is affected, and the results are best summarized in the form of a table.

Type of Coupling.	Frequency Peak Affected.	Effect of increase in Component on Frequency.	Peak Separation.
1. Shunt inductance	lower	decreases	increases
2. Shunt capacitance	higher	decreases	decreases
3. Series inductance	higher	decreases	decreases
4. Series capacitance	lower	decreases	increases

All possible forms of coupling have not been exhausted and there is yet mutual inductance coupling. The equivalent circuit is that of Fig. 7.4*a*, and the limit frequencies are

$$f_2 = \frac{1}{2\pi\sqrt{(L-M)C}}, \qquad f_1 = \frac{1}{2\pi\sqrt{(L+M)C}}$$

$$k = \frac{M}{L} = \frac{f_2^2 - f_1^2}{f_2^2 + f_1^2}$$

$$f_m = \frac{1}{2\pi\sqrt{LC}}, \quad f_2 = \frac{f_m}{\sqrt{1-k}}, \quad f_1 = \frac{f_m}{\sqrt{1+k}}.$$

Both limit frequencies are dependent on the coupling impedance and the pass-band range increases as M increases. If k is small ($< 0{\cdot}1$), $f_2 - f_m \simeq f_m - f_1$, so that symmetrical variable selectivity may be obtained by variation of M. With L and C coupling variation, f_m varies and the I.F. carrier must be varied if it is to be located in the centre of the pass-band, i.e., asymmetrical variable selectivity is obtained.

In Fig. 7.4*a* the mutual inductance is given a positive value,

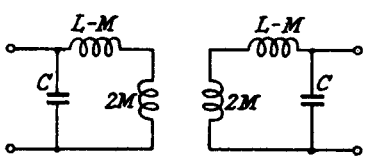


FIG. 7.4*a*.—Mutual Inductance Coupling with the Circuits Separated.

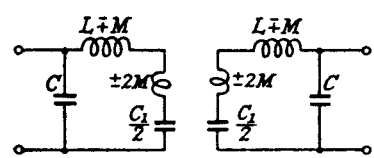


FIG. 7.4*b*.—Combined Mutual Inductance and Shunt Capacitance Coupling with the Circuits separated.

but it might equally be designated as $-M$. The horizontal inductances then become $(L+M)$ and the frequencies f_1 and f_2 merely reverse values. As long as no other coupling exists between the circuits the sign of M is unimportant. If, however, there is capacitance coupling, as in Fig. 7.4*b*, the sign of M is important. When M is positive the coupling impedance is $j\left(\omega M - \frac{1}{\omega C_1}\right)$ and the limit frequencies are

$$f_2 = \frac{1}{2\pi\sqrt{\frac{(L+M)CC_1}{2C+C_1}}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{(L-M)C}}$$

with a rejection frequency

$$f_3 = \frac{1}{2\pi\sqrt{MC_1}}$$

produced by series resonance of the coupling arm. If M is negative $Z_1 = -j\left(\omega M + \frac{1}{\omega C_1}\right)$ and

$$f_2 = \frac{1}{2\pi\sqrt{\frac{(L-M)CC_1}{2C+C_1}}} \text{ and } f_1 = \frac{1}{2\pi\sqrt{(L+M)C}},$$

but there is no rejection frequency, for f_3 is imaginary. The sign of M is changed by reversing the primary or secondary coil connections as explained in Section 3.4.2.

A typical selectivity curve for positive M is shown in Fig. 7.4c, the rejection frequency f_3 can be varied by changing C_1 and can occur at any desired position. It may be used to increase attenua-

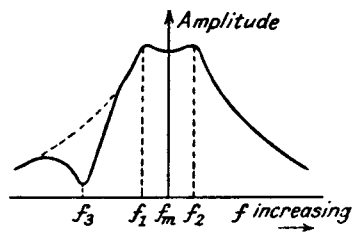


FIG. 7.4c.—A Typical Selectivity Curve for Combined Positive Mutual Inductance and Shunt Capacitance Coupling

tion at the edge of the pass-band, though the tendency is then to reduce the pass-band width and response nearest f_3 and to cause asymmetrical pass response.

If there is series capacitance as well as mutual inductance coupling, the equivalent circuit is that of Fig. 7.5. One limit frequency is obtained from resonance of the circuit excluding the mutual inductance (C_2 is doubled). Thus when M is positive

$$f_2 = \frac{1}{2\pi\sqrt{(L-M)(C+2C_2)}}.$$

The other limit frequency is given by resonance of the circuit excluding the series capacitance. The mutual inductance is doubled in separation and

$$f_1 = \frac{1}{2\pi\sqrt{(L+M)C}}.$$

Since there is mixed coupling, a rejection frequency is possible and its value may be found as follows.

For zero voltage across the second capacitance in Fig. 7.5

$$I_2 j\omega_3(L - M) + I_1 j\omega_3 M = 0$$

or
$$I_1 = -I_2\frac{(L - M)}{M}$$

also

$$\frac{I_2}{j\omega_3 C_2} + I_2 j\omega_3(L - M) + (I_2 - I_1)j\omega_3(L - M) = 0$$

$$jI_2\left[\omega_3\left(2(L - M) + \frac{(L - M)^2}{M}\right) - \frac{1}{\omega_3 C_2}\right] = 0$$

from which
$$f_3 = \frac{1}{2\pi\sqrt{\frac{(L+M)(L - M)}{M}C_2}}.$$

The selectivity curve is similar to that shown in Fig. 7.4*c*. The frequencies for negative M are found by replacing $+M$ by $-M$ in

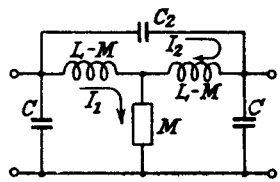


FIG. 7.5.—A Circuit with Combined Positive Mutual Inductance and Shunt Capacitance Coupling.

the above expressions and, as for shunt capacitance and mutual inductance coupling, there is no rejection frequency.

We see from this preliminary examination that variable pass-band width is most satisfactorily obtained by mutual inductance coupling, because increase of M moves both limit frequencies away from the mid-frequency. We will therefore proceed to a detailed analysis of the I.F. tuned transformer with mutual inductance coupling.

7.3. The Design of an I.F. Transformer with Mutual Inductance Coupling. The equivalent circuit diagram for the double-tuned transformer with mutual inductance coupling is shown in Fig. 7.6 ; the valve is represented as a constant current generator of $I_a = g_m E_g$, and its resistance is assumed to be much greater than the impedance of the primary of the transformer. The procedure for calculating the transfer impedance $Z_T\left(\frac{E_2}{I_a}\right)$ is considerably simplified by replacing the actual circuits by Z_1, Z_2, etc.

Then $$\frac{E_2}{E_1} = \frac{Z_5}{Z_4+Z_5}$$

$$\frac{E_1}{E} = \frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5} \Big/ \frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5}+Z_2$$

$$\frac{E_2}{E} = \frac{Z_3Z_5}{Z_3(Z_4+Z_5)+Z_2(Z_3+Z_4+Z_5)}.$$

But $$E = I_aZ_1\left[Z_2+\frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5}\right] \Big/ Z_1+Z_2+\frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5}.$$

Hence $$\frac{E_2}{I_a} = \frac{Z_1Z_3Z_5}{(Z_1+Z_2)(Z_3+Z_4+Z_5)+Z_3(Z_4+Z_5)}$$

or $$Z_T = \frac{Z_1Z_3Z_5}{(Z_1+Z_2+Z_3)(Z_3+Z_4+Z_5)-Z_3^2} \quad . \quad . \quad 7.1a$$

$$= \frac{Z_1Z_3Z_5}{Z_{s1}.Z_{s2}-Z_3^2} \quad . \quad . \quad . \quad . \quad . \quad . \quad 7.1b$$

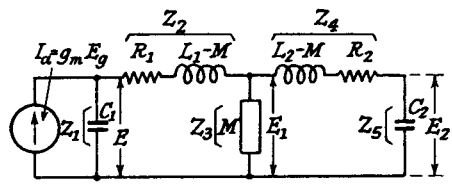


FIG. 7.6.—The Equivalent Circuit for a Double Tuned I.F. Transformer.

where Z_{s1} and Z_{s2} are the series impedances of first and second circuits taken separately, i.e.,

$$Z_{s1} = Z_1+Z_2+Z_3 \text{ and } Z_{s2} = Z_3+Z_4+Z_5.$$

But $$Z_1 = \frac{1}{j\omega C_1}, \quad Z_2 = R_1+j\omega(L_1-M), \quad Z_3 = j\omega M$$

$$Z_4 = R_2+j\omega(L_2-M) \text{ and } Z_5 = \frac{1}{j\omega C_2}.$$

It is assumed that the resistance components in the capacitive branches are negligible (this is usually true of air dielectric capacitors or of good quality mica fixed capacitors).

$$Z_{s1} = Z_1+Z_2+Z_3 = R_1+j\left(\omega L_1-\frac{1}{\omega C_1}\right)$$

$$= R_1(1+jQ_1F)$$

where $$Q_1 = \frac{\omega_m L_1}{R_1}$$

$$F = \frac{2\Delta f}{f_m} \text{ (see Section 4.2.3)}$$

and f_m = central or mid-band frequency

$$= \frac{1}{2\pi\sqrt{L_1C_1}} = \frac{1}{2\pi\sqrt{L_2C_2}}.$$

$$Z_{s2} = R_2 + j\left(\omega L_2 - \frac{1}{\omega C_2}\right)$$

$$= R_2(1 + jQ_2F).$$

Hence
$$Z_T = \frac{-\dfrac{j\omega M}{\omega^2 C_1C_2}}{R_1R_2(1+jQ_1F)(1+jQ_2F)+\omega^2M^2}$$

$$= \frac{-\dfrac{j\omega M}{\omega^2 C_1C_2R_1R_2}}{(1+jQ_1F)(1+jQ_2F)+\dfrac{\omega^2M^2}{R_1R_2}}.$$

But $k = \dfrac{M}{\sqrt{L_1L_2}}$ and $\dfrac{\omega^2M^2}{R_1R_2} = \dfrac{\omega_mL_1}{R_1}\cdot\dfrac{\omega_mL_2}{R_2}\cdot\dfrac{M^2}{L_1L_2}\cdot\left(\dfrac{\omega}{\omega_m}\right)^2 \simeq Q_1Q_2k^2$

$$\text{if } \frac{\omega}{\omega_m} \simeq 1$$

$$\frac{\omega M}{\omega^2 C_1C_2R_1R_2} = \frac{\omega_m}{\omega}\left[\frac{M}{\omega_m C_1C_2R_1R_2}\right] = \frac{\omega_m}{\omega}\left[\frac{\omega^3{}_mL_1L_2M}{R_1R_2}\right]$$

for $$\omega_mL_1 = \frac{1}{\omega_mC_1} \text{ and } \omega_mL_2 = \frac{1}{\omega_mC_2}.$$

$$\therefore \frac{\omega M}{\omega^2 C_1C_2R_1R_2} = \frac{\omega_m}{\omega}\left[\frac{\omega_m{}^2L_1{}^2}{R_1}\cdot\frac{\omega_mL_2}{R_2}\cdot\frac{M_1}{L_1}\right]$$

$$= \frac{\omega_m}{\omega}\left[R_{D1}\cdot Q_2\cdot\frac{M}{\sqrt{L_1L_2}}\cdot\sqrt{\frac{L_2}{L_1}}\right]$$

$$= \frac{\omega_m}{\omega}\left[R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}\right]$$

where R_{D1} = dynamic resistance of the primary when not coupled to the secondary circuit.

Generally we shall be concerned only with frequencies close to f_m (within $\pm$ 30 kc/s) so that $\dfrac{\omega_m}{\omega} \simeq 1$ and

$$\frac{\omega M}{\omega^2 C_1C_2R_1R_2} \simeq R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}$$

Thus
$$Z_T = \frac{-jR_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{(1+jQ_1F)(1+jQ_2F)+Q_1Q_2k^2} \quad . \quad . \quad 7.2a$$

$$|Z_T| = \frac{R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{\sqrt{[1+Q_1Q_2(k^2-F^2)]^2+(Q_1+Q_2)^2F^2}} \quad . \quad 7.2b.$$

The conditions for maximum value of $|Z_T|$ are found by differentiating 7.2*b* with respect to F and equating to zero.

Thus $2[1+Q_1Q_2(k^2-F^2)](-2Q_1Q_2F)+2(Q_1+Q_2)^2F=0$

or
$$F_{(max.\,|Z_T|)} = \pm\sqrt{k^2-\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}.$$

We have now to estimate the importance of k, Q_1 and Q_2 in determining the behaviour of the coupled circuits. Three conditions arise, viz.,

$$k<\sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)};\ k=\sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)} \text{ and } k>\sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}.$$

When $k<\sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}$, $F_{(max.\,|Z_T|)}$ is imaginary and a maximum occurs at $F=0$, when

$$|Z_T| = \frac{R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{1+Q_1Q_2k^2} \quad . \quad . \quad . \quad 7.3a.$$

For $k=\sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}$; $F_{(max.\,|Z_T|)}=0$ and

$$|Z_T| = \frac{R_{D1}\sqrt{\frac{1}{2}\left(\frac{Q_2^2}{Q_1^2}+1\right)}\sqrt{\frac{L_2}{L_1}}}{1+\frac{Q_1^2+Q_2^2}{2Q_1Q_2}} \quad . \quad . \quad 7.3b.$$

The above value of k represents critical coupling and greater values produce a double-peaked response.

When $Q_1=Q_2$

$$|Z_T| = \frac{R_{D1}\sqrt{\frac{L_2}{L_1}}}{2} \quad . \quad . \quad . \quad . \quad 7.3c.$$

Two maxima result when $k > \sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}$

and $$F_{(max.)} = \pm\sqrt{k^2 - \frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}.$$

This corresponds to off-tune frequencies

$$\Delta f_{(max.\ |Z_T|)} \text{ of } \pm\frac{f_m}{2}\sqrt{k^2 - \frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}.$$

Replacing F by $\pm\sqrt{k^2 - \frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}$ in expression 7.2*b* gives after simplification

$$|Z_T| = \frac{R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{\sqrt{1 - \frac{(Q_1^2+Q_2^2)^2}{4Q_1^2Q_2^2}+(Q_1+Q_2)^2k^2}} \quad . \quad 7.3d.$$

This reduces to

$$|Z_T| = \frac{R_{D1}\sqrt{\frac{L_2}{L_1}}}{2}$$

when $Q_1 = Q_2$, which is the same result as for $k = \sqrt{\frac{1}{2}\left(\frac{1}{Q_1^2}+\frac{1}{Q_2^2}\right)}$, i.e., after critical coupling is exceeded two-peaked response is obtained, but the two maxima have amplitudes equal to the amplitude at critical coupling. Increased coupling merely increases the frequency separation of the two peaks, but does not change their amplitude unless coupling is very large when the term $\frac{\omega_m}{\omega}$ in the expression $\frac{\omega M}{\omega^2 C_1 C_2 R_1 R_2}$ given above can no longer be neglected. The effect of this term is to increase the lower frequency maximum and decrease the higher frequency maximum. For couplings much less than critical, the transfer impedance at any frequency is

$$|Z_T| = \frac{R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{\sqrt{(1+Q_1Q_2(-F^2))^2+(Q_1+Q_2)^2F^2}}$$

$$= \frac{R_{D1}Q_2k\sqrt{\frac{L_2}{L_1}}}{\sqrt{(1+Q_1^2F^2)(1+Q_2^2F^2)}}.$$

The selectivity characteristic (Section 4.2.3), or the variation of $|Z_T|$ with respect to F, is equivalent to that of two circuits of Q_1 and Q_2 separated by a valve. The transfer impedance is however directly proportional to k, so that under these conditions amplification (in association with a valve) is low.

In most cases the primary and secondary circuits are identical and the expression for transfer impedance becomes

$$|Z_T| = \frac{R_D Qk}{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}} \quad . \quad . \quad 7.2c$$

and its maximum value is obtained when

$$F = \pm\sqrt{k^2 - \frac{1}{Q^2}}$$

when $k < \frac{1}{Q}$, $F_{(max.\,|Z_T|)} = 0$ and

$$|Z_T| = \frac{R_D Qk}{1+Q^2k^2} \quad . \quad . \quad . \quad . \quad 7.3e.$$

For $k = \frac{1}{Q}$, $F_{(max.\,|Z_T|)} = 0$

$$|Z_T| = \frac{R_D}{2} \quad . \quad . \quad . \quad . \quad 7.3f$$

and $k > \frac{1}{Q}$ gives $F_{(max.\,|Z_T|)} = \pm\sqrt{k^2 - \frac{1}{Q^2}}$ and

$$|Z_T| = \frac{R_D}{2}.$$

It may be noted that $Qk = \frac{\omega L}{R} \cdot \frac{M}{L} = \frac{\omega M}{R} = \frac{\text{coupling reactance}}{\text{coil resistance}}$

This shows that the maximum amplification obtainable from a pair of coupled circuits is only one-half that obtained with one circuit. The loss of amplification is however offset by increased selectivity and pass-band response.

7.4. Generalized Selectivity Curves for Mutual Inductance Coupling. A series of generalized selectivity curves [5] may be obtained for a pair of identical coupled circuits by plotting $20 \log_{10} \frac{|Z_T|}{|Z_T|_{max.}}$ *db* against QF for selected values of Qk. The reference level, 0*db*, is conveniently taken as $|Z_T|_{(max.)}$, i.e., $\frac{R_D}{2}$ and the decibel loss at any particular frequency f is

$$20 \log_{10} \frac{R_D}{2|Z_T|_f} = 20 \log_{10} \frac{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}}{2Qk} \quad 7.4.$$

This expression is independent of the sign of QF and it means that the selectivity curve is symmetrical about f_m. By plotting expression 7.4 against QF (to a logarithmic scale) we have the generalized curves of Fig. 7.7. The great advantage of these curves lies in the ease with which QF may be converted to the off-tune frequency Δf as soon as we have fixed Q and f_m. As indicated in

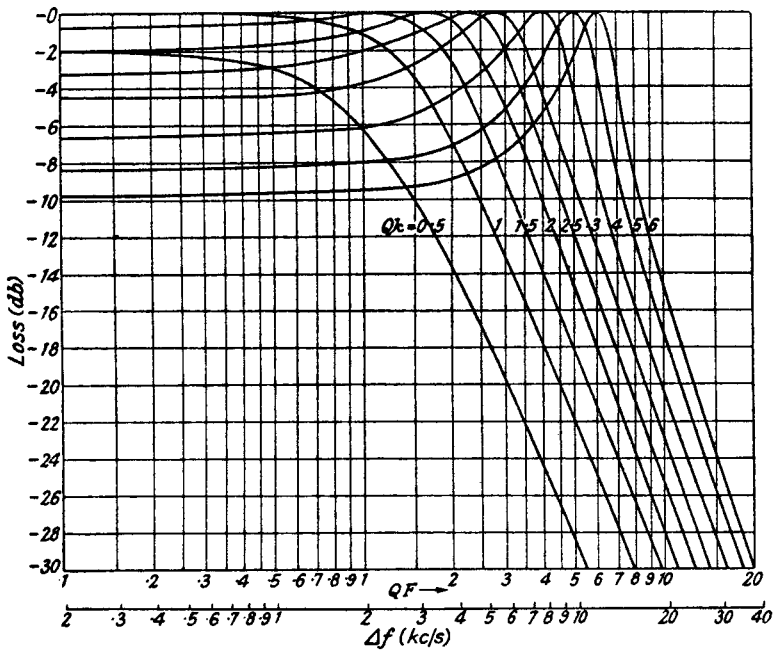


FIG. 7.7.—Generalized Selectivity Curves for an I.F. Transformer with two Identical Coupled Circuits.

Section 4.2.3, to obtain the selectivity curve against off-tune frequency it is only necessary to slide a logarithmic Δf scale underneath the QF scale until $\frac{f_m}{2Q}$ on the former corresponds to $QF = 1$. As an example of the use of these curves, let us consider coupled circuits having $Q = 120$ and $k = 0{\cdot}025$ at an I.F. of $f_m = 465$ kc/s.

$$Qk = 3 \text{ and } \frac{f_m}{2Q} = 1{\cdot}94.$$

The logarithmic Δf scale is adjusted so that $\Delta f = 1{\cdot}94$ lies immediately under $Qk = 1$ (Fig. 7.7). The decibel loss at any off-tune frequency is read directly from the curve $Qk = 3$. Thus

maximum $|Z_T|$ occurs at $\Delta f = 5{\cdot}5$ kc/s, whilst at $\Delta f = 15$ kc/s the frequency response is reduced by 18 dbs.

The assumptions made in order to obtain the generalized curves lead to only small errors with normal couplings at an I.F. of 465 kc/s or over and off-tune frequencies up to ± 50 kc/s. For lower intermediate frequencies (110 kc/s) and similar off-tune frequency limits, the generalized curves are useful but their accuracy is limited.

7.5. Generalized Selectivity Curves for Shunt and Series Coupling. Let us again assume that the primary and secondary circuits are identical, and that shunt capacitance (C_1) coupling is employed. The transfer impedance is found by inserting the following values for Z_1, Z_2, etc., in expression 7.1*b*.

$$Z_1 = Z_5 = \frac{1}{j\omega C}; \quad Z_3 = \frac{1}{j\omega C_1}$$

$$Z_{s1} = Z_{s2} = R + j\left(\omega L - \frac{1}{\omega C C_1 / (C+C_1)}\right)$$

From Section 7.2

$$\omega_m = \frac{1}{\sqrt{\frac{LCC_1}{C+C_1}}} \quad \text{and} \quad k = \frac{C}{C+C_1}$$

$$\therefore Z_{s1} = Z_{s2} = R(1+jQF)$$

and

$$Z_T = \frac{-\frac{1}{j\omega^3 C_1 C^2 R^2}}{(1+jQF)^2 + \frac{1}{\omega^2 C_1{}^2 R^2}} \qquad 7.5$$

$$\frac{1}{\omega^3 C_1 C^2 R^2} = \left(\frac{\omega_m}{\omega}\right)^3 \frac{C+C_1}{\omega_m C C_1 R} \cdot \frac{C}{C+C_1} \cdot \frac{(C+C_1)^2}{\omega_m{}^2 C^2 C_1{}^2 R} \cdot \frac{C_1{}^2}{(C+C_1)^2}$$

$$= \left(\frac{\omega_m}{\omega}\right)^3 . QkR_D \frac{C_1{}^2}{(C+C_1)^2}$$

$$\simeq QkR_D$$

since $\frac{\omega_m}{\omega}$ over the normal response range is nearly equal to unity and $C_1 \gg C$.

Similarly

$$\frac{1}{\omega^2 C_1{}^2 R^2} = \left(\frac{\omega_m}{\omega}\right)^2 \left[\frac{C+C_1}{\omega_m C C_1 R} \cdot \frac{C}{C+C_1}\right]^2$$

$$= \frac{\omega_m{}^2}{\omega^2} Q^2 k^2$$

$$\simeq Q^2 k^2.$$

Hence $$|Z_T| = \frac{R_D Q k}{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}}$$

which is identical with expression 7.2c. The generalized selectivity curves are therefore applicable to this form of coupling so long as we note that

$$f_m = \frac{1}{2\pi\sqrt{\frac{LCC_1}{C+C_1}}},$$

i.e., it varies as the coupling $k = \frac{C}{C+C_1}$ is varied.

The curves may also be used to estimate the performance of shunt inductance coupling L_1 and in this instance

$$f_m = \frac{1}{2\pi\sqrt{(L+L_1)C}} \text{ and } k = \frac{L_1}{L+L_1}.$$

The generalized curves are also applicable to series coupling provided the coupling reactance is large compared with the reactance of the same kind in either of the tuned circuits. For series capacitance coupling

$$f_m = \frac{1}{2\pi\sqrt{L(C+C_2)}} \text{ and } k = \frac{C_2}{C+C_2}$$

and for series inductance

$$f_m = \frac{1}{2\pi\sqrt{\frac{LL_2}{L+L_2}C}} \text{ and } k = \frac{L}{L+L_2}.$$

7.6. The Impedance of the Primary of Two Coupled Circuits. In the solution of certain problems, e.g., the calculation of the voltage available for A.G.C., it is necessary to know the impedance (Z_p) of the primary of two coupled circuits. Referring to Fig. 7.6

$$Z_p = \frac{Z_1 Z}{Z_1+Z} \quad . \quad . \quad . \quad . \quad . \quad 7.6$$

where $$Z = Z_2 + \frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5}$$

$$\therefore Z_p = \frac{Z_1[Z_2(Z_3+Z_4+Z_5)+Z_3(Z_4+Z_5)]}{(Z_1+Z_2)(Z_3+Z_4+Z_5)+Z_3(Z_4+Z_5)}$$

$$= \frac{Z_1[(Z_2+Z_3)(Z_3+Z_4+Z_5)-Z_3^2]}{(Z_1+Z_2+Z_3)(Z_3+Z_4+Z_5)-Z_3^2}.$$

Replacing Z_1, Z_2, etc., by their appropriate components

$$Z_p = \frac{\frac{1}{j\omega C_1}[R_1R_2(1+jQ_1)(1+jQ_2F)+R_1R_2Q_1Q_2k^2]}{R_1R_2[(1+jQ_1F)(1+jQ_2F)+Q_1Q_2k^2]}$$

where $Q_1 = \frac{\omega_m L_1}{R_1}$, $Q_2 = \frac{\omega_m L_2}{R_2}$ and $\frac{\omega}{\omega_m} \simeq 1$.

Since $Q_1 \gg 1$ and $\gg Q_1Q_2k^2$

$$Z_p \simeq \frac{Q_1(1+jQ_2F)}{\omega C_1[(1+jQ_1F)(1+jQ_2F)+Q_1Q_2k^2]}$$

$$\simeq \frac{\frac{\omega_m}{\omega}R_{D1}(1+jQ_2F)}{(1+jQ_1F)(1+jQ_2F)+Q_1Q_2k^2}$$

$$\simeq \frac{R_{D1}(1+jQ_2F)}{(1+jQ_1F)(1+jQ_2F)+Q_1Q_2k^2} \quad . \quad . \quad 7.7a$$

when $\omega_m \simeq \omega$

$$|Z_p| = \frac{R_{D1}\sqrt{1+Q_2{}^2F^2}}{\sqrt{[1+Q_1Q_2(k^2-F^2)]^2+(Q_1+Q_2)^2F^2}} \quad . \quad 7.7b$$

$$= \frac{|Z_T|\sqrt{1+Q_2{}^2F^2}}{Q_2k\sqrt{\frac{L_2}{L_1}}} \quad . \quad . \quad . \quad . \quad 7.7c.$$

When the primary and secondary circuits are identical

$$Z_p = \frac{R_D(1+jQF)}{(1+jQF)^2+Q^2k^2} \quad . \quad . \quad . \quad 7.7d.$$

At the mid-frequency f_m, $F = 0$ and

$$|Z_p|f_m = \frac{R_{D1}}{1+Q_1Q_2k^2},$$

from which we see that the primary impedance decreases progressively as k is increased. At critical coupling

$$k = \sqrt{\frac{1}{2}\left(\frac{1}{Q_1{}^2}+\frac{1}{Q_2{}^2}\right)}$$

$$|Z_p|f_m = \frac{R_{D1}}{1+\frac{1}{2}\left(\frac{Q_2}{Q_1}+\frac{Q_1}{Q_2}\right)}$$

$$\simeq \frac{R_{D1}}{2}$$

when $Q_1 \simeq Q_2$.

It is sometimes helpful to use the equivalent primary circuit developed from

$$Z_p = \frac{Z_1 Z}{Z_1 + Z}.$$

Now
$$Z = Z_2 + \frac{Z_3(Z_4+Z_5)}{Z_3+Z_4+Z_5}$$

$$= Z_2 + Z_3 - \frac{Z_3^2}{Z_3+Z_4+Z_5} = Z_2 + Z_3 - \frac{Z_3^2}{Z_{s2}}.$$

The term $-\dfrac{Z_3^2}{Z_3+Z_4+Z_5}$ can be separated into a resistive and reactive component connected in series.

$$-\frac{Z_3^2}{Z_3+Z_4+Z_5} = \frac{\omega^2 M^2}{R_2(1+jQ_2F)} = \frac{\omega^2 M^2}{R_2(1+Q_2^2F^2)} - \frac{j\omega^2 M^2 Q_2 F}{R_2(1+Q_2^2F^2)}.$$

This result may be compared with that indicated by expression 3.18 in Section 3.4.2. The equivalent primary circuit is shown in

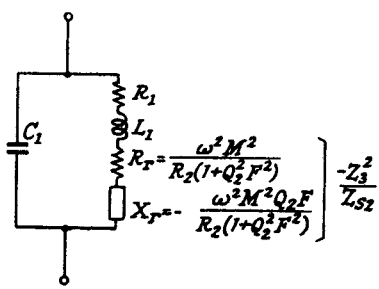


FIG. 7.8.—The Equivalent Circuit for the Primary of an I.F. Transformer.

Fig. 7.8, the total resistance and reactance in the inductive arm becoming

$$R = R_1 + R_r$$

$$= R_1\left[1 + \frac{\omega^2 M^2}{R_1 R_2(1+Q_2^2F^2)}\right]$$

$$\simeq R_1\left[1 + \frac{Q_1 Q_2 k^2}{1+Q_2^2F^2}\right] \quad . \quad . \quad . \quad 7.8a.$$

$$X = j(\omega L_1 + X_r)$$

$$\simeq j\omega L_1\left[1 - \frac{Q_2^2 k^2 F}{1+Q_2^2F^2}\right] \quad . \quad . \quad . \quad 7.8b$$

when $\dfrac{\omega_m}{\omega} \simeq 1$.

7.7. Variable Selectivity.

7.7.1. Introduction. The degree of selectivity required to suppress adjacent channel interference depends to a large extent on the strength of the received signal. For local station reception the selectivity can be reduced with consequent improvement in fidelity, but for distant station reception very high selectivity is often necessary. Hence some means of varying selectivity is an advantage, and in a superheterodyne receiver a very simple method is to control the selectivity of the I.F. amplifier.

Variable selectivity may conveniently be divided into asymmetrical (one side-band only is affected) and symmetrical variation (both side-bands are affected). Each has its merits and their advantages are listed below.

Asymmetrical.	*Symmetrical.*
1. One side-band only is reduced, that suffering from interference.	1. Both side-bands are reduced.
2. Tuning is critical and the carrier must be located correctly.	2. Tuning is not very critical.
3. Amplitude (harmonic) distortion can be high if not tuned correctly.	3. Amplitude distortion is small.
4. Frequency distortion is not so appreciable as only the high frequencies in one side-band are affected.	4. Frequency distortion may be considerable because the high frequencies in both side-bands are reduced.
5. Selectivity is better for given fidelity (audio frequency response) when interference is on one side of the carrier.	5. Selectivity is better when interference occurs on both sides of the carrier.
6. Signal to noise ratio is approximately − 3 dbs. down on that for symmetrical variable selectivity.	

Amplitude distortion due to asymmetrical variable selectivity can be reduced to a low value if there is symmetrical response for side-bands up to $\pm 1{,}500$ c.p.s.[6] on either side of the carrier. The audio frequency output with two side-bands is +6 dbs. above that with one, whilst noise is only +3 dbs. up due to the increased band width. Sections 4.9.2 and 4.9.3 show that circuit and valve noise is proportional to the square root of the band width, whereas the audio signal is directly proportional.

7.7.2. Asymmetrical Variation. Asymmetrical variation of selectivity is rarely used because of the need for skilled operation if satisfactory results are to be realized. It may be achieved by

detuning the oscillator so as to locate the carrier on the side of the selectivity curve, or by detuning one side of the I.F. coupled circuits. The second method varies the total pass-band, whereas the first varies only the relative widths of the pass region of each side-band.

7.7.3. Symmetrical Variation. Symmetrical variation of selectivity may be accomplished by varying the mutual inductance coupling between the I.F. circuits, or the damping of the circuits. The latter is not very satisfactory as it appreciably reduces the discrimination against frequencies outside the pass-band.

7.7.4. Variable Selectivity by Mutual Inductance Variation.[3, 9, 10] We have already seen from Section 7.2 that a double-peaked response may be obtained from two tuned circuits coupled by mutual inductance. Increase of coupling causes these two peaks

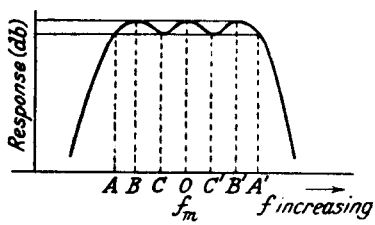


FIG. 7.9.—The Overall Selectivity Curve for an Overcoupled I.F. Transformer and a Single $\frac{1}{2}Q$ Circuit.

to travel by almost equal amounts from the mid-frequency f_m. It is important to note that the mutual inductance coupling between the circuits can only be satisfactorily varied by changing the relative positions of the two coils. The insertion of a shielding coil,[14] paralleled by a resistance between the two coils, is equivalent to a resistance shunting the mutual inductance arm. The effect is to reduce the lower frequency peak at $f_1 = \frac{1}{2\pi\sqrt{(L+M)C}}$, and if the resistance is made small enough this peak is completely suppressed. One disadvantage of mutual inductance coupling variation is that a trough is produced between the peaks and its depth increased as the peaks separate. This results in variable frequency response over the pass region between the peaks. It is, however, possible to overcome this difficulty if the overcoupled circuits are used in conjunction with a single tuned circuit of half the Q of the coupled circuits taken separately. The overall frequency response of the three circuits has three maxima of equal amplitude (see Fig. 7.9), and the frequency separation between the two outside maxima (contributed by the overcoupled circuits) is increased as the coupling

between the two circuits is increased. At the same time the two minima are reduced in amplitude so that an absolutely flat response over the pass-band is not obtained by this arrangement. Nevertheless, for normal pass-bands of ± 10 kc/s the variation can be reduced to very small proportions, less than 2 dbs. The proof that the maxima have equal amplitudes is as follows : the frequency response for the two coupled circuits is

$$20 \log_{10} \frac{R_D Q k}{\sqrt{[1+Q^2(k^2-F^2)]^2+4Q^2F^2}}$$

and that for the single tuned circuit

$$20 \log_{10} \frac{R_{D1}}{\sqrt{1+Q_1^2F^2}}.$$

The overall frequency response of the two circuits coupled by a valve is

$$20 \log_{10} \frac{R_D Q k R_{D1}}{\sqrt{([1+Q^2(k^2-F^2)]^2+4Q^2F^2)(1+Q_1^2F^2)}} \quad . \quad 7.9.$$

The variable factor is F, and since it occurs only in the denominator we can differentiate the latter and equate to zero in order to find the maxima and minima of the overall curve. We will replace Q_1 by $\frac{Q}{2}$ and check the final result by noting if three equal maxima are obtained.

$$(\text{Denominator})^2 = ([1+Q^2(k^2-F^2)]^2+4Q^2F^2)\left(1+\frac{Q^2F^2}{4}\right). \quad 7.10$$

$$\frac{d(D)^2}{dF} = \left(1+\frac{Q^2F^2}{4}\right)(-4Q^2F(1+Q^2(k^2-F^2))+8Q^2F) + ([1+Q^2(k^2-F^2)]^2+4Q^2F^2)\left(\frac{Q^2F}{2}\right)$$

$$= \frac{Q^2F}{2}[(9-6Q^2k^2+Q^4k^4)+Q^2F^2(12-4Q^2k^2)+3Q^4F^4] \quad . \quad 7.11$$

$$= 0$$

when $F = 0, \pm\sqrt{k^2-\frac{3}{Q^2}}$ and $\pm\sqrt{\frac{k^2}{3}-\frac{1}{Q^2}}$.

The three maxima in the overall frequency response curve, B, O and B' in Fig. 7.9, are given by the first three values of F $\left(0 \text{ and } \pm\sqrt{k^2-\frac{3}{Q^2}}\right)$ and the two minima, C and C', by the last

two values $\left(\pm\sqrt{\frac{k^2}{3}-\frac{1}{Q^2}}\right)$. Replacing F by 0 and $\pm\sqrt{k^2-\frac{3}{Q^2}}$ in expression 7.9 gives three equal maxima of

$$20\log_{10}\frac{R_DQkR_{D1}}{1+Q^2k^2} \quad . \quad . \quad . \quad . \quad 7.12$$

thus proving the statement that one circuit of $\frac{Q}{2}$ in conjunction with two coupled circuits, each of magnification Q, gives three equal maxima in the response curve. The value at the minima, obtained by replacing F by $\pm\sqrt{\frac{k^2}{3}-\frac{1}{Q^2}}$ in expression 7.9, is

$$20\log_{10}\frac{3\sqrt{3}R_D.R_{D1}}{9+Q^2k^2} \quad . \quad . \quad . \quad 7.13.$$

Thus the ratio maximum-to-minimum response is

$$20\log_{10}\frac{(9+Q^2k^2)Qk}{3\sqrt{3}(1+Q^2k^2)} \quad . \quad . \quad . \quad 7.14$$

and from this expression the decibel ratio variation of frequency response over the pass range can be determined for any value of Qk.

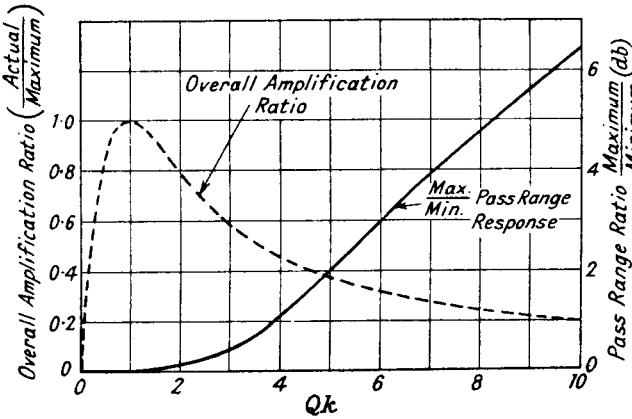


FIG. 7.10.—The Overall Amplification and Pass-band Ratios for an Overcoupled I.F. Transformer and a Single $\frac{1}{2}Q$ Circuit.

It is plotted in Fig. 7.10 against Qk. Expression 7.12 clearly shows that overall amplification is a function of Qk, and the condition for maximum is found by differentiating $\frac{R_DQkR_{D1}}{1+Q^2k^2}$ with respect to Qk and equating to 0. This gives $Qk = 1$ and maximum amplification proportional to $\frac{R_DR_{D1}}{2}$. The ratio loss of amplification from the

maximum value is shown against Qk as the dotted curve in Fig. 7.10.

The pass-band is not strictly the frequency separation between the two maxima, B and B', but is more correctly that between the two points A and A' in Fig. 7.9, i.e., the extreme frequencies at which the response is equal to that at the minima, C and C'. The values of F corresponding to A and A' are found by equating expressions 7.9 and 7.13, which means that

$$[[1+Q^2(k^2-F^2)]^2+4Q^2F^2]\left(1+\frac{Q^2F^2}{4}\right)=\frac{Q^2k^2(+9Q^2k^2)^2}{27}$$

or

$$\left[Q^2F^2+1-\frac{Q^2k^2}{3}\right]^2\left(Q^2F^2+4-\frac{4Q^2k^2}{3}\right)=0$$

$$F=\pm 2\sqrt{\frac{k^2}{3}-\frac{1}{Q^2}} \quad . \quad . \quad . \quad 7.15a.$$

Since $F=\dfrac{2\Delta f}{f_m}$, we may write for the maximum off-tune frequency

$$\Delta f_{(max.)}=f_m\sqrt{\frac{k^2}{3}-\frac{1}{Q^2}}=\frac{R}{2\pi L}\sqrt{\frac{Q^2k^2}{3}-1} \quad . \quad 7.15b.$$

To illustrate the design of such a stage let us assume that the coupled circuits are preceded and succeeded by valves and that the $\dfrac{Q}{2}$ single circuit is in the anode of the second valve.

Let g_{m1} = 1 mA/volt = mutual conductance of the first valve.
g_{m2} = 3 mA/volt = ,, ,, ,, ,, second ,,
f_m = 465 kc/s
$\Delta f_{(max.)}$ = 10 kc/s.

The following requirements are to be met:

The variation of the pass-band is not to exceed 2 dbs., the amplification (A_1) of the stage supplying the coupled circuits is to be 50 at critical coupling and that (A_2) of the second stage 100.

Figure 7.10 gives $Qk = 5$ for a 2-db. difference between the maxima and minima of the response curve.

Hence from expression 7.15*b*.

$$\frac{L}{R}=\frac{1}{2\pi 10^4}\sqrt{\frac{25-3}{3}}=4{\cdot}31\times 10^{-5}$$

$$Q=\frac{\omega_m L}{R}=126.$$

The amplification of the coupled circuit stage at critical coupling is

$$A_1 = g_{m1}\frac{\omega_m LQ}{2}$$

or

$$L = \frac{2A_1}{g_{m1}\omega_m Q} = 272\ \mu\text{H}$$

$$R_D = \omega_m LQ = 100{,}000\Omega$$

$$C = 430\ \mu\mu\text{F}$$

$$R = \frac{\omega_m L}{Q} = 6{\cdot}28\Omega$$

$$Qk = 5 = \frac{\omega_m M}{R}$$

$$M = 10{\cdot}75\ \mu\text{H}.$$

The amplification of the single circuit stage is

$$A_2 = 100 = g_{m2}\omega_m L_1 \times \frac{Q}{2}\left[Q_1 = \frac{Q}{2}\right]$$

thus

$$L_1 = 181\ \mu\text{H}$$

$$R_{D1} = 33{,}000\Omega$$

$$C_1 = 645\ \mu\mu\text{F}$$

$$R_1 = \frac{2\omega_m L_1}{Q} = 8{\cdot}4\ \Omega$$

Referring to the curve for overall amplification given in Fig. 7.10, we note that the actual operating overall amplification at maximum desired coupling is only 0·38 of its maximum value ; i.e., the operating amplification of the first stage is $0{\cdot}38 \times 50 = 19$. If the pass-band is reduced by reducing Qk, overall amplification is increased as far as critical coupling, and the variation over the pass-band is reduced. When variable selectivity is obtained by this method there is often very little change in volume when Qk exceeds 1, because the increase in side-band width tends to offset the loss of amplification at the lower audio frequencies. Automatic gain control will also tend to compensate for the reduction in operating amplification.

In order to obtain sufficient maximum selectivity, more than two coupled and one single circuit are required in an I.F. amplifier, but to preserve the flat pass-band we must maintain the ratio of two coupled to one single circuit. Using two pairs of coupled circuits and two single circuits each separated by a valve is not very practical, and it is essential to consider a compromise. If we

couple two tuned circuits loosely together, we obtain a selectivity curve which approximates to that of two separate tuned circuits connected by a valve, having zero anode-grid capacitance. It is therefore possible to couple the tuned circuit of $\frac{1}{2}Q$ to the secondary of the overcoupled circuits, and to obtain an overall selectivity curve which very nearly approaches that of the circuits separated by a valve. The actual effect of coupling on the pass-band response is to depress the side maxima below the mid-frequency maximum. The chief problem in design is to obtain a sufficiently high transfer impedance with loose coupling.

By assuming that all three circuits have equal values of L and C but that the Q of the last loosely coupled circuit is only $\frac{1}{2}$ the Q of the overcoupled circuit, the current in the secondary of the overcoupled circuits is (from 7.1*a*)

$$I_2 = \frac{E_2}{Z_5} = \frac{I_a Z_1 Z_3}{(Z_1+Z_2+Z_3)(Z_3+Z_4+Z_5+Z') - Z_3^2} \quad . \quad 7.16a$$

$$= \frac{I_a \omega M_1}{\omega C[R(1+jQF)(R(1+jQF)+Z') + \omega^2 M_1^2)} \quad . \quad 7.16b$$

where M_1 = mutual inductance between the overcoupled circuits

,, R = effective series resistance of these two circuits

,, Z' = impedance reflected into the secondary of the two circuits from the loosely coupled $Q/2$ circuit

,, $= \frac{\omega^2 M_2^2}{Z''}$ (Section 7.6)

,, Z'' = series impedance of the $Q/2$ circuit

and M_2 = mutual inductance between the $Q/2$ circuit and the secondary of the overcoupled circuits

$$Z'' = 2R + j\omega_m L\left(\frac{\omega}{\omega_m} - \frac{\omega_m}{\omega}\right)$$

$$= R(2+jQF).$$

The current-voltage equation for the $Q/2$ circuit is

$$I_3 R(2+jQF) + I_2 j\omega M_2 = 0$$

where I_3 = the current in the $Q/2$ circuit.

The output voltage across the capacitance is

$$E_3 = \frac{I_3}{j\omega C} = \frac{-I_2 j\omega M_2}{j\omega C R(2+jQF)}$$

and the transfer impedance is

$$Z_T = \frac{E_3}{I_a}$$

$$= \frac{-\omega^2 M_1 M_2}{\omega^2 C^2 R(2+jQF)\left[R(1+jQF)\left[R\left(1+jQF+\frac{\omega^2 M_2{}^2}{R^2(2+jQF)}\right)\right]+\omega^2 M_1{}^2\right]}$$

$$= \frac{-\omega^2 M_1 M_2}{\omega^2 C^2 R^3\left[(2+jQF)(1+jQF)^2+\frac{\omega^2 M_2{}^2}{R^2}(1+jQF)+\frac{\omega^2 M_1{}^2}{R^2}(2+jQF)\right]}$$

But $$\frac{\omega^2 M_1 M_2}{\omega^2 C^2 R^3} = \frac{\omega_m{}^2 L^2}{R^2}\cdot\frac{1}{\omega_m{}^2 C^2 R}\cdot\frac{M_1 M_2}{L^2} = Q^2 R_D k_1 k_2$$

where $$k_1 = \frac{M_1}{L}, \text{ and } k_2 = \frac{M_2}{L}$$

$$\frac{\omega^2 M_2{}^2}{R^2} = \left(\frac{\omega}{\omega_m}\right)^2\cdot\frac{\omega_m{}^2 L^2}{R^2}\cdot\frac{M_2{}^2}{L^2} = \left(\frac{\omega}{\omega_m}\right)^2 Q^2 k_2{}^2$$

$$\simeq Q^2 k_2{}^2$$

if $$\frac{\omega}{\omega_m} \simeq 1.$$

Similarly $$\frac{\omega^2 M_1{}^2}{R^2} \simeq Q^2 k_1{}^2.$$

Thus

$$Z_T = \frac{-Q^2 k_1 k_2 R_D}{2\left[1+Q^2\left(k_1{}^2+\frac{k_2{}^2}{2}-2F^2\right)\right]+jQF[5+Q^2(k_1{}^2+k_2{}^2-F^2)]} \qquad 7.17a$$

or $$|Z_T| = \frac{Q^2 k_1 k_2 R_D}{\sqrt{4\left[1+Q^2\left(k_1{}^2+\frac{k_2{}^2}{2}-2F^2\right)\right]^2+Q^2F^2[5+Q^2(k_1{}^2+k_2{}^2-F^2)]^2}} \qquad 7.17b$$

when Qk_2 is small

$$|Z_T| = \frac{Q^2 k_1 k_2 R_D}{\sqrt{4[1+Q^2(k_1{}^2-2F^2)]^2+Q^2F^2[5+Q^2(k_1{}^2-F^2)]^2}} \qquad 7.17c.$$

The square of the denominator of expression 7.17c can be expanded as follows

$$4\left[(1+Q^2(k_1{}^2-F^2)-Q^2F^2)^2+\frac{Q^2F^2}{4}(4+1+Q^2(k_1{}^2-F^2))^2\right]$$

$$= 4\left[(1+Q^2(k_1{}^2-F^2))^2-2Q^2F^2(1+Q^2(k_1{}^2-F^2))+Q^4F^4\right.$$

$$\left.+\frac{Q^2F^2}{4}(1+Q^2(k_1{}^2-F^2))^2+\frac{Q^2F^2}{4}[16+8(1+Q^2(k_1{}^2-F^2))]\right]$$

$$= 4\left[(1+Q^2(k_1^2-F^2))^2\left(1+\frac{Q^2F^2}{4}\right)+Q^2F^2(4+Q^2F^2)\right]$$

$$= 4[(1+Q^2(k_1^2-F^2))^2+4Q^2F^2]\left[1+\frac{Q^2F^2}{4}\right]$$

and this is 4 × expression 7.10 ; $|Z_T|$ therefore reduces to

$$|Z_T| = \frac{Q^2k_1k_2R_D}{2\sqrt{[(1+Q^2(k_1^2-F^2))^2+4Q^2F^2]\left[1+\frac{Q^2F^2}{4}\right]}} \quad . \quad 7.17d.$$

Hence the frequency response characteristic is the same as that when the overcoupled and single circuits are separated by a valve, and, provided Qk_2 is small, the curves for the loss over the pass range, and overall amplification ratio in Fig. 7.10 are applicable. The value of Qk_1 at which maximum amplification is realized, is moved to a slightly higher value by the coupling to the single tuned circuit and it can be found by differentiating the value of $|Z_T|$, for $F = 0$ ($f = f_m$), with respect to k_1. Thus from expression 7.17*b*

$$|Z_T|_{f_m} = \frac{Q^2k_1k_2R_D}{2\left[1+Q^2\left(k_1^2+\frac{k_2^2}{2}\right)\right]} \quad . \quad . \quad 7.18a$$

and $$\frac{d|Z_T|_{f_m}}{dk_1} = 0 \text{ when } Qk_1 = \sqrt{1+\frac{Q^2k_2^2}{2}}$$

so that if $Qk_2 = 0{\cdot}2$, $Qk_1 = 1{\cdot}01$, which is not very different from 1.

Replacing Qk_1 in expression 7.18*a* gives

$$|Z_T|_{f_m} = \frac{Qk_2R_D}{4\sqrt{1+\frac{Q^2k_2^2}{2}}} \quad . \quad . \quad . \quad 7.18b.$$

If the required maximum overall amplification, A, of the stage is known, the value of Qk_2 may be calculated from 7.18*b*, for

$$A = g_m|Z_T|_{f_m} = \frac{g_mQk_2R_D}{4\sqrt{1+\frac{Q^2k_2^2}{2}}}$$

from which

$$Qk_2 = \frac{4A}{\sqrt{g_m^2R_D^2-8A^2}} \quad . \quad . \quad . \quad 7.19.$$

To illustrate the design of the circuits, let us consider the problem given in the previous example, viz., the overall response in the pass range (±10 kc/s) is not to exceed 2 dbs., the amplification required at critical coupling is 50, $f_m = 465$ kc/s, and $g_m = 1$ mA/volt. If

it is assumed that the coupling k_2 is sufficiently loose to allow the curves in Fig. 7.10 to be used, the 2 dbs. variation gives

$$Qk_{1(max.)} = 5, \frac{L}{R} = 4{\cdot}31 \times 10^{-5} \text{ and } Q = 126.$$

The value of L must be as high as possible to obtain the highest transfer impedance. Let us assume that the maximum permissible value is 1,000 μH, which gives a tuning capacitance of 117 $\mu\mu$F.

Thus $L = 1{,}000\ \mu\text{H}. \quad C = 117\ \mu\mu\text{F}, \quad R = \frac{\omega L}{Q} = 23{\cdot}2\Omega.$

$$R_D = \omega L.Q = 368{,}000\Omega$$
$$g_m.R_D = 368.$$

From 7.19

$$Qk_2 = \frac{200}{\sqrt{368^2 - 8(50)^2}} = 0{\cdot}588.$$

This value of Qk_2 is too large, and the single circuit will affect the response of the overcoupled circuits to too great an extent. A more suitable value of Qk_2 is about 0·2. This entails increasing either R_D or g_m, or decreasing A. Increase of R_D involves an increase of L and decrease of C, and it is not very practicable since stray capacitance becomes an appreciable proportion of the total capacitance. A large ratio of stray-to-total capacitance increases the probability of mistuning during operation (due to valve or wiring lead changes, and to temperature effects). Furthermore, an increase in R_D may not result in a proportionate increase in amplification, due to R_D becoming comparable with the valve slope resistance R_a; when R_D approaches R_a the frequency response of the overcoupled circuits is affected and the variation over the pass-band increased. The only course is therefore to reduce A, if g_m cannot be increased; let us therefore take A as 17 as this gives $Qk_2 = 0{\cdot}2$. The critical value of Qk_1 corresponding to maximum amplification is

$$Qk_{1(crit.)} = \sqrt{1 + \frac{Q^2k_2^2}{2}} = 1{\cdot}01$$

$$M_{1(crit.)} = \frac{Qk_1R}{\omega_m} = \frac{1{\cdot}01 \times 23{\cdot}2 \times 10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5}$$
$$= 8{\cdot}02\ \mu\text{H}$$

$$M_{1(max.)} = \frac{5}{1{\cdot}01} M_{1(crit.)} = 39{\cdot}7\ \mu\text{H}$$

$$M_2 = \frac{Qk_2R}{\omega_m} = \frac{0{\cdot}2 \times 23{\cdot}2 \times 10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5}$$
$$= 1{\cdot}59\ \mu\text{H}.$$

The following component values satisfy all the requirements except that of amplification, the value of which cannot exceed 17 at critical coupling :

for the overcoupled circuits

$L = 1{,}000\ \mu H, \quad C = 117\ \mu\mu F, \quad R = 23{\cdot}2\Omega, \quad R_D = 368{,}000\Omega$
$Q = 126$
$M_1 = 8{\cdot}02\ \mu H$ (critical coupling)
$= 39{\cdot}7\ \mu H$ (for the required pass-band of $\pm$ 10 kc/s)

and for the single circuit

$L = 1{,}000\ \mu H, \quad C = 117\ \mu\mu F, \quad R = 46{\cdot}4\Omega, \quad R_D = 184{,}000\Omega$
$Q = 63$
$M_2 = 1{\cdot}59\ \mu H.$

Another possible arrangement for reducing the number of stages is to follow two pairs of overcoupled circuits by a pair of loosely coupled circuits of magnification $Q/2$, each pair being separated from the next by a valve. Since the $Q/2$ circuits form the last stage of the amplifier, the detector is connected across the secondary, and an A.G.C. detector will probably be connected across the primary. As a basis of design let us assume that each circuit has initially the same Q as the overcoupled circuits, and that the detector damping, and A.G.C. detector and amplifier anode damping reduce the magnification of secondary and primary to $Q/2$. We must now decide on the minimum coupling. If it is made too small in order that the selectivity curve of the coupled circuits may approach that of two valve-separated circuits, there is considerable loss of amplification to the detector. As a compromise let us take the coupling coefficient as one half critical. From section 7.3 critical coupling is realized when the coupling coefficient is equal to the reciprocal of the magnification of either circuit, therefore the required coupling coefficient is

$$k = \frac{0{\cdot}5}{Q/2} = \frac{1}{Q}$$

A lower value of coupling, for example $\frac{0{\cdot}2}{Q/2} = 0{\cdot}4Q$ is preferable for approaching the selectivity characteristics of two separated circuits, but this reduces appreciably the voltage to the detector.

The values of the circuit components are estimated as follows :

Let R_{DET} = parallel damping resistance due to the detector
R_p = parallel damping resistance due to the A.G.C. detector and the amplifier valve resistance R_a

and R_1 and R_2 = primary and secondary coil series resistances.

If the magnification of each circuit is reduced to one-half by damping, it follows that the initial resonant impedance of each circuit equals the parallel damping resistance, thus the circuit constants are (Section 4.2.2)

$$\omega L_1 Q = R_p\,; \quad \omega L_2 Q = R_{DET}$$

where Q = initial undamped magnification of either circuit.

$$L_1 = \frac{R_p}{\omega Q} \qquad L_2 = \frac{R_{DET}}{\omega Q}$$

$$C_1 = \frac{1}{\omega^2 L_1} \qquad C_2 = \frac{1}{\omega^2 L_2}$$

$$R_1 = \frac{\omega L_1}{Q} \qquad R_2 = \frac{\omega L_2}{Q}$$

$$Qk = \frac{\omega M}{\sqrt{R_1 R_2}} = 1 \text{ or } M = \frac{\sqrt{R_1 R_2}}{\omega}.$$

The damping resistance due to the detector (R_{DET}) and that in R_p due to the A.G.C. detector may be taken as one-half of the A.C. load resistance (see Section 8.2.10). If the A.G.C. detector is delayed, its damping resistance only becomes fully effective when the R.F. voltage is much greater than the delay voltage. However, the maximum band width is not likely to be used unless strong signals are being received. Taking the value of Q obtained in the previous example, i.e., 126, and assuming that the A.C. load resistance of A.G.C. and detector diodes is 0·5 MΩ and that the valve slope resistance R_a is 1 MΩ, we have

$$R_p = \frac{1 \times 0{\cdot}25}{1{\cdot}25} = 0{\cdot}2 \text{ M}\Omega. \qquad R_{DET} = 0{\cdot}25 \text{ M}\Omega$$

$$L_1 = \frac{0{\cdot}2 \times 10^{12}}{6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 126} \qquad L_2 = \frac{0{\cdot}25 \times 10^{12}}{6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 126}$$

$$= 544\ \mu\text{H} \qquad = 680\ \mu\text{H}$$

$$R_{D1} = \tfrac{1}{2}\omega L_1 Q = 100{,}000\Omega$$

$$C_1 = 215\ \mu\mu\text{F} \qquad C_2 = 172\ \mu\mu\text{F}$$

$$R_1 = 12{\cdot}6\Omega \qquad R_2 = 15{\cdot}75\Omega$$

$$M = \frac{\sqrt{R_1 R_2}}{\omega Q k} = 4{\cdot}82\ \mu\text{H}.$$

Owing to coupling between the two circuits there is a greater variation of overall frequency response over the pass-range, and the

maximum at f_m is less than the maxima at the side frequencies. The effect is not very serious if a coupling of one-half critical is not exceeded.

The number of stages may be still further reduced by coupling the two pairs of overcoupled circuits loosely together as shown in Fig. 7.11*a*. The equivalent circuit is shown in Fig. 7.11*b* and the

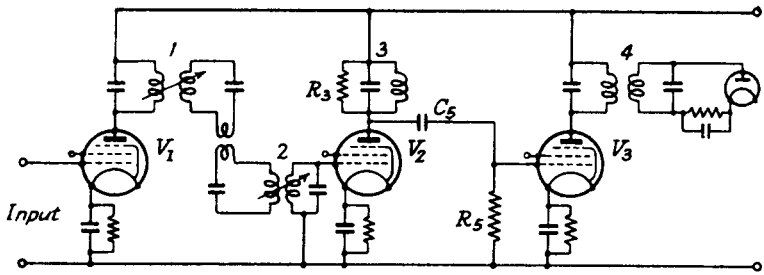


FIG. 7.11*a*.—A Variable Selectivity I.F. Amplifier with almost Constant Pass-band Response.

transfer impedance can be calculated by the method used in Section 7.3. By replacing the impedances by Z_1, Z_2, etc., the general expression for Z_T is

$$Z_T = \frac{E_2}{I_a}$$

$$= \frac{Z_1Z_3Z_5Z_7Z_9}{(Z_1+Z_2+Z_3)(Z_3+Z_4+Z_5)(Z_5+Z_6+Z_7)(Z_7+Z_8+Z_9) - Z_3^2(Z_5+Z_6+Z_7)(Z_7+Z_8+Z_9) - Z_5^2(Z_1+Z_2+Z_3)(Z_7+Z_8+Z_9) - Z_7^2(Z_1+Z_2+Z_3)(Z+Z_4+Z_5) + Z_3^2Z_7^2.}$$

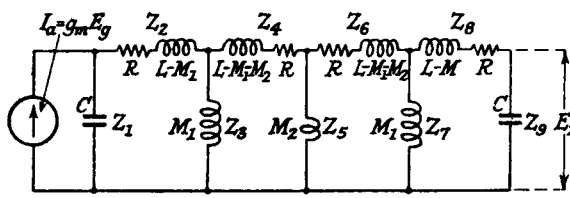


FIG. 7.11*b*.—The Equivalent Circuit for Two Loosely Coupled I.F. Transformers. [For Z_s read $L - M_1$ and R.]

If all the circuits have identical component values, as indicated in Fig. 7.11*b*

$$Z_1 = Z_9 = \frac{1}{j\omega C}; \quad Z_2 = Z_8 = R + j\omega(L - M_1)$$

$$Z_3 = Z_7 = j\omega M_1; \quad Z_4 = Z_6 = R + j\omega(L - M_1 - M_2), \quad Z_5 = j\omega M_2,$$

$$Z_1+Z_2+Z_3 = R+j\left(\omega l - \frac{1}{\omega C}\right) = Z_3+Z_4+Z_5 = Z_5+Z_6+Z_7$$
$$= Z_7+Z_8+Z_9 = R(1+jQF)$$

$$\therefore\ Z_T = \frac{j\dfrac{\omega^3 M_1^2 M_2}{\omega^2 C^2}}{R^4(1+jQF)^4+R^2\omega^2M_1^2(1+jQF)^2+R^2\omega^2M_2^2(1+jQF)^2 \\ +R^2\omega^2M_1^2(1+jQF)^2+\omega^4M_1^4.}$$

$$\text{Numerator of } Z_T = \frac{\omega M_1^2 M_2}{C^2} = \frac{\omega}{\omega_m}R^4\left[\frac{\omega_m^3L^3}{R^3}\cdot\frac{1}{\omega_m^2C^2R}\cdot\frac{M_1^2}{L^2}\cdot\frac{M_2}{L}\right]$$
$$= \frac{\omega}{\omega_m}R^4Q^3R_Dk_1^2k_2$$
$$\simeq R^4Q^3k_1^2k_2R_D.$$

Denominator of Z_T

$$= R^4\left[(1+jQF)^4+\frac{2\omega^2M_1^2}{R^2}(1+jQF)^2+\frac{\omega^2M_2^2}{R^2}(1+jQF)^2+\frac{\omega^4M_1^4}{R^4}\right]$$
$$\simeq R^4[(1+jQF)^4+2Q^2k_1^2(1+jQF)^2+Q^2k_2^2(1+jQF)^2+Q^4k_1^4]$$
$$\simeq R^4[(1+Q^2(k_1^2-F^2))^2-4Q^2F^2+Q^2k_2^2(1-Q^2F^2) \\ +j2QF(2+Q^2k_2^2+2Q^2(k_1^2-F^2))].$$

Hence

$$|Z_T| = \frac{Q^3k_1^2k_2R_D}{\sqrt{[(1+Q^2(k_1^2-F^2))^2-4Q^2F^2+Q^2k_2^2(1-Q^2F^2)]^2 \\ +4Q^2F^2[2+Q^2k_2^2+2Q^2(k_1^2-F^2)]^2}} \quad . \quad 7.20a$$

when k_2 is small this approaches

$$|Z_T| = \frac{Q^3k_1^2k_2R_D}{[1+Q^2(k_1^2-F^2)]^2+4Q^2F^2}$$

and the selectivity curve, which is determined by the denominator, is the same as that for two separated pairs.

At the mid-frequency f_m, $F = 0$ and

$$|Z_T| = \frac{Q^3k_1^2k_2R_D}{(1+Q^2k_1^2)^2+Q^2k_2^2} \quad . \quad . \quad . \quad 7.20b$$

differentiating this with respect to k_1 and equating to zero gives for a maximum

$$Qk_{1(crit.)} = \sqrt[4]{1+Q^2k_2^2} \quad . \quad . \quad . \quad 7.21.$$

Replacing Qk_1 in 7.20*b* by this value

$$|Z_T| = \frac{Qk_2R_D\sqrt{1+Q^2k_2^2}}{(1+\sqrt{1+Q^2k_2^2})^2+Q^2k_2^2} = \frac{Qk_2R_D}{2(1+\sqrt{1+Q^2k_2^2})}. \quad 7.20c.$$

The overall amplification of the stage is

$$A = g_m \mid Z_T \mid = \frac{g_m Q k_2 R_D}{2(1+\sqrt{1+Q^2k_2^2})}$$

or $4\ A^2(1+Q^2k_2^2) = (g_m R_D Q k_2 - 2A)^2$

whence

$$Qk_2 = \frac{4Ag_mR_D}{g_m^2R_D^2 - 4A^2} \quad . \quad . \quad . \quad 7.22.$$

Taking the component values already used for the overcoupled circuits and loosely coupled single circuit, viz.,

$L = 1{,}000\ \mu\text{H}$, $C = 117\ \mu\mu\text{F}$, $Q = 126$, $Qk_2 = 0{\cdot}2$, $g_m = 1$ mA/volt
$R = 23{\cdot}2\Omega$, $R_D = 368{,}000\Omega$

$$A = g_m \mid Z_T \mid = \frac{1 \times 10^{-3} \times 368{,}000 \times 0{\cdot}2}{2(1+\sqrt{1{\cdot}04})}$$

$$= 18{\cdot}2.$$

This represents the maximum gain which can be obtained from the stage with loose enough coupling between the pairs of overcoupled circuits.

$$Qk_{1(crit.)} = \sqrt[4]{1+Q^2k_2^2} = \sqrt[4]{1{\cdot}04}$$

$$= 1{\cdot}01$$

$$M_{1(crit.)} = \frac{Qk_1R}{\omega_m} = \frac{1{\cdot}01 \times 23{\cdot}2 \times 10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5}$$

$$= 8{\cdot}02\mu\text{H}.$$

For 2db. loss at ± 10 kc/s, Fig. 7.10 gives $Qk_{1(max.)} = 5$

$$\therefore\ M_{1(max.)} = 39{\cdot}7\ \mu\text{H}$$

$$M_2 = \frac{0{\cdot}2 \times 23{\cdot}2 \times 10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5} = 1{\cdot}59\ \mu\text{H}.$$

The coupling reactance between the two circuits may, if desired, be a capacitance C_1, its value being determined from $Qk_2 = \dfrac{1}{\omega C_1 R}$,

thus

$$C_1 = \frac{10^6}{\omega R Q k_2} = \frac{10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 23{\cdot}2 \times 0{\cdot}2}$$

$$= 0{\cdot}0738\ \mu\text{F}.$$

The three compromise circuits give a frequency response over the pass-band which is almost equal to that of the separated circuits if the loose coupling coefficient does not exceed appreciably the value given in the calculations. Stray capacitance coupling, in addition to the mutual inductance coupling, may occur and produce an asymmetrical overall frequency response. The effect can often

be reduced by reversing the sign of each mutual inductance element with respect to the preceding one.

7.8. Valve Input Admittance and Frequency Response. In the preceding theory we have assumed that the input admittance of each valve amplifier is zero over the pass-band. In practice the input admittance due to the valve anode-grid capacitance may be comparable with that of the circuits across which it is connected. This input admittance may be separated into two parallel components, resistive and reactive (see Section 2.8.2). The resistive component is from expression 2.9*a* (neglecting B_{g_a} in comparison with B_0)

$$R_g = \frac{(G_a+G_0)^2+B_0^2}{g_m B_{g_a} B_0} \quad . \quad . \quad . \quad 7.23a$$

where $G_a = \dfrac{1}{R_a}$ = anode slope conductance

G_0 = conductance of the external anode load admittance

B_0 = susceptance ,, ,, ,, ,, ,, ,,

g_m = mutual conductance of the valve

B_{g_a} = susceptance of the anode-grid capacitance.

The sign of R_g depends on B_0, being negative when B_0 is inductive and positive when B_0 is capacitive. With tuned anode circuits B_0 is inductive at frequencies lower than resonance, and capacitive for higher frequencies. The value of R_g is minimum when $B_0 = G_a+G_0$ and, for single or coupled circuits of normal Q, the minimum occurs in the pass-band of the amplifier. This tends to produce an overall lop-sided frequency response curve higher on the low-frequency side of resonance where R_g is negative. The grid parallel reactive component is always capacitive with a value (see expression 2.9*b*, again neglecting B_{g_a} in comparison with B_0).

$$C_g = C_{g_a}\left[1+\frac{g_m(G_a+G_0)}{(G_a+G_0)^2+B_0^2}\right] \quad . \quad . \quad 7.23b.$$

The variation of this component over the pass-band is generally negligible and compensation for its mistuning effect can be made on the grid-coil tuning capacitance. The resistive component cannot be ignored, and the only satisfactory solution is to isolate each stage from the next by a semi-aperiodic circuit. The dynamic resistance of this circuit must be much less than the minimum input resistance component of the succeeding valve, and it must also have a sufficiently low value to make the input resistance component of the valve, to which it is connected, very much greater than the dynamic resistance of the preceding circuit.. The ratio of minimum

grid input resistance to tuned circuit dynamic resistance R_D should not be less than 10 : 1. In order, therefore, to obtain satisfactory results it is necessary to insert an isolator semi-aperiodic stage between the variable coupling and the fixed coupling stages. The modified circuit appears as in Fig. 7.11*a*.

In order to determine the constants of the aperiodic circuit, we must calculate the input resistance component of valve V_3, connected to the two $Q/2$ loosely coupled circuits. Using the equivalent primary circuit developed in Section 7.6, the series resistive and reactive components reflected from the secondary into the primary inductance branch are

$$R_S = \frac{2\omega^2M^2}{R_2(4+Q^2F^2)} \text{ and } X_S = \frac{-\omega^2M^2QF}{R_2(4+Q^2F^2)}.$$

(Note that the equivalent secondary series resistance is $2R_2$ because the circuit is damped to $Q/2$, and that $Q = \frac{\omega L_1}{R_1} = \frac{\omega L_2}{R_2}$.)

The equivalent primary series resistance is $2R_1 + \frac{2\omega^2M^2}{R_2(4+Q^2F^2)}$ and the reactance of the inductive arm is

$$j\omega L_1 - \frac{j\omega^2M^2QF}{R_2(4+Q^2F^2)} \simeq j\omega L_1$$

when F is small, i.e., in the pass-band range.

Hence the total equivalent parallel resistance of the primary is very nearly (see Section 4.2.2.)

$$\frac{\omega^2L_1^2}{2R_1 + \frac{2\omega^2M^2}{R_2(4+Q^2F^2)}}$$

and, as this includes the amplifier valve anode resistance and A.G.C. damping because the primary series resistance is taken as $2R_1$, it is equal to $\frac{1}{G_a+G_0}$.

or

$$(G_a+G_0) = \frac{2R_1\left[1+\frac{\omega^2M^2}{R_1R_2(4+Q^2F^2)}\right]}{\omega^2L_1^2} = \frac{2\left[1+\frac{\omega^2M^2}{R_1R_2(4+Q^2F^2)}\right]}{\frac{\omega^2L_1^2}{R_1}} = \frac{2\left[1+\frac{Q^2k^2}{4+Q^2F^2}\right]}{R_{D1}} \quad . \quad . \quad . \quad 7.24a.$$

$$B_0 = \omega C_1 - \frac{1}{\omega L_1 - \dfrac{\omega^2 M^2 QF}{R_2(4+Q^2F^2)}}.$$

The reflected reactance term cannot now be neglected since ωC_1 is comparable with $\dfrac{1}{\omega L_1}$

$$B_0 = \frac{\omega^2 L_1 C_1 - 1 - \dfrac{\omega^2 M^2 QF \omega C_1}{R_2(4+Q^2F^2)}}{\omega L_1 - \dfrac{\omega^2 M^2 QF}{R_2(4+Q^2F^2)}}.$$

But $$\omega^2 L_1 C_1 - 1 = \left[\frac{\omega^2}{\omega_m{}^2} - 1\right] = F$$

and $$Q\omega C_1 = \frac{\omega^2 L_1 C_1}{R_1} = \frac{\dfrac{\omega^2}{\omega_m{}^2}}{R_1} \simeq \frac{1}{R_1}$$

$$\therefore\ B_0 \simeq \frac{F - \dfrac{\omega^2 M^2 F}{R_1 R_2(4+Q^2F^2)}}{\omega L_1}$$

$$= \frac{F}{\omega L_1}\left[1 - \frac{Q^2k^2}{4+Q^2F^2}\right]$$

$$= \frac{QF}{R_{D1}}\left[1 - \frac{Q^2k^2}{4+Q^2F^2}\right] \quad . \quad . \quad . \quad 7.24b.$$

Note that F is negative at frequencies below resonance, and hence B_0 is negative (inductive).

If we consider the case of the two loosely coupled $Q/2$ circuits, as shown in the anode circuit of V_3 in Fig. 7.11*a*, $Qk = 1$ and expressions 7.24*a* and 7.24*b* become

$$G_a + G_0 = \frac{2}{R_{D1}}\left[1 + \frac{1}{4+Q^2F^2}\right] \quad . \quad . \quad . \quad 7.25a$$

and $$B_0 = \frac{QF}{R_{D1}}\left[1 - \frac{1}{4+Q^2F^2}\right] \quad . \quad . \quad 7.25b$$

and replacing (G_a+G_0) and B_0 in expression 7.23*a* we have

$$R_g = \frac{4(5+Q^2F^2)^2 + Q^2F^2(3+Q^2F^2)^2}{g_m B_{g_a} R_{D1} QF (3+Q^2F^2)(4+Q^2F^2)}$$

$$= \frac{25+6Q^2F^2+Q^4F^4}{g_m B_{g_a} R_{D1} QF (3+Q^2F^2)} \quad . \quad . \quad . \quad 7.23c.$$

To find the condition for a minimum value of R_g, it is necessary to

differentiate 7.23c with respect to F and equate to 0. A cubic equation in Q^2F^2 results:

$$Q^6F^6+3Q^4F^4 - 57Q^2F^2 - 75 = 0.$$

One solution of which is

$$Q^2F^2 = 6{\cdot}86 \text{ or } QF = \pm 2{\cdot}62.$$

Replacing QF in 7.23c by 2·62 gives

$$R_{g(min.)} = \frac{4{\cdot}38}{g_m B_{g_a} R_{D1}}.$$

If we assume that the anode-grid capacitance of V_3 (Fig. 7.11a) = 0·005 $\mu\mu$F and g_m = 3 mA/volt and that the dynamic resistance of one $Q/2$ circuit is R_{D1} = 100,000Ω (as calculated for the previous example), the minimum grid input resistance of V_3 is

$$R_{g(min.)} = \frac{4{\cdot}38 \times 10^{12}}{3 \times 10^{-3} \times 6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 0{\cdot}005 \times 100{,}000} = 1\text{M}\Omega$$

so that the dynamic resistance R_{DA} of the isolator or semi-aperiodic tuned circuit must not exceed 100,000Ω if the ratio of $\frac{R_{g(min.)}}{R_{DA}}$ is not to exceed 10.

We have yet to consider the minimum grid input resistance of the isolator valve V_2, as this must not be less than 10 times the dynamic resistance of the circuit between grid and earth. The conductance and susceptance of the isolator anode circuit are from expressions 4.7 in Section 4.2.2.

$$G_0 = \frac{R}{\omega^2 L^2} = \frac{1}{R_{DA}}$$

$$B_0 = \frac{\omega^2 LC - 1}{\omega L} = \frac{F}{\omega L} \simeq \frac{QF}{R_{DA}}.$$

Note that X_p in expression 4.7 is positive and that a positive X_p is equal to $-B_0$. Hence the change of sign in the numerator of B_0.

$$\therefore R_g = \frac{\left(\frac{1}{R_a}+\frac{1}{R_{DA}}\right)^2+\left(\frac{QF}{R_{DA}}\right)^2}{g_m B_{g_a} \frac{QF}{R_{DA}}}$$

$$= \frac{Q^2F^2+\left(1+\frac{R_{DA}}{R_a}\right)^2}{g_m B_{g_a} QF R_{DA}} \quad . \quad . \quad . \quad 7.26a$$

where $R_a = \frac{1}{G_a}$ = the slope resistance of V_2.

At frequencies below resonance F is negative and R_g is negative. Regeneration of the input circuit therefore occurs. For frequencies above resonance R_g is positive and degeneration occurs.

The condition for minimum R_g is found by differentiating 7.26*a* with respect to F and equating to 0. Thus

$$QF = 1 + \frac{R_{DA}}{R_a} \quad . \quad . \quad . \quad . \quad 7.27.$$

If $R_a \gg R_{DA}$ — in practice this is generally true — $QF = 1$ and

$$R_{g(min.)} = \frac{2}{g_m B_{g_a} R_{DA}} \quad . \quad . \quad . \quad 7.26b.$$

Taking the maximum grid circuit resistance as the dynamic resistance of one of the tuned circuits forming the two pairs of overcoupled circuits in the anode of V_1, i.e., as 368,000Ω, the minimum grid input resistance of the isolator valve V_2 must not be less than 3·68 MΩ. If the grid-anode capacitance of V_2 is 0·005 $\mu\mu$F and g_m is 3 mA/volt, the maximum permissible dynamic resistance of the isolator-tuned circuit is from 7.26*b*.

$$R_{DA} = \frac{2}{g_m B_{g_a} R_{g(min.)}}$$

$$= \frac{2 \times 10^{12}}{3 \times 10^{-3} \times 6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 0{\cdot}005 \times 3{,}680{,}000}$$

$$= 12{,}400\Omega.$$

The amplification of this stage is $g_m R_{DA} = 37{\cdot}2$. It is clear in this particular example that the factor limiting R_{DA} is the input grid resistance of the valve V_2, in the anode circuit of which it is connected. Having fixed the value of R_{DA} it is necessary now to decide on the Q, L and C of the circuit. From 7.27, the grid input resistance is a minimum at $QF = 1$, i.e., at an off tune frequency $\Delta f = \frac{f_m}{2Q}$ and the only condition governing Δf is that it should be as far outside the pass range as possible. Let us assume that $\Delta f = 30$ kc/s, then

$$Q = \frac{465}{60} = 7{\cdot}75$$

and

$$L = \frac{R_{DA}}{\omega Q} = 547 \ \mu\text{H}$$

$$C = 214 \ \mu\mu\text{F}$$

$$R = 206\Omega.$$

The isolator circuit does not give an absolutely flat response over the pass-band and the loss in decibels is $20 \log \sqrt{1 + Q^2F^2} = 0{\cdot}44$ db. for $\Delta f = 10$ kc/s. The effect of the circuit on the overall response

may therefore be neglected. The variation of R_g (calculated from 7.26*b*) over a range of modulation frequencies from 1 to 100 kc/s is shown in Fig. 7.12. The variation of C_g over the same range of frequencies is also shown; its value from 7.23*b* is

$$C_g = C_{g_a}\left[1+\frac{g_m(G_a+G_0)}{(G_a+G_0)^2+B_0^2}\right]$$

$$\simeq C_{g_a}\left[1+\frac{g_m R_{DA}}{1+Q^2F^2}\right] . \quad . \quad . \quad . \quad 7.28$$

when $$R_a \gg R_{DA}.$$

Its variation over the pass-range is seen to be negligible; since it is dependent on g_m, C_g is reduced if the g_m of the valve is reduced,

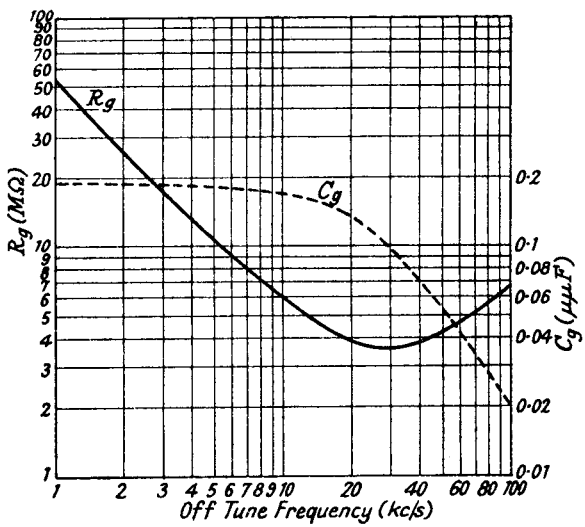


FIG. 7.12.—Curves showing the Variation of Grid Input Resistance and Capacitance for the Isolator Stage.

e.g., by A.G.C. action. A variation of g_m from 3 to 0·1 mA/volt causes a change of 0·1798 $\mu\mu$F in C_g, and this represents a mistuning of approximately 0·358 kc/s for the over-coupled circuit having a tuning capacitance of 117 $\mu\mu$F, a not very serious tuning error.

7.9. Cathode Feedback and Variable Selectivity.[16, 18] Cathode feedback may be used in an I.F. amplifier to give a level pass-band frequency response and to sharpen the attenuation outside this range, when variable selectivity is produced by changing mutual inductance between coupled tuned circuits. Two parallel-tuned circuits are included in series between the cathode and earth

as shown in Fig. 7.13*a*. One is tuned to (f_m+9) kc/s and the other to (f_m-9) kc/s. (Medium wave stations are normally separated by 9 kc/s.) These circuits provide negative feedback with considerable attenuation at the two resonant frequencies. The figure shows the tuned circuits as transformer coupled, and this is done

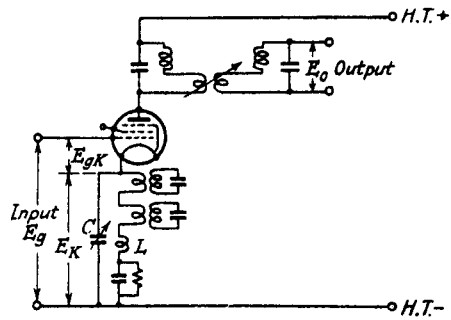


FIG. 7.13*a*.—An I.F. Amplifier with Cathode Feedback.

in order to reduce the impedance at f_m so that the loss over the pass-band range may be as small as possible. A fuller explanation is given at the end of this section. The coil L (of low inductance) and capacitance C (a small trimmer) are generally found necessary with the transformer coupling arrangement, in order to produce symmetry of the frequency response feedback characteristic. The frequency response due to feedback is shown in Fig. 7.13*b*, and by

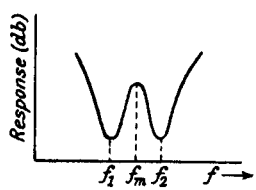


FIG. 7.13*b*.—The Frequency Response due to Cathode Feedback.

suitable design it is possible to neutralize the trough normally produced by overcoupled circuits in the anode and grid of the valve. The principle underlying the method has already been discussed in Section 2.7, where it is shown that the output voltage in the anode circuit is (expression 2.5*d*)

$$E_0 = \frac{g_m E_g Z_T}{1+g_m Z_k}$$

or overall amplification $= \dfrac{E_0}{E_g} = \dfrac{g_m Z_T}{1+g_m Z_k} = g_m' Z_T.$

By plotting $20 \log_{10} \frac{g_m}{g_{m'}}$ against frequency we obtain the frequency loss in decibels due to cathode feedback, and adding this curve to the Z_T curve gives the overall response of the stage. To do this we must obtain the generalized expression for the cathode impedance. The expression for a parallel resonant circuit impedance (obtained by rationalizing expression 4.8*b*) is

$$Z = \frac{R_D(1 - jQF)}{1+Q^2F^2}$$

where $R_D = \frac{\mathrm{L}}{CR}$; $F = \frac{2\Delta f'}{f_1}$ and $f_1 =$ the resonant frequency. However, we do not wish to refer to f_1 but to the intermediate frequency f_m. If $f_1 < f_m$

$$\frac{2\Delta f'}{f_1} = \frac{2(\Delta f_1+\Delta f)}{f_1} \quad . \quad . \quad . \quad 7.29$$

where $f_m - f_1 = \Delta f_1$ and Δf is now referred to f_m as zero, i.e., equals $-[f_m - (f_1+\Delta f')]$.

That this is correct is proved by noting that $\Delta f = -\Delta f_1$ referred to f_m gives $\Delta f'$ referred to f_1.

When Δf_1 is small, $f_1 \simeq f_m$, and 7.29 becomes

$$2\left[\frac{\Delta f_1}{f_m}+\frac{\Delta f}{f_m}\right]$$

and

$$Z_1 = \frac{R_D\left[1 - j2Q\left(\frac{\Delta f_1}{f_m}+\frac{\Delta f}{f_m}\right)\right]}{1+4Q^2\left(\frac{\Delta f_1}{f_m}+\frac{\Delta f}{f_m}\right)^2}$$

rewriting $\frac{2\Delta f_1}{f_m}$ as F_1 and $\frac{2\Delta f}{f_m}$ as F we have

$$Z_1 = \frac{R_D[1 - jQ(F_1+F)]}{1+Q^2(F_1+F)^2} \quad . \quad . \quad . \quad 7.30a.$$

Similarly for the other circuit ($f_2 > f_m$ and the j term is positive)

$$Z_2 = \frac{R_D[1+jQ(F_2 - F)]}{1+Q^2(F_2 - F)^2}$$

$$= \frac{R_D[1+jQ(F_1 - F)]}{1+Q^2(F_1 - F)^2} \quad . \quad . \quad 7.30b$$

for $f_2 - f_m = \Delta f_2 = f_m - f_1 = \Delta f_1$

$$\therefore \quad F_2 = \frac{\Delta f_2}{f_m} = F_1 = \frac{\Delta f_1}{f_m}$$

$$Z_k = Z_1 + Z_2 = \frac{2R_D[1+Q^2(F_1^2+F^2) - jQF(1 - Q^2(F_1^2 - F^2))]}{[1+Q^2(F_1+F)^2][1+Q^2(F_1 - F)^2]}$$

$$Z_k = \frac{2R_D\sqrt{[1+Q^2(F_1^2+F^2)]^2+Q^2F^2[1 - Q^2(F_1^2 - F^2)]^2}}{1+2Q^2(F_1^2+F^2)+Q^4(F_1^2 - F^2)^2} \quad . \quad 7.31a$$

$$= \frac{2R_D\sqrt{\begin{array}{r}([1+Q^2(F_1^2+F^2)]^2 - 4Q^4F_1^2F^2) \\ +Q^2F^2([1 - Q^2(F_1^2 - F^2)]^2+4Q^2F_1^2)\end{array}}}{[1+Q^2(F_1^2+F^2)]^2 - 4Q^4F_1^2F^2}$$

$$= \frac{2R_D\sqrt{\begin{array}{r}[(1+Q^2(F_1^2+F^2))^2 - 4Q^4F_1^2F^2] \\ +Q^2F^2[(1+Q^2(F_1^2+F^2))^2 - 4Q^4F_1^2F^2]\end{array}}}{[1+Q^2(F_1^2+F^2)]^2 - 4Q^4F_1^2F^2}$$

$$= \frac{2R_D\sqrt{1+Q^2F^2}}{\sqrt{[1+Q^2(F_1^2+F^2)]^2 - 4Q^4F_1^2F^2}} \quad . \quad . \quad . \quad . \quad 7.31b.$$

The above expression contains three possible variables R_D, QF_1 and QF, and it is desirable to reduce to two if generalized curves are to be obtained. It is possible to replace R_D, if the permissible loss due to feedback at f_m is fixed. Let us assume this to be 3 dbs. At $f = f_m$, $F = 0$, expression 7.31*b* becomes $\frac{2R_D}{1+Q^2F_1^2}$; for 3db. loss $\frac{g_m}{g_m'} = 1{\cdot}414$

or $$1+g_mZ_k = 1{\cdot}414$$

$$g_mZ_k = 0{\cdot}414 = \frac{g_m2R_D}{1+Q^2F_1^2}$$

or $$2R_D = \frac{0{\cdot}414(1+Q^2F_1^2)}{g_m} \quad . \quad . \quad . \quad 7.32.$$

Replacing this in 7.31*b*.

$$|Z_k| = \frac{0{\cdot}414(1+Q^2F_1^2)}{g_m}\sqrt{\frac{1+Q^2F^2}{[1+Q^2(F_1^2+F^2)]^2 - 4Q^4F_1^2F^2}}$$

and the frequency response loss

$$20\log_{10}(1+g_mZ_k)$$

$$= 20\log_{10}\left[1+\frac{0{\cdot}414(1+Q^2F_1^2)\sqrt{1+Q^2F^2}}{\sqrt{[1+Q^2(F_1^2+F^2)]^2-4Q^4F_1^2F^2}}\right] . \quad 7.33.$$

This is plotted in Fig. 7.14 against QF for different values of $Q^2F_1^2$. The QF scale can be converted to an off-tune frequency scale as shown for the generalized coupled circuit curves in Fig. 7.7. To illustrate the use of the curves let us take the following example: $f_m = 465$ kc/s, the cathode circuits are to give maximum attenuation

at ± 9 kc/s, and a flat pass-band is required over a maximum range of ± 6 kc/s when the cathode circuits are used in conjunction with a pair of coupled circuits. The design procedure is as follows : the maximum permissible Q of the cathode circuits, which is the determining factor, should be as high as possible for maximum sharpness of cut-off outside the pass-band. At the same time the maximum Q curve may not give the flattest pass-band in conjunction with

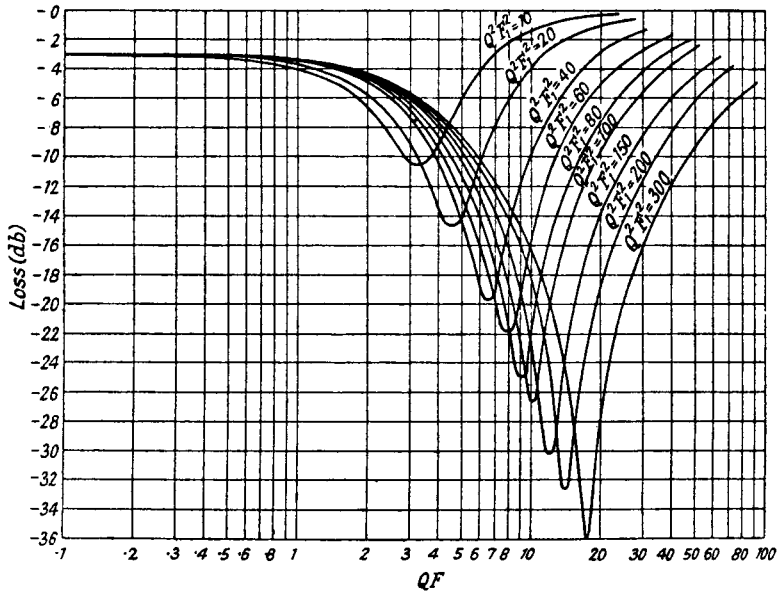


FIG. 7.14.—Generalized Selectivity Curves for Cathode Feedback.

the coupled circuits, and a compromise may be necessary. Let $Q = 225$; this automatically fixes $(QF_1)^2$ and the frequency scale

$$(QF_1)^2 = \left[\frac{225 \times 2 \times 9}{465}\right]^2 = 76 \simeq 80.$$

For a flat overall pass-band response the shape of the cathode feedback curve must be the reverse of the coupled circuit curve, and the former reversed should fit exactly over the latter in the pass-band range. The curve for $(QF_1)^2 = 80$ is therefore redrawn on tracing paper with the correct frequency scale but with the loss scale reversed. This curve is then placed on top of Fig. 7.7 and moved along until it fits as nearly as possible over a generalized coupled circuit curve up to 6 kc/s. The most suitable curve is that for $Qk = 6$ with $\Delta f = 1$ kc/s registering with $QF = 1$ on Fig. 7.7,

and the two curves are shown in Fig. 7.15. The overall response curve is obtained by plotting the difference between the two curves. The frequency scale for the coupled circuits is now fixed ($QF = 1$ when $\Delta f = 1$ kc/s), and Q and k are therefore 232·5 and 0·0258 respectively. The required value of Q for the coupled circuits is rather higher than would normally be obtainable, and a more practical value is found by assuming that two pairs of overcoupled

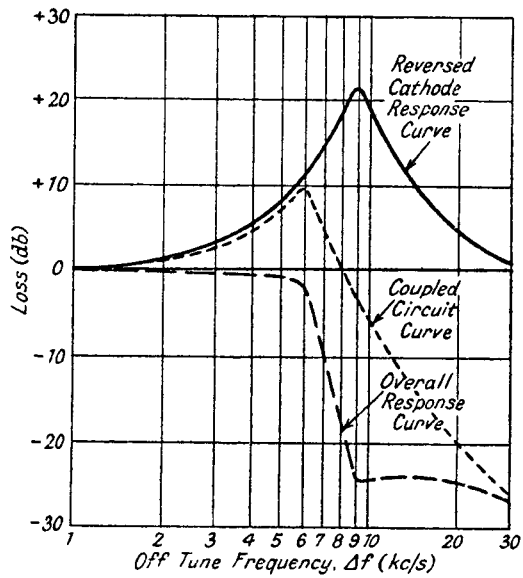


FIG. 7.15.—The Overall Frequency Response Curve for Cathode Feedback Compensation.

circuits (one in the grid and the other in the anode circuit of the valve) have to be compensated. The procedure then is to fit the curve from Fig. 7.14 over Fig. 7.7 with the loss scale of Fig. 7.7 multiplied by 2. The required Q of the overcoupled circuits is then approximately halved.

The constants for the cathode circuits can be determined by using 7.32.

$$R_D = \frac{0{\cdot}414(1+Q^2F_1^2)}{2g_m} = \frac{0{\cdot}414 \times 81}{2 \times 10^{-3}} = 16{,}800\Omega$$

if $g_m = 1$ mA/volt.

$$L = \frac{R_D}{\omega Q} = \frac{16{,}800 \times 10^6}{6{\cdot}28 \times 4{\cdot}65 \times 10^5 \times 225}\ \mu\text{H}$$
$$= 25{\cdot}5\ \mu\text{H}.$$

It is almost impossible to realize practically at 465 kc/s a Q of 225 for such a low inductance, and the most satisfactory method is to adopt transformer coupling to the tuned circuits. The resonant impedance of a transformer coupled tuned circuit is given in expression 4.29*b* (Section 4.4.2) as $\frac{\omega^2 M^2}{R_T}$ and it can be written as

$$\frac{\omega^2 M^2}{R_T} = \frac{\omega^2 M^2}{L_1 L_T} \cdot \frac{L_1 L_T}{R_T} = \frac{M^2}{L_1 L_T} \cdot \omega L_1 \cdot \frac{\omega L_T}{R_T} = k^2 \omega L_1 Q$$

where L_1 = inductance of the cathode coil
Q = the magnification of the tuned secondary coil.

By adjusting k any value of secondary coil inductance may be chosen to give the required resonant impedance.

Variable selectivity is produced in the normal manner by varying the mutual inductance between the coupled circuits, and an almost level pass-band response is maintained as selectivity is increased.

The cathode feedback curve is affected by varying the gain of the valve, and increasing bias decreases the loss. It is therefore possible to obtain variable selectivity by varying the grid bias of the valve. For example, suppose that the coupled circuits are adjusted to give a double-peaked frequency response, and the cathode circuits are designed to overcompensate at maximum gain so that a single-peaked overall response curve is obtained. Increase of bias reduces the cathode compensation and widens the pass-band. At high negative biases the cathode circuits have practically no effect, and the overall frequency response is the wide band double-peaked curve of the coupled circuits. The application of A.G.C. to such a stage automatically produces variable selectivity, with high selectivity for weak signals and low selectivity for strong signal voltages.

7.10. Automatic Variable Selectivity.[7, 11, 12] Many types of circuits have been developed for varying automatically the selectivity of a receiver. The control may be exercised by the desired signal or the interference. The ideal arrangement is by differential control from both, so that either decreasing desired signal or increasing interference causes an increase in selectivity. Methods of achieving automatic selectivity may be conveniently divided into three groups. In the first, selectivity is controlled by varying the damping of the tuned circuits,[8] in the second by varying the coupling reactance [13] between the circuits, and in the third by mistuning the circuits.

Variation of damping can be obtained by paralleling the tuned

circuits with triode valves, the anode-cathode resistance of which is controlled by varying their grid bias. The valves are initially biased so that anode current is almost cut off and the anode-cathode resistance is consequently very high. Positive bias, increasing as the desired signal increases, is derived from the detector load resistance, and the anode current of the damping valves thus increases, their resistance falls and selectivity decreases. This method is not very satisfactory because the ideal selectivity curve shape cannot be produced by damping, and widening of the pass-band is accompanied by appreciable loss of attenuation outside the band. In addition, owing to curvature of the valve characteristics, the anode-cathode resistance is not constant over the signal voltage cycle, and distortion of the latter occurs.

Automatic selectivity control by coupling reactance variation may be obtained by a relay, varying the disposition of the coupling coils between the primary and the secondary of the I.F. transformers. The relay may be controlled by the D.C. component of the input carrier at the detector, increasing the coupling as the carrier increases. Polarized variable capacitors [17] (as described in Chapter 13) may be employed as variable shunt and series couplings between the primary and secondary circuits of the I.F. transformer. Their capacitance values are controlled by varying the D.C. potential between their plates. The series capacitance coupling chiefly controls the low frequency side of the selectivity curve, whilst the shunt coupling controls the high frequency side. Increase of selectivity results from a decrease in the series capacitance and an increase in the shunt capacitance. For the most favourable conditions of reception the pass-range is a maximum, and the selectivity curve has a double-humped frequency response. Decrease of the series capacitance increases the frequency of the lower frequency peak, thus narrowing the pass-band on the low-frequency side of the carrier, whilst increase of the shunt capacitance decreases the frequency of the higher peak and narrows the pass-band on that side of the carrier. The selectivity curve can therefore be narrowed from either side, and the control voltages for the capacitors are derived from a discriminator circuit (similar to that for automatic frequency control and described in Chapter 13) responsive to the interference. Interference on the low-frequency side of the carrier produces a voltage reducing the series capacitance value, and that on the high-frequency side increases the shunt capacitance value.

Mistuning of the primary and secondary of the I.F. transformers in opposite directions, using these polarized capacitors, may also

be used to give the same result, but overall amplification is reduced.

A circuit [19] which combines mistuning with increased selectivity in the I.F. amplifier is shown in Fig. 7.16. The mistuning is brought about by moving the carrier away from the centre of the I.F. amplifier pass-band and in the direction of the interfering voltage. This asymmetric expansion is produced by the discriminator circuits T_1 and T_2 (generally tuned to $f_m \pm 9$ kc/s), the detected voltages from which are connected in series with the automatic frequency correction bias voltage to the variable reactance valve across the oscillator-tuned circuit. These discriminator circuits T_1 and T_2 are also used to produce a flat pass-band response and sharp cut-offs as described in Section 7.9. The A.F.C. discriminator circuits T_3 and

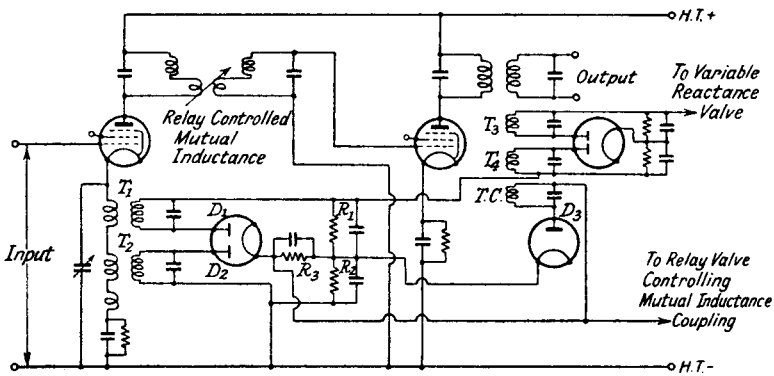


FIG. 7.16.—Automatic Variable Selectivity Control by Differential Action of the Signal and Interference Voltages.

T_4 correct for normal tuning errors and oscillator drift (see Section 13.4). The extra differential bias developed across the resistances R_1 and R_2 reacts on the oscillator frequency to move the interference more into the attenuating part of the I.F. amplifier frequency response. The overall width of the pass-band of the amplifier is controlled by a valve-operated relay connected across the resistance (R_3) in common with diodes D_1 and D_2. The relay controls the mutual inductance coupling between the I.F. tuned circuits. The circuit T.C. is tuned to the I.F. carrier frequency and its output is detected by the diode D_3, which is connected to R_3, so that the detected current opposes that due to the interference. Thus the I.F. selectivity is controlled by differential action of interference and desired signal, increase of the former and decrease of the latter increasing selectivity.

7.11. Signal Handling Capacity of the I.F. Amplifier Valve. For A.G.C. design purposes the signal-handling capacity of the I.F. amplifier valve requires to be known at different grid biases. The signal-handling capacity of a valve is defined as the maximum input (and output) carrier voltage, modulated at a given percentage, which can be accepted for a given percentage distortion of the modulation envelope. The percentage modulation should be as high as possible, and its maximum value is limited to some extent by distortion of the modulation envelope at the detector. A modulation percentage of 60 is a convenient practical value for measurement purposes, but 80% may be used if the detector circuit is designed to reduce envelope distortion (mainly due to an A.C. to D.C. load-resistance ratio less than unity) to a minimum. A usual value for audio frequency harmonic distortion is 5%, and

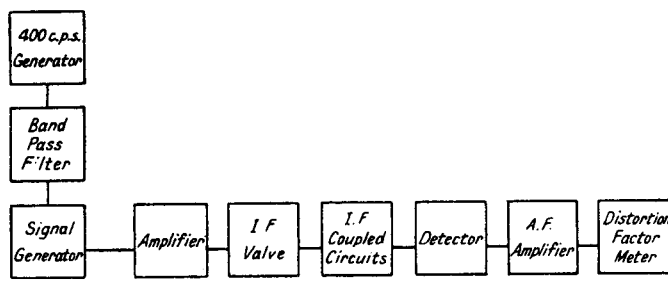


FIG. 7.17.—A Schematic Diagram for the Measurement of Signal Handling Capacity.

the results are expressed in the form of curves of input and output carrier peak voltage against grid bias for 5% total harmonic distortion of the audio output voltage from the detector.

A schematic diagram of the apparatus is shown in Fig. 7.17 and it is mainly self-explanatory. The modulating source has a frequency of 400 c.p.s., the normal test frequency for receiver measurements, and it is filtered so as to reduce distortion in the modulating voltage to very small values. The signal generator should be anode-modulated direct from the filtered 400 c.p.s., because modulation by anode voltage variation generally gives least modulation envelope distortion. An amplifier is required between the signal generator and I.F. valve because the maximum input signal carrier voltage is generally about 8 volts, and it may conveniently be a tetrode valve with an anode circuit tuned to the I.F. A diode detector, previously calibrated, is a suitable measuring device for the I.F. input carrier. The coupled circuits in the I.F. valve anode must be adjusted for single-peaked frequency response in order to

eliminate the possibility of asymmetrical side-band amplification due to mistuning. The diode detector has special features and the circuit is shown in Fig. 7.18. The D.C. load resistance consists of 0·5 MΩ fixed resistance, R_1, in series with a 20,000Ω potentiometer (R_2). A microammeter, which measures the mean current due to the output carrier voltage, is inserted in the detector cathode circuit. The A.F. output is taken from the potentiometer through a coupling capacitance C_2 and R.F. filter R_3C_3 to the grid of an A.F. amplifier. A 2 MΩ grid leak R_4 completes the D.C. path for the amplifier.

These precautions are taken in order to prevent the modulation

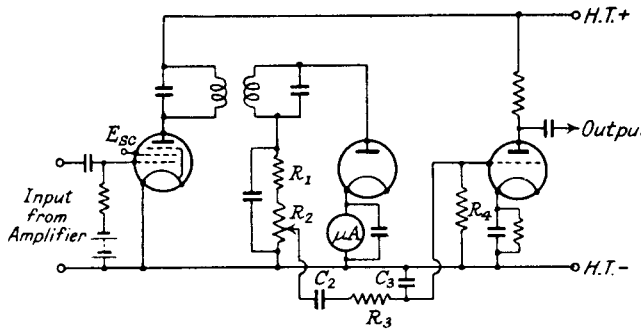


FIG. 7.18.—The A.F. Valve and Detector Circuit for the Measurement of Signal Handling Capacity

envelope distortion mentioned in Section 8.2.3. The maximum permissible modulation percentage is approximately

$$\frac{2 \times 10^6}{2{\cdot}03 \times 10^6} \times 100 = 98{\cdot}5\%,$$

so that distortion due to the detector A.C. load resistance is not likely to occur at a test modulation percentage of 80%. The first A.F. amplifier valve is followed by a second, preferably a power output triode, the output from which is connected to a distortion factor meter.

Care must be taken to ensure that distortion in the apparatus other than the test valve is as small as possible ; a minimum value of about 1% may be expected at 80% modulation. Distortion in the A.F. amplifier may be checked by connecting the output from the 400 c.p.s. filter to the 20,000-ohm potentiometer. The diode detector valve must be removed from its socket when the test is made. Distortion in the amplifier following the signal generator may be checked by connecting the diode detector to its output.

The procedure is to increase the input signal until 5% total harmonic distortion is registered on the distortion factor meter. Typical curves are shown in Fig. 7.19, and their usefulness is discussed in Section 12.4.2. A point which should be noted is that at high negative biases there may be two input voltages for 5% distortion, one a high value and the other a very low one. The first should be regarded as the correct, since the low value can only be found by starting from zero signal at the high negative bias. The high value is always obtained when the input is (as is normal in A.G.C.) increased as the bias is increased.

The mutual conductance of the valve under operating conditions

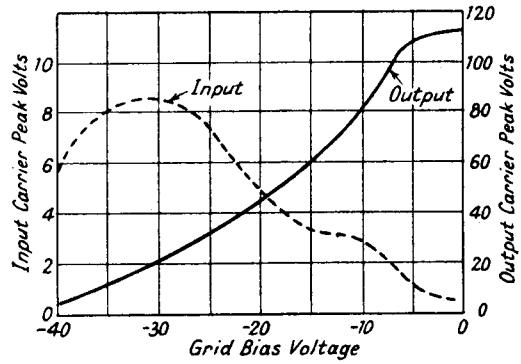


FIG. 7.19.—Typical Signal Handling Capacity Curves for an I.F. Amplifier for 5% Modulation Envelope Distortion and 80% Modulation of the Carrier.

may be calculated from the grid input carrier voltage, the output carrier voltage across the primary of the I.F. transformer, and the primary impedance. Thus

$$g_m = \frac{\hat{E}_p}{\hat{E}_s R_{Dp}}$$

where $\hat{E}_p$ = carrier peak voltage across the primary
$\hat{E}_s$ = ,, ,, ,, applied to the grid
R_{Dp} = resonant impedance of the primary.

R_{Dp} may be measured by using a peak voltmeter and a non-inductive resistance as described in Section 5.6.4. The current in the diode detector across the secondary may be used as a measure of primary volts by calibrating it against peak primary volts as measured by the peak voltmeter.

The signal handling capacity of an R.F. amplifier valve or a frequency changer may be measured in the same way. In th case

of a frequency changer a local oscillator is required and the signal generator carrier frequency is adjusted to any convenient signal frequency.

There are methods of measuring indirectly, and of calculating the signal-handling capacity of a valve, and a description of these is given in Sections 4.7.1 and 4.7.2.

BIBLIOGRAPHY

1. Analysis of Frequency in Oscillating Circuits. G. W. O. Howe, *Electrical World* (New York), August 19th, 1916.
2. Band-Pass Filters in Radio Receivers. G. W. O. Howe, *Wireless Engineer*, May 1931, p. 233.
3. The Analysis and Design of a Chain of Resonant Circuits. M. Reed, *Wireless Engineer*, May (p. 259) and June (p. 320), 1932.
4. Coupling and Coupling Coefficients. G. W. O. Howe, *Wireless Engineer*, Sept. 1932, p. 485.
5. Two Element Band Pass Filters. R. T. Beatty, *Wireless Engineer*, Oct. 1932, p. 546.
6. High Fidelity Receivers with Expanding Selectors. H. A. Wheeler and J. K. Johnson, *Proc. I.R.E.*, June 1935, p. 594.
7. Automatic Selectivity Control. B. D. Corbett, *Wireless World*. Nov. 29th, 1935, p. 554.
8. Automatic Selectivity Control. G. L. Beers, *Proc. I.R.E.*, Dec. 1935, p. 1,425.
9. Variable Selectivity and the I.F. Amplifier. W. T. Cocking, *Wireless Engineer*, March (p. 119), April (p. 179), May (p. 237), 1936.
10. Variable Selectivity. H. J. Benner, *Radio Engineering*, May 1936, p. 12.
11. Automatic Band Width Regulation. O. Köhler, *Funktechnische Monatshefte*, Sept. 1936, p. 337.
12. Variable Fidelity Control and Automatic Selectivity Control. A. W. Barber, and H. F. Mayer, *Radio Engineering*, Jan. 1937, p. 5.
13. American Automatic Selectivity Control. *Wireless World*, March 26th, 1937, p. 296.
14. Resistance Control of Mutual Inductance Coupling. K. R. Sturley, *Marconi Review*, March–April 1937, p. 1.
15. The Design of Coupling Filters in Broadcast Receivers. G. W. O. Howe, *Wireless Engineer*, June (p. 289) and July (p. 347), 1937.
16. Frequency Selective Feedback Applied to the Design of Band Pass Amplifiers. J. D. Brailsford, *Marconi Review*, January–March 1938, p. 10.
17. Broadcast Receivers. A Review. N. M. Rust, O. E. Keall, J. F. Ramsay and K. R. Sturley, *Jour. I.E.E.*, June 1941, Part III, p. 59.
18. British Patent No. 490,819. N. M. Rust, J. D. Brailsford and Marconi's Wireless Telegraph Coy.
19. British Patent No. 501,760. O. E. Keall, N. M. Rust and Marconi's Wireless Telegraph Coy.

CHAPTER 8

DETECTION

8.1. Introduction. To obtain the audio frequency intelligence conveyed in an amplitude modulated carrier it is necessary to employ a detector to suppress or reduce the amplitude of one-half of the modulation envelope. An example of a modulated wave before and after detection has already been given in Figs. 1.1 and 1.4. The suppression of the negative modulation envelope (shown in Fig. 1.4) produces a mean voltage which is an exact reproduction, to a smaller amplitude, of the modulation envelope. This mean voltage variation can be applied to an A.F. amplifier terminated by

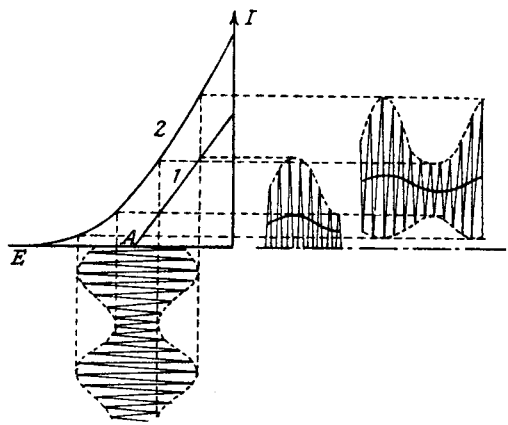


FIG. 8.1.—Detection with Linear (1) and Parabolic (2) Characteristics.

telephones or a loudspeaker, from which sound waves corresponding to the original audio frequencies modulating the carrier will be heard.

In the case of frequency and phase modulated transmission the modulation is first converted to amplitude modulation before application to the detector.

All detectors possess one feature in common, viz., their current-voltage characteristic curve is non-linear over the operating range, having either a point of discontinuity (Curve 1 in Fig. 8.1) or a curved shape (Curve 2 in the same figure). A detector having a discontinuous characteristic is generally called a linear [6] detector if it is biased to the point of discontinuity A ; the positive modulation envelope is reproduced without distortion and the shape of the mean current wave follows exactly that of the modulation envelope.

A detector having a characteristic similar to that of Curve 2 is generally (somewhat loosely) termed a "square law" or parabolic detector because detection depends on the squared term (a_2E^2) in the expression $I = f(E)$ connecting current and voltage. Such a detector distorts the modulation envelope, and the shape of the mean current wave is not an exact copy of the modulation envelope.

In this chapter we shall consider only valves as detectors, though it should be remembered that detection is not solely a property of valves but also of certain types of crystals and metallic surfaces (the copper oxide detector is an example of the latter). The fundamental principles are, however, the same whatever form the detector may take.

Valve detectors may be conveniently divided into four groups as follows:

(1) Diode
(2) Cumulative or leaky grid
(3) Power grid
(4) Anode bend.

We shall start with the diode as the simplest form of detector.

8.2. Diode Detection.

8.2.1. Introduction. A diode detector consists of a valve having two electrodes. Current flow is unidirectional from anode to cathode and a characteristic I_aE_a curve is shown in Fig. 2.3 (curve 3). The characteristic approximates to a parabolic shape for small values of E_a due to space charge and electron velocity effects, but for large values of E_a it becomes a straight line. In the preliminary theory which follows we shall neglect this initial curvature and assume a straight-line characteristic. Now suppose we apply an A.C. voltage to the detector. If the generator supplying this voltage has zero internal resistance, the current through the diode consists of a series of pulses of a shape approximating to a half-sine wave, and a D.C. milliammeter placed in series with the diode measures a mean current of approximately $\frac{\hat{I}}{\pi}$* where $\hat{I}$ is the maximum current through the diode.

When the generator has an internal resistance R_0, the pulses of

* The mean current in a half-sine wave of peak current $\hat{I}$

$$= \frac{1}{2\pi}\int_0^{\pi} \hat{I} \sin\theta \, d\theta$$

$$= \frac{\hat{I}}{2\pi}\Big[-\cos\theta \Big]_0^{\pi} = \frac{\hat{I}}{\pi}.$$

current are still half sinusoidal in shape but reduced in amplitude, giving a lower reading in the D.C. milliammeter.

Now let us assume that the voltage applied to the diode is a modulated R.F. carrier $\hat{E} \cos \omega t\ (1+M \cos pt)$. The amplitude variations of the carrier produce corresponding variations in the detected current pulses, but a D.C. milliammeter will give a reading which is dependent only on the " average " carrier voltage and not upon the modulation. This is so because a decrease of carrier

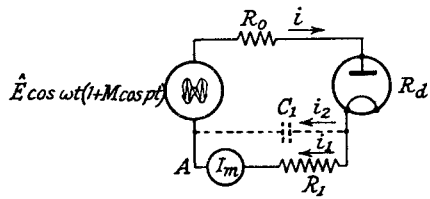


FIG. 8.2.—A Diode Detector.

amplitude (due to modulation) at one time instant is exactly counterbalanced by an increase at another time instant. Since, however, the pulses of current are varying in amplitude at an audio frequency there must be an A.C. current of that frequency in addition to the D.C. current measured by the milliammeter. The presence of this audio frequency current can be proved by noting that a thermal milliammeter, connected in series with the D.C. milliammeter, gives a higher reading than the latter.

An audio amplifier following the detector requires a voltage input so that a resistance (R_1) must be inserted in the diode circuit as in Fig. 8.2, in order that the mean current fluctuations may be converted into mean voltage variations. The voltage developed

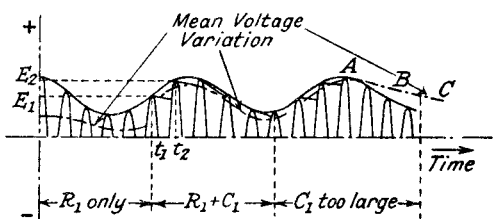


FIG. 8.3.—The Wave Shape of the Voltage across the Load Resistance in a Diode Detector.

across R_1 takes the form of half-sinusoidal pulses as shown to the left in Fig. 8.3. The reference point is A (Fig. 8.2), so that the pulses are in a positive direction. Their amplitude is dependent on R_1, the diode resistance R_d, the generator resistance R_0 and the

modulation envelope. If $R_1 \gg R_d + R_0$ the mean value of the voltage across R_1 is the mean value of half the input modulated carrier, i.e.,

$$\frac{\hat{E}(1+M\cos pt)}{2\pi}\int_{-\frac{\pi}{2\omega}}^{\frac{\pi}{2\omega}} \cos \omega t.dt = \frac{\hat{E}}{\pi} + \frac{\hat{E}M}{\pi}\cos pt,$$

the first term in the expression is the D.C. component and the second the A.C. component.

It is clear that such a method of detection is inefficient since the available output A.F. signal is reduced to approximately one-third $\left(\frac{1}{\pi}\right)$ of the input modulation envelope. The problem becomes one of increasing the mean voltage across R_1 by filling in the gaps between the pulses. This can be achieved by connecting a capacitance C_1 across R_1. Owing to its low reactance at radio frequencies it effectively short circuits R_1 and allows almost all the R.F. input voltage to be applied to the diode. Hence, during the diode conduction period, the detected current is only limited by the diode and generator resistance. This current is used to charge the capacitance C_1, the voltage across which rises nearly to the maximum value E_1 of the R.F. carrier, reached at some time instant t_1 (Fig. 8.3). After the time instant t_1 the voltage across R_1 tends to fall, but now the capacitance C_1 charged to E_1, begins to discharge through the resistance R_1, thus tending to maintain the voltage. The capacitance C_1 continues to discharge until a point on the next positive cycle when the input voltage equals that across C_1, and the diode conducts charging C_1 up to E_2. Then discharge begins again and the cycle is repeated. It will be seen that the discharge of C_1 has filled in the gaps between the voltage pulses and the mean voltage variation across R_1 is very little less than the peak voltage variation of the input wave. Thus the loss in the detector stage is almost entirely eliminated. The maximum permissible value of capacitance C_1 depends on the highest modulation frequency employed and the maximum modulation percentage, since if C_1 is too large the voltage across R_1 may be unable to fall to a value lower than the next positive peak. In this circumstance the mean voltage variation across R_1 is not a faithful reproduction of the modulation envelope but will be distorted when the modulation envelope decreases as shown by the line ABC in Fig. 8.3. This distortion is often called " non-tracking " distortion. If the value of C_1 is too small,[7] the detected voltage will be reduced because the gaps between the pulses are inadequately filled.

8.2.2. Diode Detection Characteristic Curves. If we assume that C_1 fills in completely the gaps between successive voltage pulses, the D.C. current through R_1 is determined solely by R_1 and the diode conduction resistance R_d. This means that we

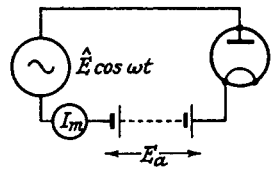


FIG. 8.4*a*.—The Circuit for Obtaining I_mE_a Curves for a Diode Detector.

may replace R_1C_1 by a fixed voltage, i.e., from a battery, and obtain a series of I_mE_a curves (I_m is the mean D.C. current through the diode), which have properties similar to those of the I_aE_a curves for a triode. It is possible on these curves (just as for triode I_aE_a curves) to draw load lines corresponding to the load resistance R_1 and to determine the output voltage for given input conditions. The circuit for obtaining a series of these curves (Fig. 8.4*b*) is shown in Fig. 8.4*a*; a low-frequency voltage [10] —the 50-c.p.s. mains supply—may be employed, but if it is desired to include such effects as anode cathode and stray capacitance, these capacitances must be artificially increased to make their reactances at 50 c.p.s. equal to their reactances at the particular radio frequency being considered, e.g., 0·0001 μF at 1,000 kc/s is equivalent to 2 μF at 50 c.p.s. Care must be taken to ensure that the D.C. resistance in the circuit is much less than the diode conduction resistance, R_d, or the curves will be incorrect. For this reason the microammeter measuring mean current should not be a multirange instrument with considerable variation of D.C. resistance between ranges. The battery voltage simulating the voltage produced across R_1 applies a negative voltage to the anode and each curve in Fig. 8.4*b* is obtained by maintaining the input signal voltage constant (at convenient peak values of, for example, 0·5, 1, 1·5, etc.) and varying the battery voltage E_a. The curves of Fig. 8.4*b* are typical of indirectly heated diodes with conduction current beginning at a negative voltage. The curve for $\hat{E}_1 = 0$ is not parallel to the other curves, which actually merge into this curve at certain positive voltages, and the reason for this is explained in Section 8.2.16.

We can now draw across these curves a line corresponding to R_1, making an angle to the horizontal axis of $\theta = \cot^{-1} R_1$, and in the absence of any fixed biasing voltage this starts from $E_a = 0$.

Subject to the condition stated at the beginning of the section, viz., $R_1 \gg R_d + R_0$, we can determine the D.C. voltage, developed across R_1 in the actual detector circuit in which C_1R_1 replace the battery, for any given peak input signal. If we consider the resistance R_1 to be represented by the line OB, an output peak voltage $\hat{E}_1$ gives a mean current in the resistance R_1 of I_1 and a mean voltage across it of E_1. If the input peak voltage rises to $\hat{E}_2$ or falls to $\hat{E}_3$, the mean voltage across R_1 rises to E_2 or falls to E_3. The line OB is therefore the locus for any variation of carrier amplitude. For example, if the input carrier peak voltage $\hat{E}_1$ is

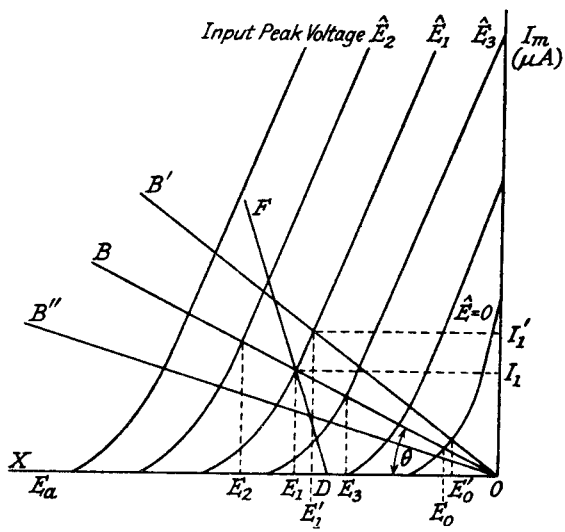


FIG. 8.4*b*.—Typical I_mE_a Curves for a Diode Detector.

modulated such that at the peak of the modulation envelope it reaches $\hat{E}_2$ and at the trough falls to $\hat{E}_3$, the mean voltage varies from E_2 to E_3. Since OB is a straight line, the mean voltage variations are an exact reproduction of the input modulation envelope reduced in amplitude in the ratio $\frac{E_1 - E_0}{\hat{E}_1}$ (E_0 = voltage across R_1 when $\hat{E}_1 = 0$), provided the carrier voltage lines are parallel and equally spaced. The peak value of the audio frequency voltage across R_1 is $\frac{E_2 - E_3}{2}$, since the modulation envelope varies by the voltage difference between positive and negative peaks of the audio frequency wave. The reduction ratio $\frac{E_1 - E_0}{\hat{E}_1}$ is generally called

the detection efficiency η_d and for normal values of R_1 (0·5 MΩ) it is about 0·9. The resistance R_1, in conjunction with R_d, determines η_d and the latter increases as R_1 increases. In most practical circuits C_1 has a low enough reactance and high enough time constant with R_1 to permit Fig. 8.4*b* to be used, but when C_1 is too small, mean detection voltage-carrier peak voltage curves can be used to estimate detector performance. The battery in Fig. 8.4*a* is replaced by the load resistance R_1, paralleled by a capacitance C_1, having a reactance value at 50 c.p.s. equal to the reactance of the actual value of C_1 at the desired radio frequency. Typical mean detection voltage-carrier peak voltage curves for decreasing values of C_1 are shown in Fig. 8.4*c* ; the loss of detection efficiency as C_1 is reduced is clearly indicated. The A.F. output voltage across the

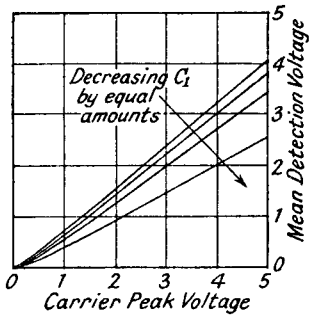


FIG. 8.4c.—Typical Detection Voltage-Carrier Peak Voltage Curves for a Diode Detector.

detector load resistance is estimated as follows : for a carrier peak voltage of 2 volts modulated 50%, carrier amplitude varies between 3 volts and 1 volt peak and the mean detection voltage change corresponding to this carrier change is for the top curve from 0·7 to 2·4 volts, a difference of 1·7 volts. Hence the A.F. peak output voltage for a sinusoidal modulation envelope is $\frac{1{\cdot}7}{2} = 0{\cdot}85$ volts, i.e., detection efficiency is 85%.

8.2.3. Effect of the Coupling Impedance from Diode to A.F. Amplifier.[23] The audio output from the detector is transferred to the A.F. amplifier through a capacitance-resistance coupling such as C_2R_2 shown in Fig. 8.5. Capacitive coupling prevents application of the D.C. component across R_1 to the grid of the amplifier valve, which is biased separately to the optimum point. In addition there is generally a resistance-capacitance filter (R_3C_3) to prevent the passage of R.F. voltages to the audio amplifier,

where they are likely to produce over-loading and perhaps instability. The capacitance of C_3 is about equal to C_1, so that its effect may be neglected at audio frequencies, and R_3 is usually much less than R_1.

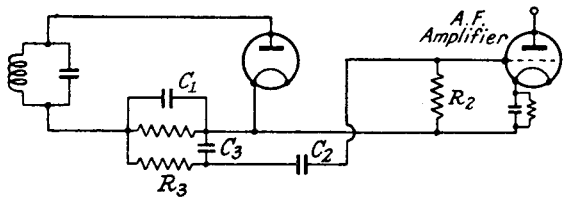


FIG. 8.5.—A Diagram of Connections for the Coupling from a Diode Detector to an A.F. Amplifier.

The actual value of R_3 is to a large extent controlled by R_2, since it forms with R_2 a potentiometer which reduces the A.F. voltage across R_2 to $\frac{R_2}{R_2+R_3}$ of that across R_1. Generally it has a value of about $\frac{R_2}{10}$. Adequate R.F. filtering is then obtained without appreciable reduction of A.F. voltage, or frequency discrimination against

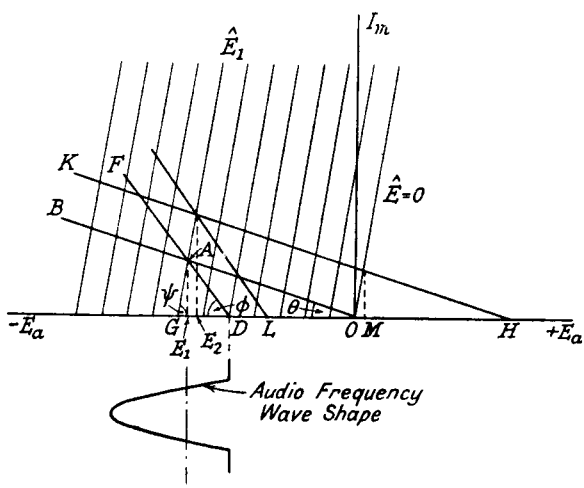


FIG. 8.6.—Ideal I_mE_a Curves for a Diode Detector showing the Effect of the Coupling Resistance R_2 and Positive Bias.
[OH = Positive Bias Voltage.]

the higher audio frequencies by C_3. The coupling reactance of C_2 must be small in comparison with R_2 at the lowest audio frequencies or frequency discrimination results. The resistances

R_2 and R_3 are thus in parallel with R_1 for all audio and radio frequencies. Referring to Fig. 8.6 (the I_mE_a curves are for convenience considered as straight lines) for a modulated input voltage to the detector of $\hat{E}_1[1+M\cos pt]\cos\omega t$, the D.C. operating point is at A, the point of intersection between OB and the voltage line $\hat{E}_1$, and the modulation envelope will travel up and down OB on either side of the point A as long as R_2 is infinite. As R_2 is decreased the locus of the modulation envelope changes to the line DF, where $\cot \widehat{FDG} = \dfrac{R_1R_2}{R_1+R_2}$. We are assuming for simplicity that R_3 is included in R_2. The first effect of R_2 is to reduce the audio voltage from the detector. The second effect, which is much more serious, occurs when the modulation trough falls below the carrier value corresponding to OD. The modulation envelope is cut off and the audio frequency wave shape is distorted as shown. The modulation ratio which can be accepted without distortion is thus limited by R_2. If we assume the I_mE_a characteristic curves for a diode to be straight lines as in Fig. 8.6, the critical modulation ratio is given by $\dfrac{DG}{OG}$.

But
$$DG = DE_1+E_1G$$
$$= AE_1\cot\phi+AE_1\cot\psi \quad . \quad . \quad 8.1$$

and
$$OG = OE_1+E_1G$$
$$= AE_1\cot\theta+AE_1\cot\psi. \quad . \quad . \quad 8.2.$$

From 8.1 and 8.2, the critical modulation ratio is

$$M = \frac{DG}{OG} = \frac{\cot\phi+\cot\psi}{\cot\theta+\cot\psi}$$
$$= \frac{\dfrac{R_1R_2}{R_1+R_2}+R_d'}{R_1+R_d'} \qquad 8.3.$$

It is important to note that R_d', the inverse of the slope of the diode I_mE_a characteristic curve, is the diode equivalent resistance to changes of voltage across R_1 and is not the diode conduction resistance R_d of Section 8.2.5. Its relationship to R_d is discussed in Sections 8.2.15 and 16.

Since detection efficiency $\eta_d = \dfrac{OE_1}{OG} = \dfrac{\cot\theta}{\cot\theta+\cot\psi}$

$$= \frac{R_1}{R_1+R_d'} \quad . \quad . \quad . \quad . \quad 8.4$$

rearranging expression 8.3

$$M = \frac{R_1 + R_d' + \dfrac{R_1 R_2}{R_1 + R_2} - R_1}{R_1 + R_d'} = 1 - \frac{R_1^2}{(R_1 + R_2)(R_1 + R_d')}$$

$$= 1 - \eta_d \cdot \left[\frac{R_1}{R_1 + R_2}\right] \quad . \quad . \quad . \quad . \quad . \quad . \quad 8.5a.$$

It is clear from expression 8.5*a* that a high critical modulation ratio requires a high value of R_2. A reduction of detection efficiency also leads to a higher critical modulation ratio. This is to be expected since lower detection efficiency means either a reduced value for R_1 or increased value for R_d', i.e., a smaller angle ψ, which brings a lower carrier voltage line to the point D on Fig. 8.6. An equivalent increase in R_2 can be realized by tapping down the detector D.C. resistance R_1. Suppose R_2 is tapped across a part $KR_1 (K < 1)$ only, the A.C. load resistance is

$$(1 - K)R_1 + \frac{KR_1 R_2}{KR_1 + R_2} = R_1 - \frac{(KR_1)^2}{KR_1 + R_2}$$

and from 8.3 the critical modulation ratio is

$$M = \frac{R_1 - \dfrac{(KR_1)^2}{KR_1 + R_2} + R_d'}{R_1 + R_d'}$$

$$= 1 - \eta_d \frac{K^2 R_1}{KR_1 + R_2}$$

$$= 1 - \eta_d K \cdot \left(\frac{R_1}{R_1 + \dfrac{R_2}{K}}\right) \quad . \quad . \quad . \quad 8.5b.$$

Not only is R_2 equivalently increased to $\dfrac{R_2}{K}$, but the negative term is also reduced by the multiplying factor K.

The critical modulation ratio may also be increased by applying a positive bias [29] to the anode of the diode. The carrier locus line takes up a new position such as HK where OH is the positive bias voltage. The critical modulation ratio may be increased to 1 for any given value of R_2 and carrier peak voltage $\hat{E}_1$ by a suitable choice of OH. Expression 8.5*a* is modified by the addition of positive bias to

$$M' = \frac{LG}{HG} = 1 - \eta_d \cdot \left(\frac{R_1}{R_1 + R_2}\right)$$

where $\eta_d = \dfrac{ME_2}{OG} = \dfrac{OE_1}{OG} = \dfrac{R_1}{R_1 + R_d'}$ since HK is parallel to OB.

The carrier modulation ratio, which is not M', is given by

$$M = \frac{LG}{OG} = \frac{LG}{HG} \cdot \frac{HG}{OG}$$

$$M = M'\left(\frac{\hat{E}_1 + E_b}{\hat{E}_1}\right) \text{ where } E_b = OH$$

thus

$$M = \left[\frac{\hat{E}_1 + E_b}{\hat{E}_1}\right]\left(1 - \eta_d\left(\frac{R_1}{R_1 + R_2}\right)\right) \quad . \quad . \quad 8.6.$$

The critical modulation ratio is now dependent on the carrier peak voltage and the positive bias applied to the diode. When this critical modulation ratio is exceeded the D.C. current in the resistance R_1 is increased owing to the fact that the bottom of the audio frequency voltage wave is cut off, i.e., is itself being detected. Thus any change of D.C. current in the diode circuit of a receiver, when a programme is being received from a steady carrier voltage source such as a local station, indicates distortion of the audio frequency output voltage by the detector stage.

There is a disadvantage to the use of positive bias due to the fact that the detector is generally supplied from a high impedance generator (a tuned circuit in the anode of a valve). A positively biased diode conducts in the absence of an input signal and so presents a low input resistance to the generator. This heavy damping is reduced as the input signal is increased, and the diode begins to function as a unidirectional device (see Section 8.2.6). Two effects are obtained; the detector is insensitive to signal inputs less than a certain value (about 0·3 of the positive bias) and it also distorts the modulation envelope of the applied signal. Hence for really satisfactory operation the positive bias needs to be variable and directly controlled by the input signal (decreasing as the signal decreases and vice versa).

In the above analysis we have assumed the diode $I_m E_a$ characteristic curves to be straight lines, but in Section 8.2.16 it is shown that a curved shape is obtained even for a diode with a linear $I_a E_a$ characteristic. This curvature tends to give a higher critical modulation ratio than that given in expressions 8.5*a*, 8.5*b* and 8.6.

8.2.4. Damping of the Input Circuit due to the Diode Conduction Current. The conduction current of the diode constitutes a load which may considerably modify the sensitivity and selectivity characteristics of the input, when the latter contains a tuned circuit. The load effect may be estimated by averaging over the complete R.F. cycle of the input voltage the resistance introduced

during the conduction period. In this way the equivalent damping resistance shunting the input can be calculated.

Two cases will be considered; in the first the diode will be assumed to have a linear I_aE_a characteristic given by $I_a = KE_a$, and in the second the diode characteristic will be assumed to be parabolic and given by $I_a = KE^2{}_a$.

8.2.5. Equivalent Damping Resistance due to a Diode with Linear I_aE_a Characteristic. We shall assume that the I_aE_a curve ($\hat{E}_1 = 0$) is a straight line given by the relationship

$$I_a = \frac{E_a}{R_d}.$$

Where R_d = diode conduction resistance, i.e., the inverse of the slope of this characteristic.

The diode conducts when the input voltage $\hat{E}_1 \cos \omega t$, which for convenience may be re-written $\hat{E}_1 \cos \theta$, exceeds the voltage across the resistance R_1 (Fig. 8.7).

The D.C. voltage (E_1) across the resistance-capacitance parallel

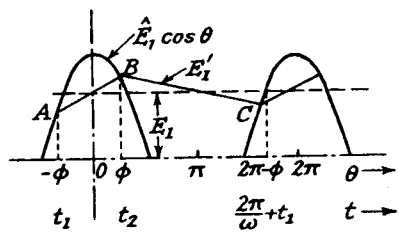


FIG. 8.7.—The Voltage Wave Shape Across the Load Resistance R_1 of a Diode Detector.

combination will not in actual practice remain constant but will fluctuate in the manner shown by the lines AB and BC in Fig. 8.7. This fluctuation, if taken into account, considerably complicates the analysis and its effect may be neglected except when the detection efficiency is much lower than is normally encountered in practice. It will therefore be assumed that E_1 is a constant voltage.

Mean Current I_m, and the Relationship between R_1, R_d, η_d.[20]

The expression for the mean current in the circuit is

$$I_m = \frac{1}{2\pi}\int_{-\phi}^{\phi} I_a \, d\theta$$

$$= \frac{1}{2\pi}\int_{-\phi}^{\phi} \frac{(\hat{E}_1 \cos\theta - E_1)d\theta}{R_d} \quad . \quad . \quad 8.7$$

where $\hat{E}_1 \cos \phi = E_1$, the voltage at which the diode conducts.

The limits of equation 8.7 may conveniently be changed to ϕ and 0 and the mean current then becomes

$$I_m = \frac{1}{\pi}\int_0^{\phi} \frac{(\hat{E}_1 \cos\theta - E_1)}{R_d} d\theta$$

$$= \frac{1}{\pi R_d}\{\hat{E}_1 \sin\phi - E_1\phi\}$$

$$= \frac{\hat{E}_1}{\pi R_d}\left\{\sin\phi - \frac{E_1}{\hat{E}_1}\phi\right\} \quad . \quad . \quad . \quad 8.8.$$

But $\cos\phi = \dfrac{E_1}{\hat{E}_1} = \eta_d$ = detection efficiency

thus $\sin\phi = \sqrt{1 - \eta_d{}^2}$

and $\phi = \cos^{-1}\eta_d$.

Replacing the expressions containing ϕ in 8.8 we get

$$I_m = \frac{\hat{E}_1}{\pi R_d}\{\sqrt{1 - \eta_d{}^2} - \eta_d . \cos^{-1}\eta_d\} \quad . \quad . \quad 8.9.$$

But $I_m = \dfrac{E_1}{R_1}$

$$\therefore \frac{R_1}{R_d} = \frac{\pi\eta_d}{(\sqrt{1 - \eta_d{}^2} - \eta_d \cos^{-1}\eta_d)} \quad . \quad . \quad 8.10.$$

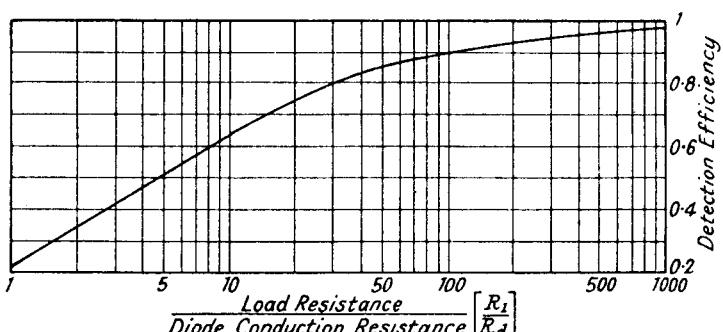


FIG. 8.8.—A Curve of Detection Efficiency against the Ratio of Load to Diode Conduction Resistance.

Expression 8.10 shows the relationship between the detection efficiency and the ratio $\dfrac{R_1}{R_d}$ and it is plotted in Fig. 8.8 for various values of $\dfrac{R_1}{R_d}$.

Effective Input Resistance R_E.

The effective input resistance may be determined by equating the power absorbed in the circuit to $\frac{\hat{E}_1^2}{2R_E}$, where R_E is the equivalent shunt resistance due to the diode conduction current.

Power absorbed in diode load resistance R_1

$$= \frac{1}{\pi}\int_0^{\phi} \hat{E}_1 \cos\theta \, I_a \, d\theta \qquad . \quad . \quad . \quad . \quad 8.11$$

$$= \frac{1}{\pi R_d}\int_0^{\phi} \hat{E}_1 \cos\theta [\hat{E}_1 \cos\theta - E_1] \, d\theta$$

$$= \frac{\hat{E}_1^2}{\pi R_d}\int_0^{\phi}\left[\frac{1+\cos 2\theta}{2} - \frac{E_1}{\hat{E}_1}\cos\theta\right] d\theta$$

$$= \frac{\hat{E}_1^2}{2\pi R_d}\left\{\phi + \frac{\sin 2\phi}{2} - \frac{2E_1}{\hat{E}_1}\sin\phi\right\}$$

$$= \frac{\hat{E}_1^2}{2\pi R_d}\{\cos^{-1}\eta_d - \eta_d\sqrt{1-\eta_d^2}\}$$

$$= \frac{\hat{E}_1^2}{2R_E}$$

$$\therefore \; R_E = \frac{\pi R_d}{\cos^{-1}\eta_d - \eta_d\sqrt{1-\eta_d^2}} \qquad . \quad . \quad . \quad 8.12.$$

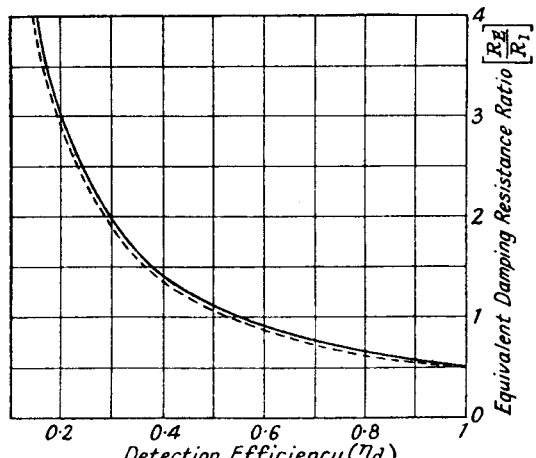


FIG. 8.9.—Curves of Damping Resistance Ratio against Detection Efficiency.
(Full Line—Linear Characteristic.
Dotted Line—Parabolic Characteristic.)

Inserting the value of R_d from 8.10 in 8.12

$$\frac{R_E}{R_1} = \frac{[\sqrt{1-\eta_d{}^2} - \eta_d \cos^{-1}\eta_d]}{\eta_d(\cos^{-1}\eta_d - \eta_d\sqrt{1-\eta_d{}^2})} \quad . \quad . \quad 8.13$$

The ratio $\frac{R_E}{R_1}$ is plotted in Fig. 8.9 for different values of η_d and it can be noted that the equivalent resistance R_E approaches $\frac{1}{2}R_1$ as η_d approaches unity. As η_d decreases R_E increases and at $\eta_d = 0{\cdot}55$, R_E is equal to R_1.

It is not usual for the diode to conduct exactly at $E_a = 0$, and in the more general case conduction starts when E_a is either positive or negative. In directly heated filament valves E_a is most often positive, whereas for indirectly heated mains valves with equipotential cathodes E_a is usually negative. The two conditions are more easily considered separately. Let us take first the case of the mains valve with current starting at $-E_0$.

8.2.6. Equivalent Damping Resistance for Conduction Current Beginning at a Negative Anode Voltage. The characteristic curve is given by line AB in Fig. 8.10, and the I_aE_a relationship by

$$I_a = \frac{(E_a+E_0)}{R_d}$$

where $-E_0$ is the voltage at which current starts. The expression E_0 thus represents a numerical and not an algebraic value.
The expression for mean current becomes:

$$I_m = \frac{1}{\pi R_d}\int_0^{\phi}(\hat{E}_1 \cos\theta - E_1+E_0)\,d\theta \quad . \quad . \quad 8.14a$$

$$= \frac{\hat{E}_1}{\pi R_d}\left[\sin\phi - \frac{E_1 - E_0}{\hat{E}_1}\phi\right] \quad . \quad . \quad 8.14b$$

where E_1 = D.C. mean voltage developed across R_1.

The signal voltage must exceed a certain value before detection can take place, since the diode conduction current must be cut off at some part of the cycle. Even though no detection takes place there is still a mean current flowing round the diode circuit due to the negative start of current. Its value, assuming the input circuit to have no resistance, will be $\frac{E_0}{R_1+R_d}$, and ϕ in expression 8.14*b* will be π

thus
$$I_m = \frac{E_0}{R_1+R_d} = \frac{\hat{E}_1}{\pi R_d}\left(-\frac{E_1' - E_0}{\hat{E}_1}\right)\pi$$
$$= \frac{E_0 - E_1'}{R_d}$$

where E_1' = voltage across the load resistance R_1 for no detection,

thus
$$E_1' = \frac{E_0 R_1}{R_1 + R_d} \quad . \quad . \quad . \quad . \quad 8.15.$$

The value of $\hat{E}_1$ which just fails to produce detection may be found by making the mean current (8.14a) due to detection equal to 0

or
$$\hat{E}_1' \cos\theta - E_1' + E_0 = 0$$

but
$$\cos\theta = \frac{E_1' - E_0}{\hat{E}_1'} = -1$$

when $\theta = \pi$ (the condition for no detection).

$$\therefore \hat{E}_1' = E_0 - E_1' = E_0 - \frac{E_0 R_1}{R_1 + R_d}$$
$$= \frac{E_0 R_d}{R_1 + R_d}.$$

That this is the voltage required to start detection may be seen from Fig. 8.10. AB is the $I_a E_a$ characteristic curve for the diode and AG the D.C. line for the diode resistance R_d and R_1. The

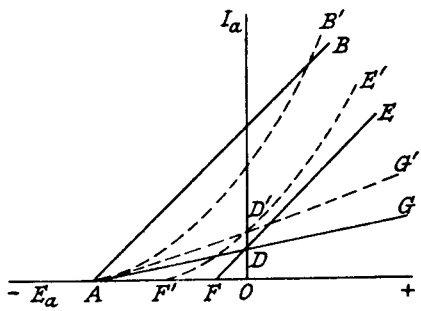


FIG. 8.10.—Operating Conditions for a Diode with Negative Start of Conduction Current.
(Full Line—Linear Conduction Characteristic.
Dotted Line—Parabolic Conduction Characteristic.)

reactance of the capacitance C_1 at radio frequencies is assumed to be negligible compared with R_d. Hence the only impedance to R.F. voltages in the circuit is the diode resistance R_d. The locus line for small variations of R.F. voltages is the line FE, which is parallel to AB and passes through D, the D.C. operating point. The current represented by OD is $OA \tan \widehat{DAO}$ which equals $\frac{E_0}{R_1 + R_d}$, so that the voltage OF is

$$OF = OD \cot \widehat{DFO}.$$
$$= \frac{E_0}{R_1 + R_d} . R_d$$

which confirms that no detection occurs until the input peak voltage $\hat{E}_1$ exceeds $\frac{E_0 R_d}{R_1+R_d}$, since only then can the input voltage take the diode into cut-off.

When the input voltage exceeds this critical value, detection takes place and the mean current is given by expression 8.14*b*. We may note that

$$\cos\phi = \frac{E_1 - E_0}{\hat{E}_1} \quad . \quad . \quad . \quad . \quad 8.16a$$

and

$$\eta_d = \frac{E_1 - E_1'}{\hat{E}_1} \quad . \quad . \quad . \quad . \quad 8.16b.$$

Replacing E_1' by $\frac{E_0 R_1}{R_1+R_d}$ and combining expressions 8.16*a* and 8.16*b*.

$$\cos\phi = \eta_d - \frac{E_0 R_d}{(R_1+R_d)\hat{E}_1} = \eta_d - A$$

where

$$A = \frac{E_0 R_d}{(R_1+R_d)\hat{E}_1}.$$

We may now re-write 8.14*b* as follows:

$$I_m = \frac{\hat{E}_1}{\pi R_d}\left[\sqrt{1-(\eta_d - A)^2} - (\eta_d - A)\cos^{-1}(\eta_d - A)\right]$$
$$= \frac{E_1}{R_1}.$$

From which

$$\frac{R_1}{R_d} = \frac{\pi\frac{E_1}{\hat{E}_1}}{\sqrt{1-(\eta_d - A)^2} - (\eta_d - A)\cos^{-1}(\eta_d - A)}$$
$$= \frac{\pi\left(\eta_d + \frac{AR_1}{R_d}\right)}{[\sqrt{1-(\eta_d - A)^2} - (\eta_d - A)\cos^{-1}(\eta_d - A)]} \quad . \quad 8.17a.$$

When $R_d \ll R_1$ and η_d approaches 1 this reduces to

$$\frac{R_1}{R_d} = \frac{\pi\left(\eta_d + \frac{E_0 R_1}{\hat{E}_1(R_1+R_d)}\right)}{\sqrt{1-\eta_d^2} - \eta_d\cos^{-1}\eta_d} \quad . \quad . \quad 8.17b.$$

Detection efficiency is plotted against $\frac{\hat{E}_1}{E_0}$ in Fig. 8.11*a* (curve 1) for $R_d = 5{,}000\Omega$ and $R_1 = 1$ MΩ; $-E_0 = -1$ volt.

The equivalent resistance R_E can be found by the same method as that used in Section 8.2.5, and is given by

$$R_E = \frac{\pi R_d}{\cos^{-1}(\eta_d - A) - (\eta_d - A)\sqrt{1-(\eta_d - A)^2}}$$

replacing R_d by the value derived from 8.17*a*.

$$\frac{R_E}{R_1} = \frac{[\sqrt{1-(\eta_d - A)^2} - (\eta_d - A)\cos^{-1}(\eta_d - A)]}{\left(\eta_d + \frac{AR_1}{R_d}\right)[\cos^{-1}(\eta_d - A) - (\eta_d - A)\sqrt{1-(\eta_d - A)^2}]} \quad . \quad 8.18$$

except where η_d is small, this reduces to

$$\frac{R_E}{R_1} = \frac{\sqrt{1-\eta_d{}^2} - \eta_d \cos^{-1}\eta_d}{\left[\eta_d + \frac{E_0 R_1}{\hat{E}_1(R_1+R_d)}\right][\cos^{-1}\eta_d - \eta_d\sqrt{1-\eta_d{}^2}]}.$$

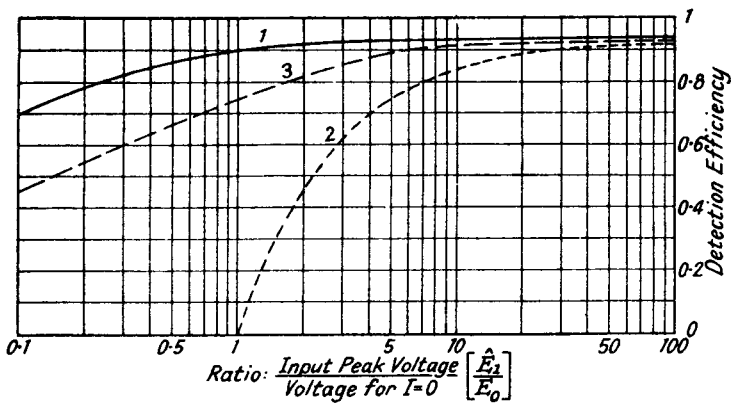


FIG. 8.11*a*.—Curves of Detection Efficiency against Voltage Ratio $\frac{\hat{E}_1}{E_0}$.

Curve 1. $-E_0 = -1$. Linear Characteristic.
Curve 2. $+E_0 = +1$. Linear Characteristic.
Curve 3. $-E_0 = -1$. Parabolic Characteristic.
($R_1 = 1M\Omega$. $R_d = 5{,}000\Omega$.)

It may be noted that when no detection takes place $\eta_d = 0$; $\hat{E}_1' \leq \frac{E_0 R_d}{R_1 + R_d}$, $A = 1$, and $R_E = R_d$.

The value of R_E for $R_d = 5{,}000\Omega$, $R_1 = 1\ M\Omega$, and $-E_0 = -1$ is plotted for various ratios of $\frac{\hat{E}_1}{E_0}$ as curve 1 in Fig. 8.11*b*. The variation of effective resistance with signal voltage is quite marked for $\hat{E}_1 < 10\ E_0$. When $\hat{E}_1 = 0{\cdot}005$ volt the effective resistance is very low indeed and almost equal to the diode resistance R_d. This

undesirable feature can be eliminated by neutralizing the negative start of current with negative bias to the diode anode of $-E_0$ volts. It should be noted that positive bias applied to a diode with conduction current starting at zero voltage has the same effect as current starting at a negative anode voltage.

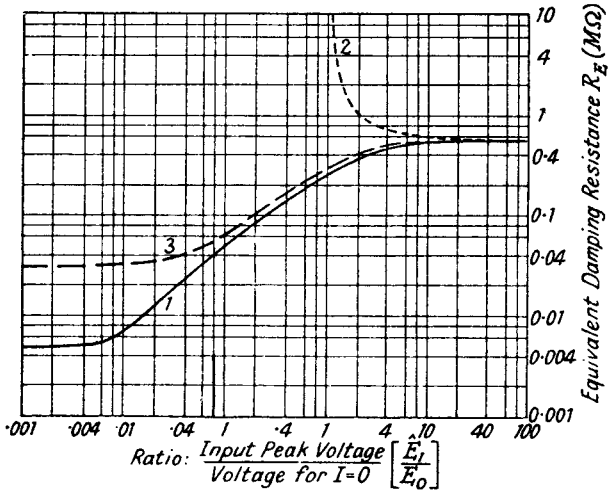


FIG. 8.11*b*.—Curves of Equivalent Damping Resistance against Voltage Ratio $\frac{\hat{E}_1}{E_0}$.

Curve 1. $-E_0 = -1$. Linear Characteristic.
Curve 2. $+E_0 = +1$. Linear Characteristic.
Curve 3. $-E_0 = -1$. Parabolic Characteristic.
($R_1 = 1$ MΩ. $R_d = 5{,}000\Omega$.)

8.2.7. Conduction Current Beginning at a Positive Anode Voltage. The equation to the characteristic curve is

$$I_a = \frac{(E_a - E_0)}{R_d}$$

where $+E_0$ is the positive voltage at which current starts.

By applying the procedure given above

$$I_m = \frac{\hat{E}_1}{\pi R_d}\left[\sin\phi - \frac{E_1+E_0}{\hat{E}_1}\phi\right]$$

$$\cos\phi = \frac{E_1+E_0}{\hat{E}_1}$$

$$\eta_d = \frac{E_1}{\hat{E}_1}$$

$$\cos\phi = \eta_d + \frac{E_0}{\hat{E}_1}.$$

Thus

$$I_m = \frac{\hat{E}_1}{\pi R_d}\left[\sqrt{1-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)^2}-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\cos^{-1}\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\right]. \quad 8.19$$

$$= \frac{E_1}{R_1}.$$

$$\frac{R_1}{R_d} = \frac{\pi\eta_d}{\left[\sqrt{1-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)^2}-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\cos^{-1}\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\right]}.$$

In a similar manner

$$R_E = \frac{\left[\sqrt{1-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)^2}-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\cos^{-1}\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\right]R_1}{\eta_d\left[\cos^{-1}\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)\sqrt{1-\left(\eta_d+\frac{E_0}{\hat{E}_1}\right)^2}\right]} \quad 8.20.$$

The variation of η_d and R_E with $\frac{\hat{E}_1}{E_0}$ for $R_d = 5{,}000\Omega$, $R_1 = 1\ \mathrm{M}\Omega$, $E_0 = +1$ is shown in Curve 2, Figs. 8.11*a* and 8.11*b*, respectively. R_E rises rapidly to infinity when $\hat{E}_1$ approaches E_0 and detection efficiency falls to zero. The condition of positive start of current is thus much less desirable than negative start. A very much larger input voltage is required to begin detection, and the maximum modulation percentage which may be accepted is less, because the trough of the modulation envelope $\hat{E}_1(1 - M)$ must not fall below E_0, whereas for a negative value of E_0 the minimum value of $\hat{E}_1(1 - M)$ is $\frac{E_0 R_d}{R_1+R_d}$. This undesirable feature may be eliminated by the addition of a positive biasing voltage equal to E_0 in series with the diode, which then operates as if current starts at zero anode voltage.

It is interesting to note that this explains why with small signal voltages improved detection is generally obtained by returning the grid leak of a battery valve acting as a cumulative grid detector to the positive lead of the filament. This is equivalent to supplying the grid, which is acting as the anode of a diode, with positive bias, and removing the start of grid current from a positive to a negative value of E_0.

8.2.8. Equivalent Damping Resistance due to a Diode with a Parabolic I_aE_a Characteristic Curve. It is not usual to find that the I_aE_a characteristic curve of a diode is linear for all values of E_a and the relationship is more nearly represented by

$I_a = KE_a^2$ for small signal voltages, approaching $I_a = KE_a$ for large signals. We will therefore examine the effect of a parabolic characteristic on the equivalent damping resistance R_E.

Let us assume that the current-voltage relationship is given by :

$$I_a = \frac{E_a^2}{R_d}$$

where R_d is the inverse of the slope of the I_aE_a characteristic at $E_a = 1$ volt.

The mean current is

$$I_m = \frac{1}{\pi R_d}\int_0^{\phi}(\hat{E}_1 \cos\theta - E_1)^2\, d\theta$$

$$= \frac{E_1^2}{2\pi R_d}\left[\left(1+\frac{2E_1^2}{E_1^2}\right)\phi + \frac{\sin 2\phi}{2} - \frac{4E_1}{\hat{E}_1}\sin\phi\right]$$

but $$\cos\phi = \frac{E_1}{\hat{E}_1} = \eta_d.$$

$$\therefore\ I_m = \frac{\hat{E}_1^2}{2\pi R_d}[(1+2\eta_d^2)\cos^{-1}\eta_d - 3\eta_d\sqrt{1-n_d^2}]. \quad 8.21$$

$$= \frac{E_1}{R_1}.$$

$$\therefore\ \frac{R_1}{R_d} = \frac{2\pi\eta_d}{\hat{E}_1[(1+2\eta_d^2)\cos^{-1}\eta_d - 3\eta_d\sqrt{1-\eta_d^2}]} \quad . \quad 8.22.$$

$$\text{Power absorbed} = \frac{\hat{E}_1^2}{2R_E} = \frac{1}{\pi}\int_0^{\phi}\hat{E}_1\cos\theta I_a\, d\theta$$

$$= \frac{1}{\pi}\int_0^{\phi}\frac{\hat{E}_1\cos\theta\,(E_1\cos\theta - E_1)^2\, d\theta}{R_d}$$

$$= \frac{\hat{E}_1^3}{\pi R_d}\left[\sin\phi\left(\frac{3}{4}+\frac{E_1^2}{\hat{E}_1^2}\right) + \frac{\sin 3\phi}{12} - \frac{E_1}{\hat{E}_1}\phi - \frac{E_1}{\hat{E}_1}\frac{\sin 2\phi}{2}\right]$$

but $$\sin 2\phi = 2\cos\phi\sin\phi = 2\eta_d\sqrt{1-\eta_d^2}$$

$$\sin 3\phi = 3\sin\phi - 4\sin^3\phi = (\sqrt{1-\eta_d^2})(4\eta_d^2 - 1)$$

or $$\frac{\hat{E}_1^2}{2R_E} = \frac{\hat{E}_1^3}{\pi R_d}\left[\left(\frac{\eta_d^2}{3}+\frac{2}{3}\right)\sqrt{1-\eta_d^2} - \eta_d\cos^{-1}\eta_d\right]$$

$$R_E = \frac{\pi R_d}{2\hat{E}_1\left[\left(\frac{\eta_d^2}{3}+\frac{2}{3}\right)\sqrt{1-\eta_d^2} - \eta_d\cos^{-1}\eta_d\right]}. \quad . \quad 8.23.$$

Replacing R_d by the value obtained from 8.22.

$$\frac{R_E}{R_1} = \frac{[(1+2\eta_d^2)\cos^{-1}\eta_d - 3\eta_d\sqrt{1-\eta_d^2}]}{4\eta_d\left[\left(\frac{\eta_d^2}{3}+\frac{2}{3}\right)\sqrt{1-\eta_d^2} - \eta_d\cos^{-1}\eta_d\right]} \quad . \quad 8.24.$$

The ratio $\frac{R_E}{R_1}$ is plotted in Fig. 8.9 against detection efficiency and the curve is almost identical with the one obtained for the linear diode. Both approach $\frac{R_E}{R_1} = \frac{1}{2}$ as η_d is increased to 1, but for low values of η_d a lower value of $\frac{R_E}{R_1}$ is obtained for the parabolic diode.

8.2.9. Conduction Current Beginning at a Negative Anode Voltage. Negative start of current has a slightly different effect from that obtained for the linear diode on account of the detection properties of the square law I_aE_a relationship. Cut-off of current is not an essential as it is in the linear case. This is explained more fully in Section 8.4.1.

The equation for the characteristic curve is

$$I_a = \frac{(E_a+E_0)^2}{R_d}$$

where $-E_0$ is the voltage at which current flow begins. When $E_a = 0$, the voltage across the diode is the difference between E_0 and the voltage across R_1 so that the current flowing in the circuit is

$$I_0 = \frac{[E_0 - I_0R_1]^2}{R_d}$$

or

$$\frac{I_0^2R_1^2}{R_d} - I_0\left[\frac{2E_0R_1}{R_d}+1\right]+\frac{E_0^2}{R_d} = 0.$$

Solving for I_0 we get for the minimum root (the maximum root gives a value for I_0 greater than $\frac{E_0}{R_1}$ which is impossible)

$$I_0 = \frac{1+\frac{2E_0R_1}{R_d} - \sqrt{1+\frac{4E_0R_1}{R_d}}}{\frac{2R_1^2}{R_d}} = \frac{R_d}{2R_1^2}+\frac{E_0}{R_1} - \sqrt{\frac{R_d^2}{4R_1^4}+\frac{E_0}{R_1^3}R_d}.$$

Generally $R_d \ll R_1$, so that

$$I_0 = \frac{E_0}{R_1}\left[1 - \sqrt{\frac{R_d}{E_0R_1}}\right] \quad . \quad . \quad . \quad 8.25a.$$

The voltage across R_1 for zero input volts

$$E_1' = E_0\left[1 - \sqrt{\frac{R_d}{E_0R_1}}\right] \quad . \quad . \quad . \quad 8.25b.$$

The mean current $I_m = \frac{1}{\pi R_d}\int_0^{\phi}[\hat{E}_1 \cos\theta - (E_1 - E_0)]^2\, d\theta$

$$= \frac{\hat{E}_1^2}{\pi R_d}\left[\left(\frac{1}{2} + \frac{(E_1 - E_0)^2}{\hat{E}_1^2}\right)\phi + \frac{\sin 2\phi}{4} - \frac{2(E_1 - E_0)}{\hat{E}_1}\sin\phi\right]$$

$$\cos\phi = \frac{E_1 - E_0}{\hat{E}_1}$$

$$\eta_d = \frac{E_1 - E_1'}{\hat{E}_1}$$

where E_1' is the voltage across R_1 with $\hat{E}_1 = 0$.

thus $$\eta_d = \frac{E_1 - E_0\left(1 - \sqrt{\frac{R_d}{E_0 R_1}}\right)}{\hat{E}_1} = \frac{E_1 - E_0}{\hat{E}_1} + \frac{1}{\hat{E}_1}\sqrt{\frac{E_0 R_d}{R_1}}.$$

If we assume $-E_0 = -1$, $R_d = 5{,}000\Omega$ and $R_1 = 1\ M\Omega$

then $\eta_d \simeq \frac{E_1 - E_0}{\hat{E}_1}$ except when E_1 is comparable with E_0.

Thus $\cos\phi \simeq \eta_d$

and $$I_m \simeq \frac{\hat{E}_1^2}{\pi R_d}\left[\left(\tfrac{1}{2} + \eta_d^2\right)\cos^{-1}\eta_d - \frac{3}{2}\eta_d\sqrt{1 - \eta_d^2}\right] \quad . \quad 8.26$$

$$= \frac{E_1}{R_1}.$$

$$\frac{R_1}{R_d} \simeq \frac{2\pi\left(\eta_d + \frac{E_0}{\hat{E}_1}\right)}{\hat{E}_1\left[\left(1 + 2\eta_d^2\right)\cos^{-1}\eta_d - 3\eta_d\sqrt{1 - \eta_d^2}\right]} \quad . \quad 8.27.$$

It may also be shown that

$$R_E \simeq \frac{R_1\left[\left(1 + 2\eta_d^2\right)\cos^{-1}\eta_d - 3\eta_d\sqrt{1 - \eta_d^2}\right]}{4\left[\eta_d + \frac{E_0}{\hat{E}_1}\right]\left[\left(\frac{\eta_d^2}{3} + \frac{2}{3}\right)\sqrt{1 - \eta_d^2} - 2\eta_d\cos^{-1}\eta_d\right]}. \quad 8.28.$$

Detection efficiency and R_E are plotted as curve 3 in Fig. 8.11*a* and 8.11*b* for $-E_0 = -1$, $R_d = 5{,}000\Omega$, $R_1 = 1\ M\Omega$ for various values of $\frac{\hat{E}_1}{E_0}$. As $\hat{E}_1$ is reduced to zero R_E approaches a limiting value, which is the A.C. resistance of the diode at $E_a = 0$. This limiting value may be found by reference to Fig. 8.10. If AB' is the original $I_a E_a$ characteristic curve of the diode where $OA = E_0$, the D.C. characteristic, which includes the effect of R_1, is represented by a curve such as AG' intersecting the I_a axis at D'. The A.C. characteristic curve is represented by $F'E'$, which is parallel to

AB'. The effective resistance of the diode when $\hat{E}_1 = 0$ is the inverse of the slope of curve $F'E'$ at the point D'.

We have already found in 8.25*a* the value of the current represented by OD'; it is

$$I_0 \simeq \frac{E_0}{R_1}\left[1 - \sqrt{\frac{R_d}{E_0 R_1}}\right].$$

The equation to curve $F'E'$ is

$$I_a = \frac{(E_a + E_3)^2}{R_d}$$

where $-E_3 = OF' = \sqrt{I_0 R_d}$. (Note $I_a = I_0$ when $E_a = 0$.)

$$R_E = \frac{dE_a}{dI_a} = \frac{R_d}{2(E_a + E_3)},$$

at $E_a = 0$

$$R_E = \frac{R_d}{2E_3} = \frac{R_d}{2\sqrt{I_0 R_d}} = \frac{1}{2}\sqrt{\frac{R_d}{I_0}}$$

$$= \frac{1}{2}\sqrt{\frac{R_d R_1}{E_0\left(1 - \sqrt{\frac{R_d}{E_0 R_1}}\right)}}$$

$$\simeq \frac{1}{2}\sqrt{\frac{R_d R_1}{E_0}} \quad . \quad . \quad . \quad . \quad . \quad 8.29$$

when $R_d \ll R_1$.

In our particular example

$$R_E \simeq \frac{1}{2}\sqrt{\frac{5{,}000 \times 10^6}{1}} = 35{,}300\Omega.$$

From curve 3, Fig. 8.11*b*, it may be noted that R_E for the parabolic diode is higher for small values of $\hat{E}_1$, but for large values of $\hat{E}_1$ there is little difference between it and the linear diode. The detection efficiency is less in the parabolic case (curve 3 in Fig. 8 11*a*), except for very small signal voltages, when it is superior to the linear diode (curve 1) because, due to curvature of the characteristic, there is always some detection For the linear diode detection ceases when $E_1 < \frac{E_0 R_d}{R_1 + R_d}$. An added advantage of the square law diode is that there is much less damping of the input tuned circuit for small signal voltages. R_E, in the particular example given, is 35,300 Ω as compared with 5,000Ω for the linear diode.

The parabolic diode with positive start of current will not be detailed here as the results follow very similar lines to those obtained for the linear diode under such conditions. For example, no detection occurs until the signal voltage exceeds the voltage E_0 at which current starts.

8.2.10. Diode Detector Damping and the Preceding R.F. Amplifier Stage.[27, 28, 30] We have seen from preceding sections that a diode detector reflects a resistance load into the input tuned circuit, causing damping with loss of amplification and selectivity. The equivalent load due to the diode for normal detection efficiencies is $\frac{R_1}{2}$, but we must note that this applies only to the carrier frequency. In Section 8.2.3 it is shown that when the carrier is modulated the coupling resistance to the first A.F. amplifier must be taken into account, and the load reflected on to the tuned circuit for the modulated signal is approximately one-half the A.C. load resistance of the diode, i.e., $\frac{1}{2}\left(\frac{R_1R_2}{R_1+R_2}\right)$ if R_2 is tapped across the full resistance R_1. This means that the amplification for the modulation envelope is less than for the carrier, or, in other words, there is a reduction of the modulation ratio from input to output. If the modulation ratio at the grid of the R.F. tetrode valve supplying the detector is M, g_m and R_a, the mutual conductance and slope resistance of the valve, and R_D the dynamic resistance of the anode tuned circuit, the amplification at the carrier voltage is

$$g_m \Big/ \left(\frac{1}{R_a}+\frac{1}{R_D}+\frac{2}{R_1}\right)$$

and for the modulation envelope

$$g_m \Big/ \left[\frac{1}{R_a}+\frac{1}{R_D}+\frac{2(R_1+R_2)}{R_1R_2}\right]$$

so that the modulation ratio at the detector is reduced to

$$\frac{M.\left(\frac{1}{R_a}+\frac{1}{R_D}+\frac{2}{R_1}\right)}{\frac{1}{R_a}+\frac{1}{R_D}+\frac{2(R_1+R_2)}{R_1R_2}}.$$

8.2.11. Effect of the Capacitance in Shunt with the Load Resistance. In the above calculations we have assumed that the reactance of the shunt capacitance C_1 at the carrier frequency is small in comparison with the diode resistance R_d. When this is not true detection efficiency is reduced and the effective resistance

R_E increased. We will consider first the case with no shunt capacitance and then proceed to a development showing the effect of the capacitance.

8.2.12. Detection Efficiency and Effective Resistance for a Linear Diode with no Shunt Capacitance. If $I_a = \frac{E_a}{R_d}$ is the equation to the diode characteristic curve, the equation to the curve representing the circuit conditions including the load resistance R_1 is $I_a = \frac{E_a}{R_1+R_d}$.

The mean current is given by

$$I_m = \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \frac{\hat{E}_1 \cos\theta}{R_1+R_d}\, d\theta,$$

the limits are from $\frac{\pi}{2}$ to 0, because there is no capacitance to maintain the voltage across R_1 and half-wave detection results.

$$I_m = \frac{\hat{E}_1}{\pi(R_d+R_1)} = \frac{E_1}{R_1}$$

thus
$$\frac{E_1}{\hat{E}_1} = \eta_d = \frac{R_1}{\pi(R_d+R_1)} \quad . \quad . \quad . \quad . \quad 8.30.$$

Power absorbed
$$= \frac{1}{\pi}\int_0^{\frac{\pi}{2}} \frac{\hat{E}_1^2 \cos^2\theta}{R_1+R_d}\, d\theta$$

$$= \frac{\hat{E}_1^2 \frac{\pi}{4}}{\pi(R_1+R_d)} = \frac{\hat{E}_1^2}{2R_E}$$

$$R_E = 2(R_1+R_d) \quad . \quad . \quad . \quad . \quad 8.31.$$

Thus when $C_1 = 0$ and $R_d \ll R_1$, detection efficiency approaches $\frac{1}{\pi}$ and R_E approaches $2R_1$. The latter result is to be expected since the diode conducts for exactly half a cycle, and during conduction the load resistance is very nearly R_1.

8.2.13. Effect of Shunt Capacitance on Detection Efficiency. When the reactance of C_1 cannot be neglected, as may be the case for a low radio frequency (110 kc/s) the curves in Fig. 8.4*b* will not give the true performance of the detector. An estimate of its behaviour may be made by obtaining a mean detection voltage-carrier peak voltage curve as described at the end of Section 8.2.2.

We will now turn to a theoretical examination of the effect of varying C_1.

In Fig. 8.7 is shown the shape of the voltage wave across the load resistance R_1. The interval during one cycle of the input wave, which will be assumed unmodulated, is divided into a charge and discharge period. The former starts at some time instant t_1 when the input voltage exceeds that across R_1, and ceases at t_2. If stable conditions prevail, charge will commence again at a time instant $t_3 = \frac{2\pi}{\omega} + t_1$ when the frequency of the input wave is $f = \frac{\omega}{2\pi}$.

Current and Voltage Relationships during Charge.

$$\hat{E}_1 \cos \omega t = E_1' + iR_d \quad . \quad . \quad . \quad 8.32$$

$$E_1' = i_1 R_1 = \frac{\int i_2 \, dt}{C_1} \quad . \quad . \quad . \quad 8.33$$

$$i = i_1 + i_2 \quad . \quad . \quad . \quad 8.34$$

where i, i_1 and i_2 are the instantaneous currents through diode, R_1 and C_1 respectively (Fig. 8.2) and E_1' is the instantaneous voltage across R_1 and C_1 (Fig. 8.7).

From 8.33 and 8.34

$$i = \frac{E_1'}{R_1} + C_1 \frac{dE_1'}{dt}.$$

Thus

$$\frac{\hat{E}_1 \cos \omega t}{C_1 R_d} = \frac{E_1'}{C_1 R_d}\left[1 + \frac{R_d}{R_1}\right] + \frac{dE_1'}{dt} \quad . \quad . \quad 8.35.$$

The particular solution is

$$E_{1p}' = K\hat{E}_1 \cos (\omega t - \phi)$$

where

$$\phi = \tan^{-1} \frac{\omega C_1}{\frac{1}{R_1} + \frac{1}{R_d}}$$

and

$$K = \frac{1}{R_d \sqrt{\omega^2 C_1^2 + \left(\frac{1}{R_1} + \frac{1}{R_d}\right)^2}} = \frac{\cos \phi}{\left(1 + \frac{R_d}{R_1}\right)}$$

$$\therefore \; E_{1p}' = \frac{E_1' \cos \phi}{\left[1 + \frac{R_d}{R_1}\right]} \cos (\omega t - \phi).$$

The complementary solution is

$$E_{1c}' = K_1 \varepsilon^{-\left(\frac{1}{R_d} + \frac{1}{R_1}\right)\frac{t}{C_1}} = K_1 \varepsilon^{-\frac{\omega t}{\tan \phi}}.$$

The complete solution is

$$E_1' = \frac{\hat{E}_1 \cos\phi}{\left[1+\frac{R_d}{R_1}\right]}\left[\cos(\omega t-\phi)+K_2\varepsilon^{-\frac{\omega t}{\tan\phi}}\right]. \quad . \quad 8.36a.$$

If diode conduction begins at some time instant t_1, the value of K_2 may be found by replacing E_1' by $\hat{E}_1 \cos \omega t_1$ in 8.36*a*

thus $$K_2 = \left\{\frac{\left(1+\frac{R_d}{R_1}\right)}{\cos\phi}\cos\omega t_1 - \cos(\omega t_1-\phi)\right\}\varepsilon^{\frac{\omega t_1}{\tan\phi}}.$$

Hence $$E_1' = \frac{\hat{E}_1\cos\phi}{\left(1+\frac{R_d}{R_1}\right)}\left\{\cos(\omega t-\phi)\right.$$
$$\left.+\left[\frac{\left(1+\frac{R_d}{R_1}\right)}{\cos\phi}\cos\omega t_1-\cos(\omega t_1-\phi)\right]\varepsilon^{\frac{-\omega(t-t_1)}{\tan\phi}}\right\}. \quad 8.36b.$$

Current and Voltage Relationships During Discharge.

$$E_1' = \frac{-\int i_2\,dt}{C_1} = R_1 i_1$$

$$i_1 = i_2$$

$$\frac{E_1'}{R_1} + C_1\frac{dE'_1}{dt} = 0$$

$$E_1'\left[D+\frac{1}{R_1C_1}\right] = 0$$

where $$D = \frac{d}{dt}.$$

The solution to this is

$$E_1' = K_3\varepsilon^{\frac{-t}{R_1C_1}}.$$

If diode conduction ceases at a time instant t_2, then replacing E_1' by $\hat{E}_1 \cos \omega t_2$ gives

$$K_3 = \hat{E}_1\cos\omega t_2\varepsilon^{\frac{t_2}{R_1C_1}}$$

$$E_1' = \hat{E}_1\cos\omega t_2\varepsilon^{-\left(\frac{t-t_2}{R_1C_1}\right)} \quad . \quad . \quad . \quad 8.37.$$

Steady State Conditions.

In the steady state condition the voltage at the end of discharge is equal to that at the beginning of charge, and vice versa.

By replacing t in equation 8.36*b* by t_2, in equation 8.37 by $\frac{2\pi}{\omega}+t_1$ and noting from Fig. 8.7 that E_1' is $\hat{E}_1 \cos \omega t_2$ and $\hat{E}_1 \cos(\omega t_1+2\pi)$ respectively, we obtain two simultaneous equations involving t_2 and t_1,

$$\text{thus } \cos \omega t_2 = \frac{\cos \phi}{\left(1+\frac{R_d}{R_1}\right)}\left\{\cos(\omega t_2 - \phi) + \left[\frac{\left(1+\frac{R_d}{R_1}\right)}{\cos \phi}\cos \omega t_1 - \cos(\omega t_1 - \phi)\right]\varepsilon^{\frac{-\omega(t_2-t_1)}{tan\,\phi}}\right\}$$

$$\text{and} \qquad \cos(\omega t_1+2\pi) = \cos \omega t_2 \varepsilon^{\frac{-\left(t_1+\frac{2\pi}{\omega}-t_2\right)}{R_1C_1}}$$

It is easier to deal with angles than time so we will replace ωt_2 by α_2 and ωt_1 by α_1, and rearrange these two equations into the following forms

$$\left\{\frac{\left(1+\frac{R_d}{R_1}\right)}{\cos \phi}\cos \alpha_1 - \cos(\alpha_1 - \phi)\right\}\varepsilon^{\frac{\alpha_1}{tan\,\phi}} = \left\{\frac{\left(1+\frac{R_d}{R_1}\right)}{\cos \phi}\cos \alpha_2 - \cos(\alpha_2 - \phi)\right\}\varepsilon^{\frac{\alpha_2}{tan\,\phi}} \quad . \quad 8.38a$$

$$\cos \alpha_1 \varepsilon^{\frac{2\pi+\alpha_1}{\omega C_1 R_1}} = \cos \alpha_2 \varepsilon^{\frac{\alpha_2}{\omega C_1 R_1}} \quad . \quad . \quad . \quad 8.38b.$$

The direct solution of 8.38*a* and 8.38*b* is not possible and the best method (suggested by Marique from whose paper [24] this analysis is taken) is by plotting the function

$$y = \left\{\left[\frac{1+\frac{R_d}{R_1}}{\cos \phi}\right]\cos \alpha - \cos(\alpha - \phi)\right\}\varepsilon^{\frac{\alpha}{tan\,\phi}}$$

for $$\alpha \text{ from } -\frac{\pi}{2} \text{ to } \frac{\pi}{2}$$

and the function

$$z = \cos \alpha \; \varepsilon^{\frac{2\pi+\alpha}{\omega R_1 C_1}} \text{ for } \alpha \text{ from } -\frac{\pi}{2} \text{ to } 0$$

and $$z = \cos \alpha \; \varepsilon^{\frac{\alpha}{\omega R_1 C_1}} \text{ for } \alpha \text{ from } 0 \text{ to } \frac{\pi}{2}.$$

Representative curves of y and z are shown in Fig. 8.12 between $-\frac{\pi}{2}$ and $\frac{\pi}{2}$; y has a positive value of $\sin\phi.\varepsilon^{-\frac{\pi}{2\tan\phi}}$ at $-\frac{\pi}{2}$ and a negative value of $-\sin\phi.\varepsilon^{\frac{\pi}{2\tan\phi}}$ at $+\frac{\pi}{2}$ whilst z is zero at both these angular values. The two values of α satisfying equations 8.38*a* and 8.38*b* for a particular value of y or z can be obtained by drawing lines parallel to the " α " axis and reading on that axis the intercepts with the y and z curves. The intercept between $-\frac{\pi}{2}$ and 0 gives α_1 and that between 0 and $\frac{\pi}{2}$, gives α_2. Two curves may be plotted of α_1 against α_2 (one for y, and one for z) obtained from the representative curves, and the point of intersection of these two curves

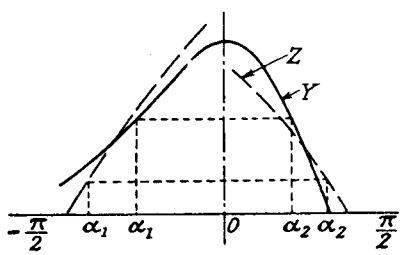


FIG. 8.12.—Typical Curves for the Charge and Discharge Functions.

for y and z gives the value of α_1 and α_2 satisfying the functions y and z simultaneously. Using these values of α_1 and α_2 and assuming the mean rectified voltage E_1 for the discharge period to be the same as for the charge period, we may calculate from 8.37 the value of this mean voltage. This assumption removes the necessity for calculating E_1 over the charge period by means of the more complicated expression 8.36*b*.

$$
\begin{aligned}
E_1 &= \frac{1}{2\pi+\alpha_1-\alpha_2}\int_{\alpha_2}^{\alpha_1+2\pi} E_1' d\alpha \\
&= \frac{\hat{E}_1\cos\alpha_2}{2\pi+\alpha_1-\alpha_2}\int_{\alpha_2}^{\alpha_1+2\pi} \varepsilon^{\frac{-(\alpha-\alpha_2)}{\omega C_1 R_1}}\, d\alpha \\
&= \frac{\omega C_1 R_1 \hat{E}_1\cos\alpha_2}{(2\pi+\alpha_1-\alpha_2)}\left[-\varepsilon^{\frac{-(\alpha-\alpha_2)}{\omega C_1 R_1}}\right]_{\alpha_2}^{\alpha_1+2\pi} \\
&= \frac{\omega C_1 R_1 \hat{E}_1\cos\alpha_2}{2\pi+\alpha_1-\alpha_2}\left(1-\varepsilon^{\frac{-(2\pi+\alpha_1-\alpha_2)}{\omega C_1 R_1}}\right)
\end{aligned}
$$

but detection efficiency $\eta_d = \dfrac{E_1}{\hat{E}_1}$

$$\therefore \eta_d = \frac{E_1}{\hat{E}_1} = \frac{\omega C_1 R_1 \cos \alpha_2}{2\pi + \alpha_1 - \alpha_2}\left[1 - \varepsilon^{\frac{-(2\pi+\alpha_1-\alpha_2)}{\omega C_1 R_1}}\right]$$

$$= \frac{R_1}{X_{C_1}} \cdot \frac{\cos \alpha_2}{2\pi + \alpha_1 - \alpha_2}\left[1 - \varepsilon^{\frac{-(2\pi+\alpha_1-\alpha_2)}{\omega C_1 R_1}}\right]. \quad . \quad 8.39$$

where X_{C_1} is the R.F. reactance of C_1.

The relationship between η_d and $\dfrac{R_1}{X_{C_1}}$ is shown in Fig. 8.13*a* for different values of $\dfrac{R_1}{R_d}$ and it will be noted that in each case there is a value of $\dfrac{R_1}{X_{C_1}}$ which, if exceeded, gives little improvement in

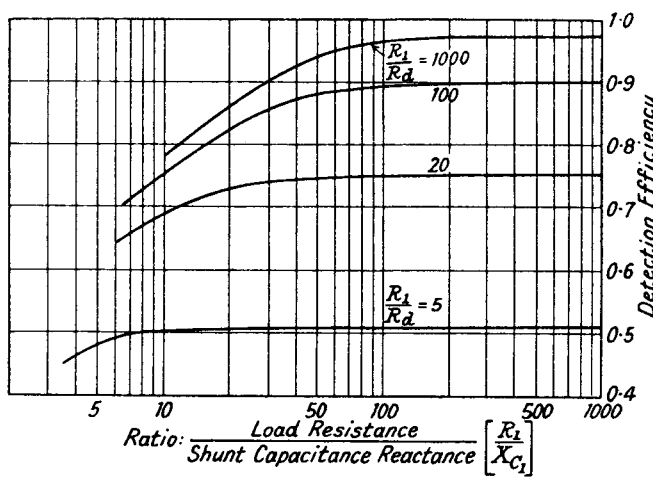


FIG. 8.13*a*.—Curves of Detection Efficiency against the Ratio of Load Resistance to Shunt Capacitance Reactance for Different Values of Load to Conduction Resistance.

detection efficiency. The curves are asymptotic to a value of η_d, which is that given in Fig. 8.8 for the particular value of $\dfrac{R_1}{R_d}$. The limiting value of $\dfrac{R_1}{X_{C_1}}$ is plotted in Fig. 8.13*b* against $\dfrac{R_1}{R_d}$, and if we take our previous case of $R_1 = 1$ MΩ and $R_d = 5{,}000\Omega$, a value

of $\frac{R_1}{X_{C_1}} = 80$ is obtained. The shunt capacitance has therefore little effect on η_d provided the following relationship is satisfied:

$$X_{C_1} \gtrsim \frac{R_1}{80}$$

or

$$C_1 \lesssim \frac{80 \times 10^6}{R_1 \times 2\pi f}\,\mu\text{F}$$

$$\lesssim \frac{12{\cdot}72 \times 10^6}{fR_1}\,\mu\text{F}$$

$$\lesssim \frac{12{\cdot}72}{f(\text{Mc/s}).R_1(\text{M}\Omega)}\,\mu\mu\text{F}.$$

If $f = 1{,}000$ kc/s, and $R_1 = 1$ MΩ, $C_1 \lesssim 12{\cdot}72$ $\mu\mu$F, whilst for $f = 100$ kc/s, $C_1 \lesssim 127{\cdot}2$ $\mu\mu$F.

In any practical case the value of C_1 is unlikely to be less than

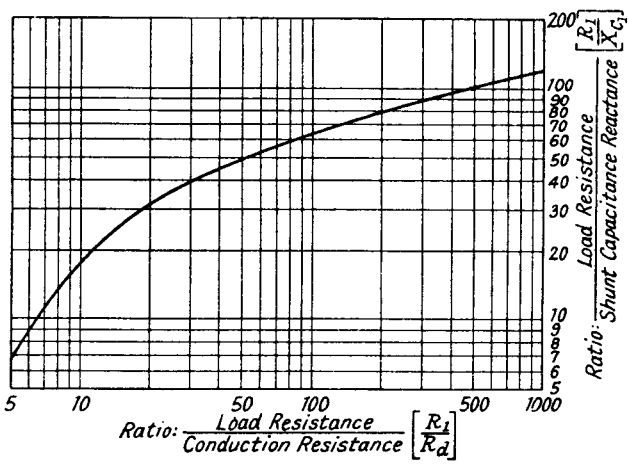


FIG. 8.13*b*.—A Curve of Load Resistance to Shunt Capacitance Reactance Ratio against Load Resistance to Conduction Resistance Ratio for Maximum Detection Efficiency.

0·0001 μF, so that we were justified in ignoring the effect of the shunt capacitance when calculating the equivalent damping resistance of the diode.

The above calculation suggests a minimum value for the capacitance C_1, and we must next consider its limiting value since too large a value can produce amplitude distortion (due to non-tracking) of the detected audio frequency wave shape.

8.2.14. Amplitude Distortion due to a Large Value of

Shunt Capacitance. Amplitude distortion [8, 17] has already been indicated in Fig. 8.3 and it is shown to be due to failure of the voltage across the capacitance to fall to a value below the succeeding carrier peak voltage. This effect only occurs as the modulation envelope is decreasing, and is worst at the point where the rate of change of the modulation envelope is greatest. Thus the rate of discharge of C_1R_1 must not be less than the maximum rate of change of the modulation envelope. If we assume a modulated input signal of $E_1(1+M\cos pt)\cos \omega t$, the rate of change of the modulation envelope is

$$\frac{dE}{dt} = -E_1Mp\sin pt \quad . \quad . \quad . \quad 8.40$$

this is a maximum when $\frac{d^2E}{dt^2} = 0$

i.e., $$-\hat{E}_1Mp^2\cos pt = 0$$

this gives $\cos pt = 0$, and $t = \frac{\pi}{2p}$ or $\frac{\pi}{p}(n+\frac{1}{2})$.

The solutions $t = \frac{3\pi}{2p}, \frac{7\pi}{2p}$, etc., are inadmissible because the modulation envelope is rising, so that the required solutions are

$$t = \frac{\pi}{2p}, \frac{5\pi}{2p}, \text{ etc.}$$

The maximum rate of change of the modulation envelope is $-\hat{E}_1Mp$.

The discharge equation for C_1R_1 is

$$E_{C_1} = E_0\varepsilon^{-\frac{t}{R_1C_1}}$$

the rate of discharge $$\frac{dE_0}{dt} = -\frac{E_0}{R_1C_1}\varepsilon^{-\frac{t}{R_1C_1}} = -\frac{E_{C_1}}{R_1C_1} \quad . \quad 8.41.$$

At the time $t = \frac{\pi}{2p}$; $\cos pt = 0$, $E_{C_1} = E_1$.

Hence $$-\frac{E_1}{R_1C_1} \geqq -\hat{E}_1Mp \quad . \quad . \quad . \quad 8.42a$$

and noting that $X_{C_1} = \frac{1}{pC_1}$, the reactance of C_1 at audio frequencies, expression 8.42*a* becomes [15]

$$\frac{X_{C_1}}{R_1} \geqq M \quad . \quad . \quad . \quad . \quad 8.42b.$$

Thus if amplitude distortion is to be avoided the ratio of the

reactance of C_1 at audio frequencies to the resistance R_1 must not be less than M, the modulation ratio.

Taking 5,000 c.p.s. as the maximum audio frequency that needs consideration (harmonics of frequencies higher than this are above 10,000 c.p.s. and are therefore unlikely to be troublesome), we have for $R_1 = 1$ MΩ and $M = 1$, a value of

$$C_1 = \frac{10^{12}}{2\pi \times 5{,}000 \times R_1} \mu\mu\text{F}$$
$$= 31{\cdot}9 \ \mu\mu\text{F}$$

which must not be exceeded if distortion is to be prevented at full modulation.

The usual practical value of C_1 is 0·0001 μF, so that the maximum acceptable modulation ratio is limited to 0·319 at 5,000 c.p.s. if distortion is to be avoided. The transmitted modulation ratio for the high audio frequency components is generally low and is normally much less than 0·319.

It is interesting to note that for a given ratio $\frac{X_{C_1}}{R_1}$ the maximum modulation ratio may be increased by applying positive bias [19] to the anode of the diode.

In the absence of a signal voltage there is a voltage across C_1 of very nearly $E_b (R_1 \gg R_d)$, and if we assume $\eta_d = 1$ the voltage rises to $\hat{E}_1 + E_b$ when a signal $\hat{E}_1 \cos \omega t$ is applied. The rate of discharge of C_1 thus becomes

$$\frac{dE_{C1}}{dt} = \frac{-(\hat{E}_1 + E_b)}{R_1 C_1}.$$

The conditions for no amplitude distortion are

$$\frac{\hat{E}_1 + E_b}{\hat{E}_1} \cdot \frac{X_{C_1}}{R_1} \geq M. \quad . \quad . \quad . \quad 8.42c.$$

8.2.15. Frequency Distortion due to the Shunt Capacitance. The capacitance C_1 can produce frequency distortion in addition to amplitude distortion if its value is increased unduly.

We have already noted in expression 8.8 that the mean current through the diode for an unmodulated carrier $\hat{E}_1 \cos \omega t$ is

$$I_m = \frac{\hat{E}_1}{\pi R_d}\left\{\sin \phi - \frac{E_1}{\hat{E}_1}\phi\right\}$$

but since $E_1 = I_m R_1$

$$I_m = \frac{1}{\pi R_d}[\hat{E}_1 \sin \phi - I_m R_1 \phi]$$

or

$$I_m = \frac{\hat{E}_1 \sin \phi}{\pi R_d + R_1 \phi} \quad . \quad . \quad . \quad . \quad 8.43.$$

If a modulated carrier $\hat{E}_1(1+M\cos pt)\cos\omega t$ is applied to the diode, substitution of $\hat{E}_1(1+M\cos pt)$ in 8.43 leads to

$$I_m = \frac{\hat{E}_1(1+M\cos pt)\sin\phi}{\pi R_d + R_1\phi}.$$

Since, however, the load resistance R_1 is shunted by the capacitance C_1, the expression for I_m is more correctly given by

$$I_m = \frac{\hat{E}_1\sin\phi}{\pi R_d + R_1\phi} + \frac{\hat{E}_1\sin\phi}{\pi R_d + Z_1\phi} M\cos pt.$$

The first part is the D.C. term produced by the carrier and the second is the A.C. term due to the modulation frequency. In this term R_1 is replaced by the impedance Z_1, which is the parallel impedance of C_1 and R_1 at the modulation frequency $\left(\frac{R_1}{jpC_1R_1+1}\right)$. The audio frequency voltage across the load resistance is

$$\begin{aligned} E_{AF} = I_m Z_1 &= \frac{E_1\sin\phi}{\pi R_d + Z_1\phi} M\cos pt . Z_1 \\ &= \frac{\hat{E}_1\sin\phi . M\cos pt}{\phi} \cdot \frac{Z_1}{\frac{\pi R_d}{\phi} + Z_1} \quad . \quad . \quad 8.44. \end{aligned}$$

The above expression 8.44 can be represented by a generator which has an open circuit voltage of $\frac{\hat{E}_1 M\sin\phi}{\phi}\cos pt$, an internal resistance $\frac{\pi R_d}{\phi}$ and an external load impedance Z_1. The internal resistance $\frac{\pi R_d}{\phi}$ is the equivalent resistance of the diode at audio frequencies. It is R_d' of Section 8.2.3 and the inverse slope of the $I_m E_a$ characteristic curves in Fig. 8.4*b*.

Two points of interest arise : (1) the resistance R_d' is dependent on detection efficiency for from Section 8.2.5. ϕ equals $\cos^{-1}\eta_d$. Expression 8.10 shows that η_d is controlled by R_d and R_1, so that R_d' itself must be dependent on R_d and R_1. The actual relationship is discussed in the next section ; (2) the reciprocal of the resistance R_d' may be obtained by partial differentiation of 8.8 with respect to E_1, thus

$$\frac{dI_m}{dE_1} = \frac{\phi}{\pi R_d} = \frac{1}{R_d'}.$$

Since the resistance R_d' is dependent on R_1, increasing as R_1 increases, the latter has a twofold effect on the frequency response

curve of the audio frequency output voltage from the detector. Increase of R_1 magnifies the influence of the capacitance C_1 and also increases R_d', and both effects tend to produce loss of high audio frequencies.

If $R_d = 5{,}000\Omega$, $R_1 = 1$ MΩ, $C_1 = 0{\cdot}0001$ μF, Fig. 8.8 gives $\eta_d = 0{\cdot}93$, and the audio output voltage is

$$E_{AF} = \frac{KZ_1}{R_d'+Z_1} = \frac{KR_1}{R_d'(1+jpC_1R_1)+R_1}$$
$$= \frac{KR_1}{\sqrt{(R_d'+R_1)^2+(pC_1R_1R_d')^2}}.$$

But
$$R_d' = \frac{R_d\pi}{\cos^{-1}\eta_d} = \frac{5{,}000\pi}{0{\cdot}3763}$$
$$= 41{,}700\Omega.$$

The reduction in voltage at 10,000 c.p.s. as compared with 50 c.p.s. is

$$\frac{E_{10,000}}{E_{50}} \simeq \frac{R_1}{\sqrt{(R_d'+R_1)^2+(pC_1R_1R_d')^2}} \cdot \frac{R_d'+R_1}{R_1}$$
$$\simeq \frac{1}{\sqrt{1+\left(\frac{pC_1R_1R_d'}{R_d'+R_1}\right)^2}} \simeq \frac{1}{\sqrt{1+(0{\cdot}2525)^2}}$$
$$= 0{\cdot}969 \text{ or } -0{\cdot}28 \text{ db.}$$

The loss of high-frequency response for the component values chosen is negligible. Should, however, the value of C_1 or the conduction resistance R_d of the diode be increased, the loss may become important. For example, suppose $C_1 = 0{\cdot}0005$ μF, then

$$\frac{E_{10,000}}{E_{50}} \simeq \frac{1}{\sqrt{1+(5 \times 0{\cdot}2525)^2}}$$
$$= \frac{1}{\sqrt{1+1{\cdot}59}} = 0{\cdot}621$$
$$\equiv -4.12 \text{ dbs.}$$

or if $R_d = 25{,}000\Omega$, $R_1 = 1$ MΩ and $C_1 = 0{\cdot}0001$ μF. Fig. 8.8 gives $\eta_d = 0{\cdot}84$

$$R_d' = \frac{25{,}000\pi}{0{\cdot}574} = 136{,}800\Omega$$

$$\frac{E_{10,000}}{E_{50}} = \frac{1}{\sqrt{1+(0{\cdot}756)^2}} = 0{\cdot}797 \equiv -1{\cdot}96 \text{ dbs.}$$

Increase of capacitance C_1 has therefore a greater effect than the same ratio increase in R_d. Decrease of R_1 reduces the high-frequency loss for given values of C_1 and R_d.

8.2.16. The I_mE_a Characteristic Curves[9] for a Linear Diode Conducting at $E_a = 0$. The mean current expression 8.8 may be written

$$I_m = \frac{\hat{E}_1}{\pi R_d}[\sin\phi - \phi\cos\phi] \quad . \quad . \quad . \quad 8.45$$

for $$\frac{E_1}{\hat{E}_1} = \cos\phi.$$

$$\therefore\ I_m = \frac{\hat{E}_1}{\pi R_d} f(\phi).$$

The function $f(\phi)$ is plotted in Fig. 8.14*a* between $\phi = 0$ and 4 radians. To obtain a particular I_mE_a characteristic curve it is

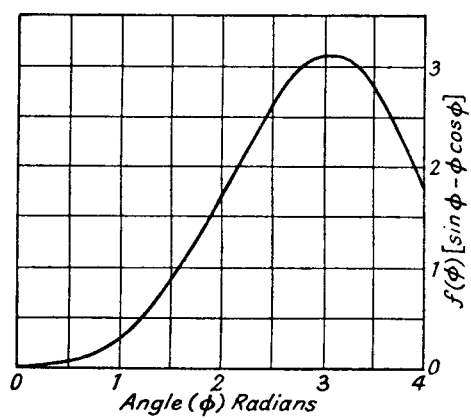


FIG. 8.14*a*.—The Curve of $f(\phi)$ against ϕ.

only necessary to replace E_1 in 8.45 by the required peak value and to note that $\frac{E_1}{\hat{E}_1} = \frac{-E_a}{\hat{E}_1} = \cos\phi$. For example, if $R_d = 5{,}000\Omega$, $E_1 = 1$ volt and $E_a = -0{\cdot}5$ volts, $\cos\phi = 0{\cdot}5$ or $\phi = \frac{\pi}{3}$, which gives a value of $f(\phi) = 0{\cdot}35$ and $I_m = 0{\cdot}022$ mA. If this is repeated for values of E_a from -1 to $+1$ the I_mE_a curve for $\hat{E}_1 = 1$ volt is obtained. When E_a exceeds $+\hat{E}_1$ each curve merges into the diode conduction curve because the diode current is not taken to cut-off and no detection occurs. In Fig. 8.14*b* I_mE_a curves are shown for $\hat{E}_1 = 1$ to 5 volts and $R_d = 5{,}000$ ohms. If a load line OB, corresponding to any particular value of R_1, is drawn across these

curves, it will be noted that I_m at the intersections is proportional to $\hat{E}_1$. Since $I_m = \frac{\hat{E}_1}{\pi R_d} f(\phi)$ it follows that $f(\phi)$ must be constant,

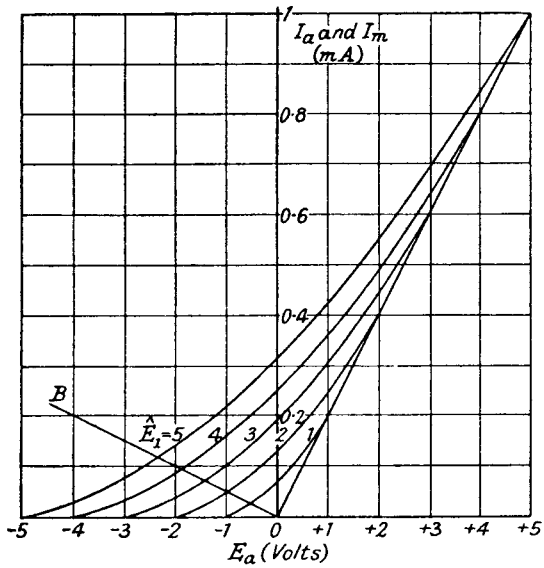


FIG. 8.14b.—Characteristic $I_m E_a$ Curves for a Diode with Linear Conduction Curve.
(R_d = 5,000Ω.)

i.e., ϕ is constant. But we have already shown in Section 8.2.15 that the slope of these curves is $\frac{1}{R_d'}$ or $\frac{\phi}{\pi R_d}$. Hence if ϕ is constant, R_d' is also constant for a given value of R_1. The relationship between R_d' and R_1 may be obtained for a fixed value of R_d by

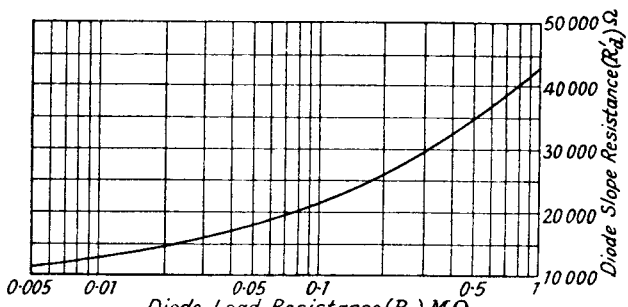


FIG. 8.14c.—A Curve of Diode Slope Resistance against Load Resistance.
(R_d = 5,000Ω.)

substituting the particular values of ϕ (obtained from the curves of $\eta_d(\phi = \cos^{-1}\eta_d)$ against $\frac{R_1}{R_d}$, Fig. 8.8), in the expression $R_d' = \frac{\pi R_d}{\phi}$, e.g., Fig. 8.8 gives $\eta_d = 0{\cdot}866$ at $\frac{R_1}{R_d} = 58$; $\phi = \cos^{-1} 0{\cdot}866 = \frac{\pi}{6}$, and replacing this in $R_d' = \frac{\pi R_d}{\phi}$ when $R_d = 5{,}000\Omega$, we find that $R_d' = 6R_d = 30{,}000\Omega$ for $R_1 = 58 \times 5{,}000 = 290{,}000\Omega$. A curve of R_d' against R_1 for $R_d = 5{,}000\Omega$ is plotted in Fig. 8.14c.

8.3. Cumulative Grid Detection.

8.3.1. Introduction. The cumulative grid detector operates in a manner similar to the diode by using the grid of a triode or multigrid valve as the diode anode. The I_gE_g characteristic curve is generally similar to curve AB' of Fig. 8.10 illustrating the parabolic detector. The audio frequency variations of small carrier amplitudes are detected due to curvature [5] of the characteristic, i.e.,

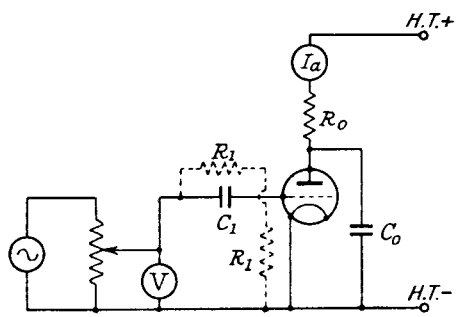


FIG. 8.15.—A Circuit for Measuring the Detection Characteristics of a Cumulative Grid Detector.

I_g [14] flows throughout the cycle and R_E is low, but for large carrier amplitudes the grid-cathode space acts as a unidirectional device and I_g flows over a small part of the cycle only.

The audio frequency variations across R_1 (Fig. 8.15) cause variations in the grid bias of the valve which in turn produce audio frequency variations of anode current. The cumulative grid detector may be viewed as a separate diode,[18] the anode of which is connected directly to the grid of an A.F. amplifier valve. Detection makes the grid voltage negative with respect to the cathode so that the D.C. anode current falls as the carrier input voltage is increased. This type of detector would therefore appear to offer the advantage of a combined detector and A.F. amplifier in one valve. It possesses,

however, two disadvantages, the D.C. grid-bias point is variable, and R.F. as well as A.F. voltages are applied to the grid. The bias point is fixed by the input carrier peak voltage, so that it is impossible to operate the valve under optimum audio frequency amplifying conditions except for a restricted range of carrier voltages. The presence of R.F. as well as A.F. voltages increases the possibility of overloading, and the A.F. output voltage is consequently restricted. Since R.F. voltages from the input tuned circuit are not filtered from the grid of the valve, the anode-grid capacitance will cause the anode circuit to reflect back a resistance and capacitance component into the grid circuit as described in Section 2.8.2. The resistance component is generally the more serious and its effect may be many times greater than that due to grid current damping.

There are two possible methods of connecting the detector load resistance R_1 as illustrated in Fig. 8.15. The only difference between the two positions so far as the operating conditions are concerned is that when R_1 is connected between grid and cathode it acts as an additional shunt resistance across the input circuit. This connection of R_1 is chiefly of advantage in battery valve detectors, having start of grid current at some positive voltage, for R_1 can be returned to the L.T. positive terminal without disturbing the valve or input circuit connections. The necessity for a positive bias voltage on the grid of the valve having a positive start of current has already been discussed in Section 8.2.7. If large signal voltages are applied, grid current is only taken on the peaks of the input signal, because the voltage across R_1 is maintained by the discharge of capacitance C_1 between the peaks in a manner exactly similar to that shown in Fig. 8.3 for the diode. This produces less damping on the grid circuit $\left(\text{see curve 1 in Fig. 8.11}b \text{ for } \frac{\hat{E}_1}{E_0} \text{ increasing}\right)$, a higher detection efficiency (curve 1 in Fig. 8.11a) and a mean grid voltage variation which is an almost exact reproduction of the modulation envelope.

Owing to the control of the bias by the input carrier voltage, though no distortion may occur in the grid circuit, distortion may be produced in the anode circuit. When the mean grid voltage variations are large the anode current variations may be carried into the bottom bend of the I_aE_g characteristic curve with resultant flattening [2] of the audio frequency wave shape. In extreme cases this may cause in a receiver a double-humped tuning effect with a distorted minimum output at the correct tuning point. This flattening of the audio frequency wave may be decreased by

increasing the detector anode voltage and reducing the mean grid voltage variations. A detector modified in this way is often called a power grid detector.

8.3.2. Power Grid Detection. The power grid detector—a modification of the cumulative grid detector—is intended for operation with large input carrier voltages, when the grid circuit acts as a unidirectional current device and current is taken only on extreme positive peaks of the signal. The grid leak resistance R_1 (Fig. 8.15) is reduced and the anode voltage increased so as to prevent distortion of the audio frequency output current in the anode circuit. The reduction of grid voltage variations by decreasing R_1 may be illustrated from Fig. 8.4*b* where it is shown that the line OB' (corresponding to a small R_1) reduces the grid voltage variations in the ratio $\frac{E_1' - E_0'}{\hat{E}_1}$.

8.3.3. Damping of the Input Circuit by a Cumulative Grid Detector. The cumulative grid detector introduces two forms of damping on the input signal circuit. The first, due to grid conduction current, has already been adequately discussed in Sections 8.2.5 to 8.2.9 on the diode. The second is due to the anode-grid capacitance of the valve. We have already developed in Section 2.8.2 expressions for the conductance and susceptance reflected into the grid circuit by feedback from the anode circuit through the anode-grid capacitance of the valve. Taking expressions 2.9*a* and 2.9*b*, substituting $\frac{1}{R_a}$ for G_a, $\frac{1}{R_0}$ for G_0 and ωC_0 for B_0, and neglecting B_{g_a} in association with other components we have

$$R_g = \frac{\left(\frac{1}{R_a}+\frac{1}{R_0}\right)^2+(\omega C_0)^2}{g_m \omega C_{g_a} \omega C_0} \quad . \quad . \quad . \quad . \quad . \quad 8.46a$$

and

$$C_g = \frac{C_{g_a}\left[\left(\frac{1}{R_a}+\frac{1}{R_0}+g_m\right)\left(\frac{1}{R_a}+\frac{1}{R_0}\right)+(\omega C_0)^2\right]}{\left(\frac{1}{R_a}+\frac{1}{R_0}\right)^2+(\omega C_0)^2} \quad . \quad 8.46b$$

where R_g and C_g are the equivalent parallel resistance and capacitance components of the grid input admittance. The load capacitance C_0 is an essential part of the circuit as it acts as a bypass for R.F. currents developed in the valve, and reduces the possibility of passing on R.F. voltages to the grid of the next valve. Its maximum value is limited by its discrimination against the high-frequency

components of the audio frequencies contained in the modulation; a suitable value is 100 $\mu\mu$F.

When $\omega C_0 \gg \frac{1}{R_a}+\frac{1}{R_0}$ and $g_m \gg \frac{1}{R_a}+\frac{1}{R_0}$.

$$R_g \simeq \frac{C_0}{g_m C_{g_a}} \quad . \quad . \quad . \quad . \quad . \quad . \quad 8.47a$$

and

$$C_g \simeq C_{g_a}\left(\frac{g_m\left(\frac{1}{R_a}+\frac{1}{R_0}\right)}{(\omega C_0)^2}+1\right) \quad . \quad . \quad 8.47b.$$

By using these expressions for R_g and C_g we can estimate the effect of g_m, C_{g_a}, R_0 and C_0 on the input admittance, and this is summarized in the table below.

Variable.	R_g	C_g
Increasing g_m	decrease	increase
Increasing C_{g_a}	decrease	increase
Increasing R_0		
(*a*) $R_0 \gg \frac{1}{\omega C_0}$	(*a*) little effect	decrease
(*b*) R_0 comparable to $\frac{1}{\omega C_0}$	(*b*) has minimum value when $\omega C_0 = \frac{1}{R_a}+\frac{1}{R_0}$	
(*c*) $R_0 \ll \frac{1}{\omega C_0}$	(*c*) decrease	
Increasing C_0	decrease to minimum when $\omega C_0 = \frac{1}{R_a}+\frac{1}{R_0}$	decrease

To show the order of R_g and C_g let us take the following example of a triode detector:

$$f = 1{,}000 \text{ kc/s}, \; R_0 = 30{,}000\Omega, \; C_0 = 100 \; \mu\mu\text{F}, \; C_{g_a} = 3 \; \mu\mu\text{F}$$

$$R_a = 10{,}000\Omega, \; g_m = 3 \text{ mA/volt}.$$

From 8.46*a* and 8.46*b*

$$R_g = 11{,}600\Omega$$
$$C_g = 5{\cdot}9 \; \mu\mu\text{F}.$$

The detuning effect of the parallel capacitance C_g is small and except for small values of tuning capacitance may be ignored. The parallel resistance R_g is so low that it shows anode circuit damping to be much more serious than grid current damping $\left(\frac{R_1}{2} \text{ approximately}\right)$. The resistance R_g can be conveniently increased only by reducing C_{g_a} or g_m. It is not possible to increase

C_0 unduly or loss of the high audio frequency modulation components results.

A reduction of g_m is not a satisfactory solution since this reduces the gain of the detector stage. (For large signal input voltages g_m is reduced because of increase of D.C. bias voltage, so that there is a tendency to decrease anode circuit damping as the input signal is increased.) The only real solution is a reduction of C_{g_a} and this can be accomplished by using a tetrode or pentode [21] valve as a detector. Care must be exercised to see that the external stray capacitance between grid and anode is very small if the full advantage of such a valve is to be gained, and it should be noted that the high internal resistance of a multigrid valve will tend to accentuate the loss of high audio frequencies due to the bypass capacitance C_0. The stage gain of a triode detector is proportional to $\frac{\mu Z_0}{R_a+Z_0}$, i.e., $\frac{g_m R_a Z_0}{R_a+Z_0}$, whilst for a tetrode it is more nearly proportional to $g_m Z_0$ or $\frac{\mu Z_0}{R_a}$; any variation of Z_0 will obviously affect the result to a much less extent in the triode case.

Anode circuit damping can be entirely eliminated when a separate diode and triode valve are used for detection and amplification. An R.F. filter (R_3C_3 in Fig. 8.5), included between the diode load resistance and the triode grid circuit, effectively isolates the latter from the tuned R.F. input circuit to the detector. The diode-triode combination usually produces a higher gain than that of the same triode operating as a cumulative grid detector because of the elimination of anode circuit damping. Furthermore, greater output voltage is possible for a given distortion since only A.F. voltages are applied to the triode grid circuit.

8.3.4. Estimation of the Performance of the Cumulative Grid Detector. The detection characteristic of the cumulative grid detector may be measured at a low frequency (50 c.p.s.) provided the capacitances C_1 and C_0 are increased to give the same reactances at 50 c.p.s. as at the normal carrier frequency. For example, 0·0001 μF at 1,000 kc/s is equivalent to 2 μF at 50 c.p.s. The circuit is given in Fig. 8.15 ; no special precautions are necessary except that since grid current flows the input potentiometer must not have a high resistance value. A series of curves of mean detection voltage plotted against peak input volts may be obtained as in Fig. 8.16. The mean detection voltage is given by $(I_{a0} - I_{a1})R_0$, where I_{a1} is the D.C. current at a particular peak input voltage,

I_{a0} is the D.C. current at zero input voltage, and R_0 is the resistance in the anode circuit. The initial curvature for small peak input voltages (less than 0·1 volt) due to square law detection is indicated in the curves.

The audio output voltage in the anode circuit is obtained by noting the change of mean detection voltage between the carrier peak voltage limits set by the modulation envelope maximum and minimum values. Taking a carrier peak voltage of 1 volt modulated by 50%, the maximum and minimum of the modulation envelope are 1·5 and 0·5 peak volts, respectively. The mean detected voltage change corresponding to these limits is from 10 to 28 volts. The relationship between carrier peak volts and mean detection volts is

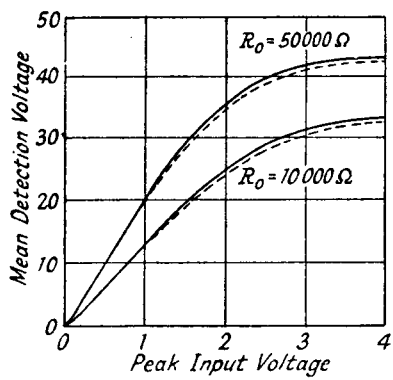


FIG. 8.16.—Typical Detection Curves for a Cumulative Grid Detector.
(Full Line—R_1 = 1 MΩ.
Dotted Line—R_1 = 0·1 MΩ.)

very nearly linear over this range, so that the peak value of the audio output voltage is one-half of the mean detection voltage change, i.e., 9·0 volts. The turn-over effect as I_a approaches the bottom bend of the I_aE_g curve is very clearly shown, and it will be observed that when the carrier voltage is increased beyond 2 volts the mean detection voltage change due to the modulation envelope decreases and becomes distorted. This is the effect producing an apparent double-humped response already discussed in Section 8.3.1.

Characteristic curves [16] rather similar to the I_mE_a characteristic curves of a diode may be obtained for a cumulative grid detector. The circuit is the same as that of Fig. 8.15, except that the anode resistance R_0 is replaced by a low-resistance milliammeter. I_mE_a curves are obtained for different values of input peak voltage for

fixed values of R_1 and C_1 (2 μF to correspond to 0·0001 μF at 1,000 kc/s), and these are indicated in Fig. 8.17. The audio output voltage for any particular H.T. voltage and anode external resistance R_0 can be obtained by drawing the appropriate load lines such as AB or AG. A separate set of curves must, of course, be drawn for each required value of R_1. These curves only represent detection conditions if the anode impedance to the carrier frequency is low. This assumption is generally valid owing to the presence of the R.F. bypass capacitor C_0. The great advantage of the curves lies in the fact that the effect of any values of R_0 with, or without, a decoupling resistance R' may be studied. When a decoupling resistance (R') is employed, a load line, such as AB, is drawn from a point

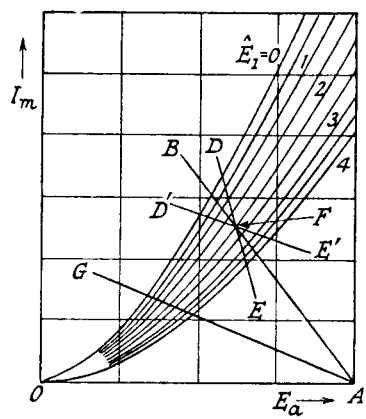


FIG. 8.17.—Typical I_mE_a Curves for a Cumulative Grid Detector.

on the E_a axis equal to the H.T. voltage at an angle $\cot^{-1}(R'+R_0)$, and then another load line DE is drawn at an angle $\cot^{-1} R_0$ through the intersection F of the line AB and the particular carrier peak voltage being considered. The A.F. output voltage is obtained from the intersections of DE with the appropriate carrier voltage lines. $D'E'$ represents the condition for a transformer load with a decoupling resistance R'; AB now has an angle given by $\cot^{-1} R'$.

8.4. Anode Bend Detection.

8.4.1. Introduction. A valve may be made to operate as an anode bend detector if the grid bias is so adjusted that the valve is working on the curved part of its I_aE_g characteristic. Such detection depends mainly on the second derivative [1] of the I_aE_g curve. If we assume that the I_aE_g relationship is given by $I_a = f(E_g)$, a

small change of grid voltage δE_g results in the following current-voltage relationship

$$(I_a+\delta I_a) = f(E_g+\delta E_g)$$

and using Taylor's expansion

$$(I_a+\delta I_a) = f(E_g)+\frac{df(E_g)}{dE_g}\delta E_g+\frac{d^2f(E_g)}{dE_g{}^2}\frac{\delta E_g{}^2}{2!}+\frac{d^3f(E_g)}{dE_g{}^3}\cdot\frac{\delta E_g{}^3}{3!}+\text{ etc.},$$

$$= f(E_g)+f'(E_g).\delta E_g+f''(E_g)\frac{\delta E_g{}^2}{2!}+f'''(E_g)\frac{\delta E_g{}^3}{3!}+\text{ etc.},$$

where $f'(E_g) = \frac{df(E_g)}{dE_g} = \frac{dI_a}{dE_g}$; $f''(E_g) = \frac{d^2I_a}{dE_g{}^2}$, etc.

If $\delta E_g = \hat{E}\cos\omega t(1+M\cos pt)$

$$\delta I_a = f'(E_g)[\hat{E}\cos\omega t(1+M\cos pt)]+f''(E_g)\frac{[\hat{E}\cos\omega t(1+M\cos pt)]^2}{2!}$$

$$+f'''(E_g)\frac{[\hat{E}\cos\omega t(1+M\cos pt)]^3}{3!} \quad . \quad . \quad 8.48.$$

The first and third terms contain only modulated radio frequency components and are therefore of no interest since we require audio frequency components. The second term is

$$\frac{f''(E_g)}{2!}\hat{E}^2\cos^2\omega t(1+M\cos pt)^2$$

$$= \frac{f''(E_g)}{2!}\left[\frac{\hat{E}^2(1+\cos 2\omega t)}{2}(1+2M\cos pt+M^2\cos^2 pt)\right].$$

The radio frequency term is not required and the D.C. and audio frequency components produced by this term are

$$\frac{f''(E_g)}{4}\hat{E}^2\left(1+2M\cos pt+\frac{M^2}{2}(1+\cos 2pt)\right). \quad . \quad 8.49a$$

of which the audio frequency component is

$$E_{AF} = \frac{f''(E_g)}{4}\hat{E}^2\left(2M\cos pt+\frac{M^2}{2}\cos 2pt\right) \quad . \quad . \quad 8.49b.$$

This contains a term of frequency $\frac{p}{2\pi}$ c.p.s., the original modulation component, and a second harmonic component of $\frac{2p}{2\pi}$ c.p.s.

The ratio of second harmonic to fundamental is

$$\frac{M^2}{2}\cdot\frac{1}{2M} = \frac{M}{4} \quad . \quad . \quad . \quad . \quad 8.50.$$

Hence, the modulation ratio of a carrier detected by a parabolic detector should not be high if distortion is to be kept low.

The coefficient $f''(E_g)$ otherwise $\frac{d^2I_a}{dE_g^2}$ is the slope of the mutual conductance-grid bias (g_mE_g) curve of the valve at the particular grid bias point. We may note that it is not only the second derivative term in the Taylor expansion which produces audio frequency components. All even derivatives can produce such components. A better understanding of the physical significance of these derivatives in the Taylor expansion may be gained by considering an I_aE_g curve represented by a power series such as

$$I_a = a_0+a_1E_g+a_2E_g^2+a_3E_g^3+a_4E_g^4+ \ . \ . \ .$$

The first derivative represents the mutual conductance

$$g_m = f'(E_g) = \frac{dI_a}{dE_g} = a_1+2a_2E_g+3a_3E_g^2+4a_4E_g^3 \ . \ . \ .$$

the second is

$$f''(E_g) = \frac{d^2I_a}{dE_g^2} = 2a_2+6a_3E_g+12a_4E_g^2+ \ . \ . \ .$$

In a true parabolic detector all terms above a_2 are zero, hence

$$f''(E_g) = 2a_2 = \text{constant}$$

and the detected audio frequency output voltage is then independent of the bias point on the I_aE_g characteristic. For an actual valve characteristic the terms a_3, a_4, etc., are rarely zero. The a_4 term is usually negative so that $f''(E_g)$ has a maximum value at a particular grid bias. This is clearly shown in the typical detection curves of Fig. 8.19*b*; a bias of -3 volts gives maximum detection. Two grid bias points can generally be found giving optimum detection conditions, for there is a position of maximum rate of change of curvature [$f''(E_g)$] near zero anode current (bottom bend) and near saturation current of the valve (top bend). The "top bend" of the I_aE_g characteristic is never used since it represents a waste of anode current and a considerable reduction in the life of the valve. Furthermore, saturation can usually be reached only by using positive grid bias, and grid current is an undesirable feature because it causes damping of the input circuit.

The operation of a valve as an anode bend detector using the "bottom bend" can be followed from the I_aE_g characteristic curve in Fig. 8.18*a*.

The curve AB is the characteristic curve for zero external anode resistance. This condition is of no use since an output voltage is

required from the detector. A resistance must therefore be inserted in the anode circuit. The operational I_aE_g curve then takes the shape shown by the line AG, and detection occurs because the negative envelope is amplified to a less extent than the positive envelope. This produces a mean current fluctuation which approximately follows the modulation envelope. Harmonics (mainly second) of the audio frequency are actually produced as shown by the mathematical analysis. As the resistance R_0 is increased, the rate of change of curvature decreases (i.e., the slope of AG is less),

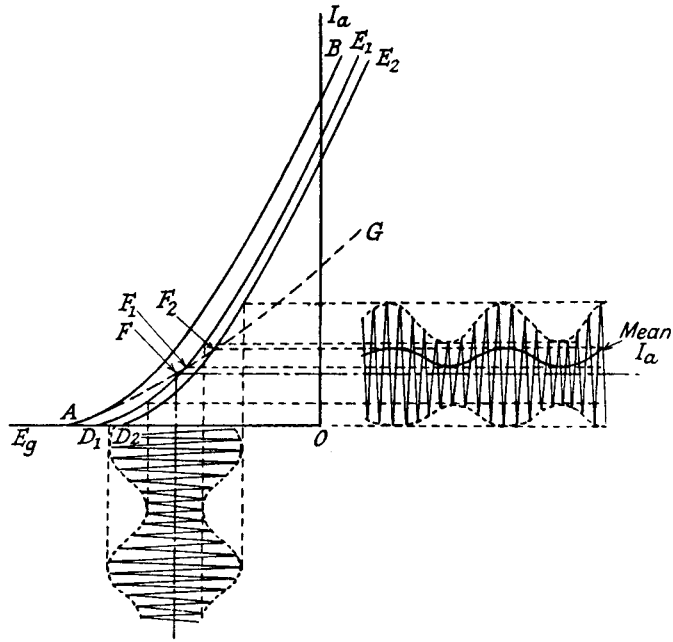


FIG. 8.18*a*.—Anode Bend Detection referred to the I_aE_g Characteristic.

thus decreasing the mean current, but the audio output voltage is actually increased. The case is analogous to that of a generator of fixed internal resistance connected to a variable resistance. As this resistance is increased the voltage across it rises, approaching in the limit the open circuit voltage of the generator.

The decrease in the rate of change of curvature as R_0 is increased is undesirable and we should prefer to maintain the original rate of change of I_a as R_0 is increased. This can be accomplished if the anode impedance is zero, or nearly so, for radio frequencies. The use of a capacitance (C_0 in Fig. 8.19*a*) in parallel with R_0 will produce this result. If the reactance of C_0 is small at radio fre-

quencies the characteristic curve along which detection takes place is one similar to D_1E_1 in Fig. 8.18*a*. It is substantially parallel to the original curve AB for $R_0 = 0$, but is moved towards zero grid volts by an amount governed by the mean current, which in turn is fixed by the carrier amplitude, R_0, and the original grid bias.

Since the mean current is proportional to the carrier amplitude, the position of the operating line varies with the modulation envelope. For minimum carrier amplitude, the mean current is minimum so that the correct line is D_1E_1, but for the maximum carrier amplitude it is D_2E_2. These two lines represent the limits of movement for the particular modulated wave shown, and the

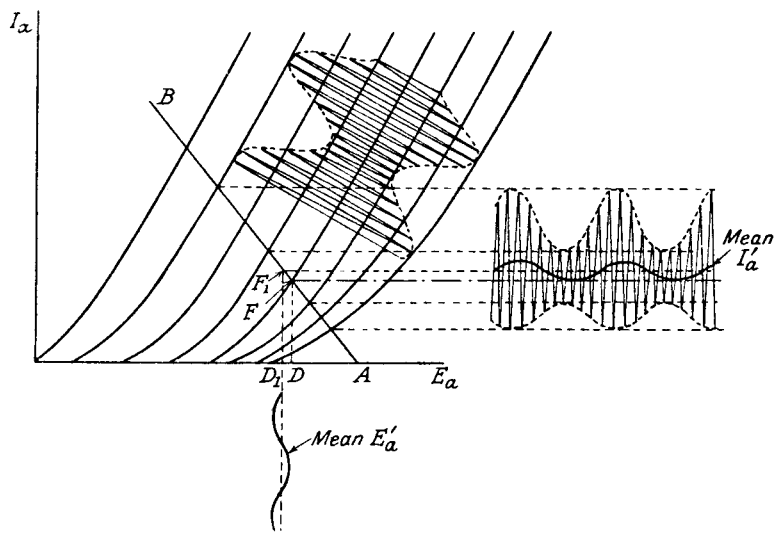


FIG. 8.18*b*.—Anode Bend Detection, with no Anode Capacitance, referred to the I_aE_a Characteristic.

R.F. detection line oscillates between the two positions at an audio frequency corresponding to that of the modulation envelope. Increase of modulation percentage causes F_1 to move nearer to F, the operating point in the absence of a signal, and F_2 to move further from F.

The effect of the capacitance C_0 may be more clearly shown by reference to the I_aE_a curves. Taking the case of the resistance R_0 only in the anode circuit, the conditions are represented in Fig. 8.18*b* where AB is the load line at an angle $\cot^{-1} R_0$. If F is the point corresponding to the grid bias, a modulated wave in the grid circuit produces a distorted modulated anode current wave

which has a mean current variation shown by the full line. The mean current variation is always above the point F unless the input signal is 100% modulated, because even the smallest signal causes some increase in anode current. The mean current variation produces a mean voltage variation given by the full line beneath the E_a axis. The D.C. anode voltage is decreased by detection from its zero signal value OD to OD_1. The mathematical analysis shows in equation 8.49*a* that detection increases the D.C. current by $\frac{f''(E_g)\hat{E}^2\left(1+\frac{M^2}{2}\right)}{4}$. Hence the true operating point moves from F to F_1; the exact position of F_1 depends on the signal voltage,

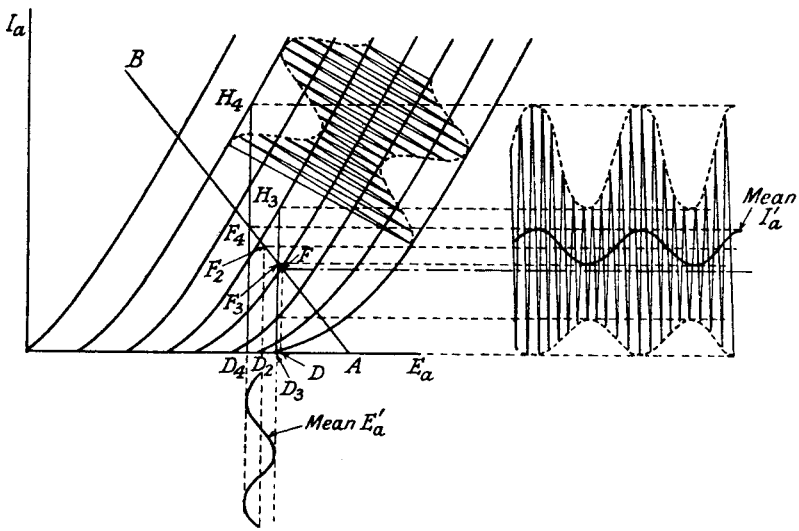


FIG. 8.18c.—Anode Bend Detection, with Anode Bypass Capacitance, referred to the I_aE_a Characteristic.

rising higher on the load line as the carrier amplitude and modulation percentage are increased.

When the resistance R_0 is paralleled by a capacitor C_0, the operating conditions become those of Fig. 8.18c. The audio frequency load line is as before, AB, because the reactance of C_0 is very much greater than R_0 at these frequencies, but the R.F. load line changes to an almost vertical line. This vertical line really approximates to a thin ellipse, the minor axis of which is small since the reactance of C_0 is much less than R_0. The position of this load line is not fixed but varies with the modulation envelope

amplitude between the two points F_3 and F_4. The load line position for maximum envelope amplitude is D_4H_4 and for minimum amplitude D_3H_3. The D.C. anode voltage, which is given by OD_2, is approximately the average of OD_4 and OD_3, and its value is determined by the carrier amplitude and modulation percentage. The anode current wave shape is shown to the right of Fig. 8.18*c*, where the thick curve indicates the fluctuations of mean current. This curve, projected on to the audio frequency load line AB, gives the mean voltage shown below the E_a axis. The addition of the capacitor C_0 has therefore two effects; it has for a given input signal: (1) decreased the D.C. anode voltage and (2) greatly increased the mean anode current and anode voltage variation. The resistance R_0 controls the mean output voltage variation, and maximum gain, which is limited by the amplification factor of the valve, is obtained when R_0 is large. Very high values of R_0 tend to reduce the amplification factor as in the analogous case of the A.F. amplifier (Section 9.3.1) and also cause discrimination against the high audio frequencies by reason of the parallel capacitance C_0. R_0 should not therefore exceed three to four times the internal resistance of the valve.

To obtain optimum detection, variation of anode or grid bias voltage is essential. With small signal voltages detection is generally a maximum for low anode voltages because the rate of change of curvature of the I_aE_g characteristic diminishes as the anode voltage increases. For the same reason a valve having high values of μ and R_a is generally better than one having low values.

If a large input signal is available an anode bend detector may be made to function in a manner approaching a linear[4] detector by biasing the valve almost to the point where $I_a = 0$. The negative modulation envelope is then suppressed and, provided the positive envelope passes over the straight part of the I_aE_g curve, the audio frequency components of the modulation envelope are reproduced without distortion. Under these conditions the output voltage approaches $\frac{\mu R_0}{R_a+R_0}$ times the input modulation envelope. The maximum signal voltage, however, must not be large enough to produce grid current.

8.4.2. Estimation of the Performance of an Anode Bend Detector. Measurements at 50 c.p.s. may be made on a valve to indicate its performance as an anode bend detector. The circuit is shown in Fig. 8.19*a*. The value of C_0 is chosen to give a reactance at 50 c.p.s. equal to the reactance of the capacitor normally used

at radio frequencies. Fig. 8.19*b* shows the presence of an optimum bias point and also the disadvantage of the anode bend detector,

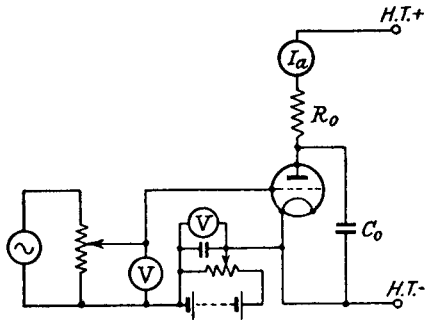


FIG. 8.19*a*.—A Circuit for determining Anode Bend Detection Characteristics.

viz., its inefficient detection of small signal voltages. This feature also limits the maximum modulation percentage of larger signals which can be accepted without serious distortion.

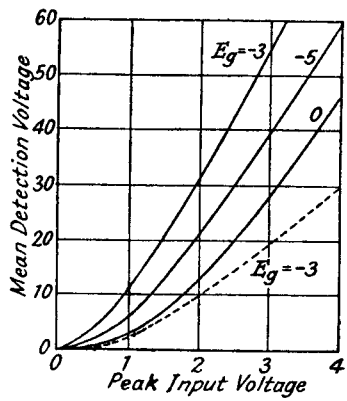


FIG. 8.19*b*.—Typical Detection Curves for an Anode Bend Detector.
(Full Line—R_0 = 50,000Ω.
Dotted Line—R_0 = 10,000Ω.)

8.4.3. Damping of the Input Circuit. The anode bend detector has the advantage that it produces no grid current damping. There is, however, damping due to the anode-grid capacitance,[3] the parallel components of resistance and capacitance being the same as for the cumulative grid detector, viz.,

$$R_g \simeq \frac{C_0}{g_m C_{g_a}}.$$

$$C_g \simeq C_{g_a} \cdot \left[\frac{g_m \left(\frac{1}{R_a} + \frac{1}{R_0} \right)}{(\omega C_0)^2} + 1 \right].$$

Since, however, the anode bend detector operates on the curved part of the characteristic, g_m will be lower than in the case of the cumulative grid detector. For small signals, therefore, the damping is small, but it increases as the signal increases because the valve begins to operate over the higher gain part of the characteristic. This is not a serious disadvantage since strong signals require generally less selectivity than weak signals. The low gain of the detector for small signals is, however, not offset by the lower damping.

8.4.4. Anode Bend Detection with Self-Bias. It is quite usual with indirectly heated anode bend detectors to provide self-bias by a resistance in the cathode circuit. A large capacitance (50 μF) must bypass the resistance if loss of the low audio frequency modulation components is to be prevented, since voltages across the self-bias resistance tend to cancel the anode current changes producing them (see Section 2.7). The resistance value should be chosen to give under no signal conditions the optimum bias. As the signal increases there is a tendency for self-bias to increase and detection efficiency to decrease.

8.5. The Advantages and Disadvantages of the Three Types of Detectors.[13] We may now summarize the advantages and disadvantages of the diode, cumulative grid, and anode bend detectors as follows :

1. *Diode Detector.*

Advantages.

(1) Distortion of the audio frequency components of the modulation envelope decreases as the carrier voltage is increased.
(2) The carrier modulation percentage may be high without distortion becoming appreciable provided certain coupling conditions are fulfilled.
(3) There is no damping of the input circuit other than conduction current damping.
(4) The maximum permissible carrier voltage is almost unlimited.

Disadvantages.

(1) The valve represents a loss of amplification ; this loss may be reduced to small proportions by a suitable choice of R_1 and C_1.
(2) Conduction current produces damping of the input circuit. The use of a high value of R_1 reduces this effect.

2. *Cumulative Grid Detector.*

Advantages.

(1) It is very sensitive for small carrier voltages.
(2) Amplification is obtained from input to output, approximating to that of the valve as an A.F. amplifier.
(3) Distortion is not very great provided the carrier voltage is not too large (the maximum permissible is generally about 1 volt).

Disadvantages.

(1) The input circuit is damped by conduction grid current and by a resistance component reflected from the anode circuit by the anode-grid capacitance.
The latter effect is generally much more serious than grid conduction current damping.
(2) The maximum carrier voltage is limited due to the bottom bend of the I_aE_g characteristic.

3. *Anode Bend Detector.*

Advantages.

(1) There is no damping of the input circuit due to grid current.
(2) Anode-grid capacitance damping is generally low owing to the fact that the valve operates on the low gain bottom bend of the I_aE_g characteristic.
(3) Amplification is obtained from input to output.

Disadvantages.

(1) Sensitivity is very low for small carrier voltages.
(2) Distortion of the audio frequency modulation components is high except for large carrier input voltages with low modulation percentages.
(3) The maximum carrier voltage is limited by grid current.

8.6. Reaction or Regeneration in Detectors. The term reaction, or more correctly regeneration, is used to describe the process by which the output voltage of a valve is coupled back to the input so as to increase the effective input voltage. Regeneration may be produced over any range of frequencies, but in a detector it is generally confined to the input carrier frequency and its modulation side-bands. The degree of regeneration depends on the input voltage, the grid and anode circuit characteristics and the phase relationship between the feedback and initial input voltages. It

leads to accentuation of any already existing frequency discrimination, i.e., it increases the selectivity of its input tuned circuit. In a detector quasi-correct phase relationship is generally established for the carrier frequency, but owing to phase changes of the side-band frequencies by the input circuit, regeneration progressively decreases at frequencies above and below the carrier. This causes a virtual increase in the selectivity of the input tuned circuit.

The improvement in selectivity due to regeneration is not as satisfactory as that produced by adding more tuned circuits. With a high degree of regeneration the virtual selectivity curve is sharply peaked with a narrow pass-band, which rapidly attenuates all the modulation side-bands except those due to the low audio frequencies. This leads to " woolly " A.F. reproduction. The virtual selectivity curve is considerably modified by the presence at the detector of a large undesired signal because the latter increases the bias on the valve (acting as a cumulative grid detector) and reduces its gain as an R.F. amplifier. Regeneration of the desired signal is decreased and the virtual selectivity reduced. To obtain maximum selectivity by regeneration a pre-detector volume control is essential so as to reduce any undesired signal. The desired is simultaneously reduced but the reduction is compensated by increasing regeneration. The need for reducing any undesired signal to the lowest possible value indicates the desirability of using an input tuned circuit having a high rather than a low Q value.

With added tuned circuits, instead of regeneration, the selectivity curve has a much wider and flatter pass-band and attenuation outside the pass region is independent of the magnitude of the undesired signal.

By regeneration the sensitivity of a cumulative grid detector may be increased and damping due to grid current and anode-grid capacitance coupling neutralized. The detector must operate as an R.F. amplifier and a suitable impedance (a R.F. choke) is inserted between the A.F. impedance and the anode. The R.F. anode voltage is returned to the input circuit via a variable capacitance (C_2) and a coil (L_1) as shown in Fig. 8.20. The feedback voltage due to the mutual inductance coupling increases as the frequency increases, and to prevent excessive regeneration at the highest frequency in the tuning range the feedback capacitor (C_2) is of the differential type. Decrease of the feedback capacitance coupling is thus accompanied by increasing shunt capacitance from anode to earth. The maximum value of the differential capacitance should not exceed 0·0003 μF as it is a shunt to the high audio frequency components

of the modulation envelope in the anode circuit. In an alternative circuit the feedback coil L_1 is inserted directly in the anode circuit and the variable capacitance, C_2, decrease of which increases the feedback voltage, is connected from anode to earth and acts as a shunt to R.F. voltages.

As the coupling between anode circuit and input is increased a point is reached where self-oscillation occurs. The transition from the stable to the oscillating condition should be smooth and free from backlash, i.e., a slight reduction of coupling from the oscillating point should stop oscillation. This is essential because the R.F. gain of a cumulative grid detector tends to fall owing to increasing bias as the input carrier signal is increased. With backlash, slight fading of the carrier voltage might cause oscillation which would persist when the carrier voltage returned to its original value.

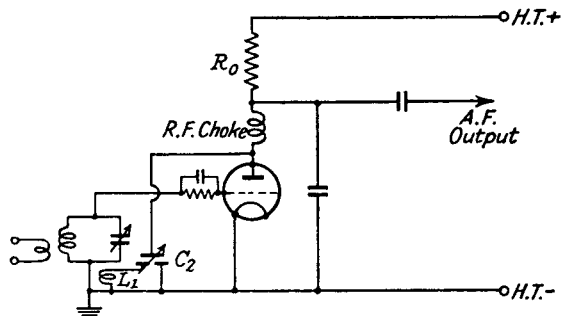


FIG. 8.20.—A Typical Circuit for a Regenerative Detector.

Unstable regeneration is generally due to incorrect phasing of the feedback and input voltages and to an R.F. amplification characteristic which increases as the input increases. Incorrect phasing may largely be avoided by making the resonant frequency of the feedback circuit much higher than the highest required input tuning frequency. The feedback coil, having a small number of turns, should therefore be tightly coupled to the earthed end (to avoid capacitance coupling) of the input coil. Combined capacitance and mutual inductance coupling may produce excessive regeneration or stop regeneration at some point in the tuning range of the input circuit.

Since increasing carrier voltage increases the bias on a cumulative grid detector, its gain as an R.F. amplifier tends to fall. This is quite different from the R.F. amplification characteristic of an anode bend detector, for owing to the curvature of the I_aE_g characteristic the R.F. anode voltage increases at a greater rate than the input

voltage, i.e., the gain increases as the signal increases. Hence unstable regeneration is usually a feature of such detectors.

Although the anode current characteristic of a cumulative grid detector tends to give stable regeneration, the grid detection characteristic can produce instability and backlash. In Fig. 8.11*b*, it is shown that the equivalent damping resistance due to detection is dependent on the input voltage when detection current begins at a voltage other than zero. If therefore grid current begins in the cumulative grid detector at some negative or equivalent negative voltage (as in an indirectly heated valve or a directly heated valve with the grid leak returned to the L.T. positive), damping of the input circuit decreases as the signal is increased. Reduced damping increases the R.F. voltage at the grid and the detector has therefore a tendency to unstable regeneration. Backlash is increased if the negative voltage start of grid current is increased. On the other hand, a detector having a positive voltage start of grid current tends to give stable regeneration since curve 2, Fig. 8.11*b*, shows increasing damping as the signal is increased. The positive voltage start is, however, undesirable because of distortion (see Section 8.2.7) and the ideal condition is obtained with zero start of grid current.

Increasing the detector grid leak resistance decreases the damping of the input tuned circuit and tends to more stable operation.

Regeneration may be applied to a diode detector followed by an A.F. amplifier valve by using the latter as the regenerator, but it is not usual because most of the advantages of diode detection are lost if R.F. voltages are passed to the succeeding A.F. amplifier. When regeneration is required the R.F. filter circuit (R_3C_3 in Fig. 8.5) between the detector and amplifier must be removed and an R.F. impedance, coupled back to the detector input tuned circuit, must be included in the A.F. amplifier anode circuit.

We may state the points for and against regeneration in detectors as follows. The advantages are : considerably increased sensitivity and selectivity (provided any undesired signal voltages are small). The disadvantages are that as regeneration is increased the tuning of the input circuit generally needs adjustment (the degree of correction depends on the phase angle between the feedback and input voltages and is small when this is small), adjustment of the regeneration control is necessary as the input tuning frequency is changed, radiation may occur if regeneration is carried as far as oscillation, and the presence of an undesired signal modifies the regeneration characteristic.

8.7. Detection with Push-Pull Output.[22] It may some-

times be advantageous to provide push-pull output from a detector and two possible circuits are shown in Figs. 8.21*a* and 8.21*b*. The first gives push-pull output from a diode, and is applicable when one side of the input circuit is earthed. When the circuit need not be earthed (the secondary of an I.F. transformer) the R.F. choke L

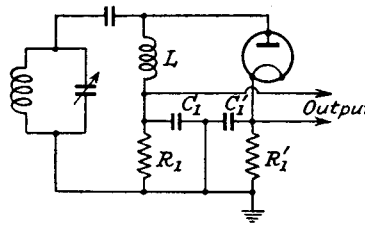


FIG. 8.21*a*.—Push-Pull Output from a Diode Detector.

may be removed and the tuned circuit inserted in its place. The two output voltages across R_1 and R_1' may be slightly unbalanced at high audio frequencies due to unequal stray capacitances, but this effect is not usually serious. Heater to cathode insulation must be high or interference from hum is likely to be experienced.

Fig. 8.21*b* shows how push-pull output may be obtained from a cumulative grid detector. Anode bend detection can equally well be employed, but both detectors require an input circuit isolated from earth. Capacitors C_0 (0·0001 μF) bypass R.F. to earth.

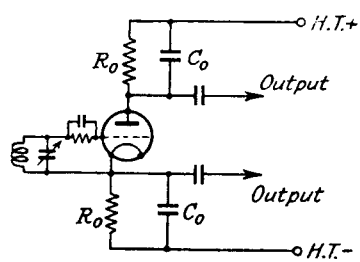


FIG. 8.21*b*.—Push-Pull Output from a Triode Detector.

8.8. Double-Wave Detection. The chief advantage of double-wave detection is the smaller R.F. ripple voltage across the load resistance R_1, which means that less R.F. filtering is required between R_1 and the A.F. amplifier. A centre tapped coil is necessary (see Fig. 8.22) and this reduces the available A.F. voltage from the load resistance to one-half that for half-wave detection. On the other hand, the equivalent damping resistance due to the diode conduction current is increased to twice that for half-wave detection. It

has been shown [25] that only under special circumstances can double-wave detection prove superior to half-wave detection, and the latter is almost universally employed in receivers.

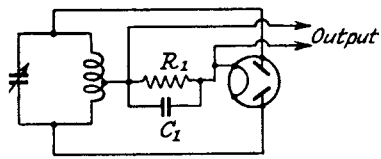


FIG. 8.22.—Double Wave Detection with a Diode.

8.9. The Anode Bend Detector with Negative Feedback.[26] The anode bend detector with negative feedback (see Fig. 8.23*a*) has an action similar to that of a diode with current starting at a high negative voltage. It has a linear detection characteristic for carrier voltages exceeding about 0·5 volts, a high critical modulation ratio (see Section 8.2.3), and the maximum permissible carrier voltage is almost unlimited. As in the case of the normal anode bend detector, detection is obtained by the parabolic or unidirectional property of the I_aE_g characteristic. No grid current is taken and there is therefore practically no damping of the input tuned circuit. In its action it is similar to the diode with a negative start of anode current, but it has no serious damping effect for small signal voltages as shown by the diode (Fig. 8.11*b*, curve 1). High input admittance, the resistance component of which may be negative, and high critical modulation ratio are the important ad-

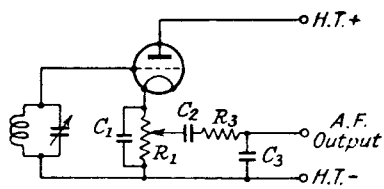


FIG. 8.23*a*.—The Anode Bend Detector with Negative Feedback.

vantages of this type of detector. Its chief disadvantage (shared by the diode) is that its amplification factor is less than 1 and its detection efficiency is of the same order as that of the diode, viz., about 90% for $R_1 = 1\ \text{M}\Omega$. Since the I_aE_g characteristic has usually a greater initial curvature than the I_aE_a characteristic of a diode the detection efficiency for small signals is lower than that of a diode with zero start of current, though not lower than that of the diode with negative start of current. We should note that the diode

with negative start of current (Section 8.2.6) is identical with the diode having zero start of current and positive bias (Section 8.2.3).

Detection $I_m E_g$ characteristic curves (Fig. 8.23*b*) similar to the diode curves of Fig. 8.4*b* may be obtained by using the 50 c.p.s. mains supply as an input and replacing $R_1 C_1$ in Fig. 8.23*a* by a microammeter and a battery as illustrated in Fig. 8.4*a* for the diode. The negative start ($-E_b$) of anode current gives a critical modulation ratio in accordance with expression 8.6, viz.,

$$M = \left[\frac{\hat{E}_1 + E_b}{\hat{E}_1}\right]\left[1 - \eta_d \frac{R_1}{R_1 + R_2}\right],$$

so that modulation distortion due to the grid leak (R_2) in the A.F. amplifier following the detector is appreciably reduced. The value

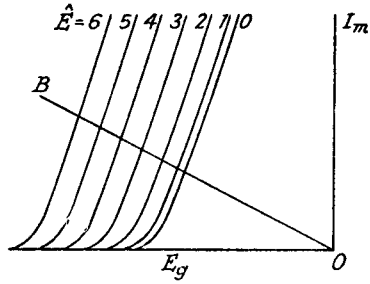


FIG. 8.23*b*.—Typical Curves for an Anode Bend Detector with Negative Feedback.

of $-E_b$, which is entirely controlled by the anode voltage, can be increased by increasing the latter. A characteristic of this type of detector is that it presents a negative input conductance, i.e., produces regeneration and improves the selectivity of the input tuned circuit across which it is connected. This negative conductance is due to feedback through the grid-cathode capacitance from the R.F. voltage developed across the capacitive cathode load impedance (see Section 2.8.3).

8.10. Interference Effects due to an Undesired R.F. Signal in the Detector Input Circuit. The desired A.F. output voltage from a detector is generally affected by the presence of an undesired R.F. voltage in the input circuit. The extent of the interference produced depends on the relative magnitudes and frequency separation of the desired and undesired carrier voltages.

If the frequency difference between the two carrier voltages is an inaudible frequency, the interference effect is entirely dependent on their relative magnitudes. An unmodulated R.F. signal compar-

able in value to the desired signal reduces and distorts the A.F. modulation components of the latter. If the former is modulated the undesired modulation (generally distorted) is also heard. The demodulation [11] effect (reduction of the A.F. output voltage from a modulated carrier due to the presence of another carrier at the detector input) occurs for either desired or undesired signals, and if the magnitude of the undesired signal is small it may be almost completely demodulated by the desired carrier. That is to say, the undesired modulation is audible in the absence of the desired carrier, but becomes inaudible when the latter is being transmitted.

If the frequency difference between the undesired and desired carrier voltages is an audible frequency, additional interference effects are observed. The difference frequency appears as a heterodyne whistle together with the modulation components of the undesired carrier, which are generally distorted. Another effect which may also occur is that of frequency inversion of the undesired modulation side-bands, the high-frequency components appearing in the A.F. output as low frequencies and vice versa. Both effects give rise to an A.F. output which is best characterized as a chatter. An example may often be found on the medium-wave range of a receiver tuned to a distant station adjacent in frequency to a strong local station. The frequency separation of 9 kc/s appears as a heterodyne whistle of 9 kc/s, and the modulation side-bands produce chatter due to distortion and frequency inversion, an undesired modulation frequency of 1 kc/s. being produced as a frequency of 8 kc/s, i.e., 9 − 1, a frequency of 3·5 kc/s as 5·5 kc/s, i.e., 9 − 3·5 ; and so on.

To examine the theory underlying these distortion effects we will consider separately the linear and parabolic detector. Let us imagine that the undesired carrier only is applied to a linear detector. It produces a detection current which causes a voltage, equivalent to a negative bias on the diode anode, to appear across the diode load resistance. As shown in Section 8.2.7, no detection occurs until the signal voltage exceeds any negative bias applied to the diode anode. The comparison of the effect produced by the undesired carrier to that produced by a fixed negative bias is not strictly correct since some detection of the desired signal takes place, even when the latter is much less than the undesired carrier. For a large desired signal, however, the undesired carrier has practically the same effect as a fixed negative bias equal in value to the D.C. voltage produced by the undesired carrier acting alone. The influence of the undesired on the mean detection voltage-desired

peak carrier curves is shown in Fig. 8.24. These are idealized curves calculated [12] on the assumption that the detector has an efficiency of $\eta_d = 100\%$. Similar curves may be obtained experimentally using the circuit in Fig. 8.2 with two signal voltages at differing frequencies connected in series in place of the generator

$$\hat{E}_1 \cos \omega t(1+M \cos pt).$$

Curves 1, 2 and 3 give the detection characteristic for undesired peak voltages of $\hat{E}_u = 0$, 0·25 and 1 volt. The A.F. output voltage wave shape for a desired carrier of $\hat{E}_d = 0{\cdot}2$ volts modulated 50% is shown to the right of the figure. Demodulation and distortion

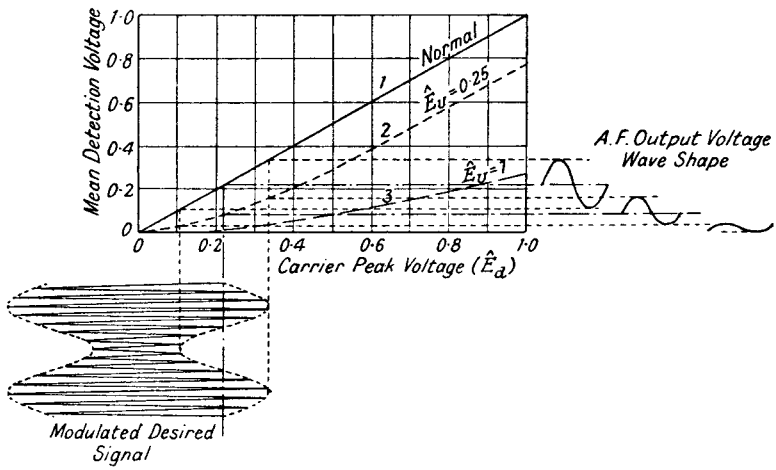


FIG. 8.24.—Demodulation of a Small by a Large Signal at the Detector.

of the modulation envelope are clearly marked as $\hat{E}_u$ is increased. Both effects are mutual and are greatest for the smaller signal, which in practice is generally $\hat{E}_u$. We see from curve 2 that when $\hat{E}_d$ exceeds $\hat{E}_u$, the detection curve becomes straight and parallel to curve 1, thus showing that $\hat{E}_u$ has an effect similar to that of negative bias. If $\hat{E}_u$ is unmodulated and differs from $\hat{E}_d$ by an inaudible frequency, the output A.F. voltage is unaffected when the modulation is restricted to operation over the straight part of curve 2, i.e., the modulation trough $\hat{E}_d(1-M)$ must not fall below about 0·4 volts.

For the parabolic detector let us assume its $I_a E_g$ characteristic to be given by

$$I_a = a_0 + a_1 E_g + a_2 E_g^2 + a_3 E_g^3 - a_4 E_g^4 + \ldots \qquad 8.51.$$

(The reason for the coefficient $-a_4$ is explained in section 8.4.1.)

But $E_g = \hat{E}_d \cos \omega_d t(1+M_d \cos p_d t)+\hat{E}_u \cos \omega_u t(1+M_u \cos p_u t) - E_b$ where $-E_b$ = the bias voltage.

Replacing E_g in equation 8.51 and analysing the anode current into its A.F. components as in Section 8.4.1, we find the following effects in the detector output.

1. Undesired modulation.

This results from the $a_2 E_g^2$ term and is

$$\frac{a_2 \hat{E}_u^2}{2}\left[2M_u \cos p_u t+\frac{M_u^2}{2} \cos 2p_u t\right].$$

The distortion term in the above expression is not for our purposes classed as such since it is due to the parabolic property and not to the interaction of the two signals.

2. Difference frequency and inversion of the undesired modulation components.

These are obtained from

$$a_2[2\hat{E}_d \cos \omega_d t(1+M_d \cos p_d t)\hat{E}_u \cos \omega_u t(1+M_u \cos p_u t)].$$

The difference frequency component

$$a_2 \hat{E}_d \hat{E}_u \cos (\omega_d - \omega_u)t$$

is contained in

$$2a_2 \hat{E}_d \cos \omega_d t \hat{E}_u \cos \omega_u t$$

for $\cos \theta \cos \phi = \frac{1}{2}[\cos (\theta - \phi)+\cos (\theta+\phi)]$

and the frequency inversion term

$$\frac{a_2 \hat{E}_d \hat{E}_u M_u}{2} \cos (\omega_d - \omega_u - p_u)t.$$

occurs in

$$a_2 2\hat{E}_d \hat{E}_u M_u \cos \omega_d t \cos \omega_u t \cos p_u t.$$

If $\frac{\omega_d - \omega_u}{2\pi} = 9$ kc/s, and $\frac{p_u}{2\pi} = 8$ kc/s

$$\frac{(\omega_d - \omega_u - p_u)}{2\pi} = 1 \text{ kc/s},$$

thus an 8 kc/s. undesired modulation frequency appears as a 1 kc/s. frequency in the output.

The desired modulation also has an inverted component for equation 8.51 gives

$$\frac{a_2 \hat{E}_d \hat{E}_u M_d}{2} . \cos (\omega_d - \omega_u - p_d)t.$$

3. Demodulation of the desired signal.

The demodulation term appears from $-a_4E_g^4$ for this contains

$$-a_4 6\hat{E}_d^2 \cos^2 \omega_d t(1+M_d \cos p_d t)^2 \hat{E}_u^2 \cos^2 \omega_u t(1+M_u \cos p_u t)^2,$$

which gives the component

$$-\frac{3a_4\hat{E}_d^2\hat{E}_u^2}{2}(1+M_d \cos p_d t)^2$$

$$= -\frac{3a_4\hat{E}_d^2\hat{E}_u^2}{2}\left(2M_d \cos p_d t+\frac{M_d^2}{2}\cos 2p_d t\right).$$

This subtracts from the A.F. term

$$\frac{a_2\hat{E}_d^2}{2}\left(2M_d \cos p_d t+\frac{M_d^2}{2}\cos 2p_d t\right)$$

in $a_2E_g^2$.

A similar demodulating term is also obtained for the undesired signal.

4. Distortion of the desired and undesired modulation components.

A distorted desired and undesired modulation component is contained in $-a_4E_g^4$

$$-\frac{6a_4\hat{E}_d^2\hat{E}_u^2}{4}(1+M_d \cos p_d t)^2(1+M_u \cos p_u t)^2$$

gives

$$-\frac{3}{4}a_4\hat{E}_d^2\hat{E}_u^2\left[M_d^2 \cos 2p_d t+M_u^2 \cos 2p_u t\right].$$

In addition there are sum and difference frequencies $\frac{(2p_d \pm p_u)}{2\pi}$, $\frac{2p_u \pm p_d}{2\pi}$, and $\frac{(2p_d \pm p_u)}{2\pi}$.

BIBLIOGRAPHY

1. A Theoretical and Experimental Investigation of Detection for Small Signals. E. L. Chaffee and G. H. Browning, *Proc. I.R.E.*, Feb. 1927, p. 113.
2. Detection by Grid Rectification with the High Vacuum Triode. S. Ballantine, *Proc. I.R.E.*, May 1928, p. 593.
3. Effect of Anode-Grid Capacity in Anode-Bend Rectifiers. E. A. Biedermann, *Wireless Engineer*, Feb., p. 71, March, 1929, p. 135.
4. Detection at High Signal Voltage—Plate Rectification with the High Vacuum Triode. S. Ballantine, *Proc. I.R.E.*, July 1929, p. 1,153.
5. The Numerical Estimation of Grid Rectification for Small Signal Amplitudes. W. A. Barclay, *Wireless Engineer*, Nov. 1929, p. 596.
6. The Theory of the Straight Line Rectifier. F. M. Colebrook, *Wireless Engineer*, Nov. 1930, p. 595.

7. The Estimation of the Sensitivity of the Grid Rectifier for Large Inputs. C. D. Hall, *Wireless Engineer*, Dec. 1930, p. 668.
8. Some Properties of Grid Leak Power Detection. F. E. Terman and N. R. Morgan, *Proc. I.R.E.*, Dec. 1930, p. 2,160.
9. Grid Circuit Power Rectification. J. R. Nelson, *Proc. I.R.E.*, March 1931, p. 489.
10. Test Procedure for Detectors with Resistance Coupled Output. G. D. Robinson, *Proc. I.R.E.*, May 1931, p. 806.
11. Mutual Demodulation and Allied Problems. G. W. O. Howe, *Wireless Engineer*, Aug. 1931, p. 405.
12. The Apparent Demodulation of a Weak Station by a Stronger One. F. M. Colebrook, *Wireless Engineer*, Aug. 1931, p. 409.
13. Quality Detectors. W. Greenwood and S. J. Preston, *Wireless Engineer*, Dec. 1931, p. 648.
14. The Graphical Solution of Detector Problems. G. S. C. Lucas, *Wireless Engineer*, April 1932, p. 202.
15. Some Notes on Grid Circuit and Diode Rectification. J. R. Nelson, *Proc. I.R.E.*, June 1932, p. 989.
16. A New Valve Characteristic. P. K. Turner, *Wireless Engineer*, July 1932, p. 384.
17. Discussion on Grid Circuit Detection and Diode Rectification. J. R. Nelson and F. E. Terman, *Proc. I.R.E.*, Dec. 1932, p. 1971.
18. Some Notes on the Use of a Diode as a Cumulative Grid Rectifier. E. A. Biedermann, *Wireless Engineer*, March 1933, p. 123.
19. Discussion on " Some Notes on Grid Circuit and Diode Rectification ". F. G. Kelly, *Proc. I.R.E.*, April 1933, p. 630.
20. Diode Detection Analysis. C. E. Kilgour and J. M. Glessner, *Proc. I.R.E.*, July 1933, p. 930.
21. Notes on Screened Grid Pentode Detectors. F. R. W. Strafford, *Wireless Engineer*, Sept. 1934, p. 484.
22. Push-Pull Input Systems. W. T. Cocking, *Wireless World*, Sept. 21st, 1934, p. 245.
23. Diode Detectors. J. B. L. Foot, *Wireless World*, Dec. 28th, 1934, p. 547.
24. Notes on the Theory of Diode Rectification. J. Marique, *Wireless Engineer*, Jan. 1935, p. 17.
25. The Detector Input Circuit. W. T. Cocking, *Wireless Engineer*, Nov. 1935, p. 595.
26. New Detector Circuit. W. N. Weeden, *Wireless World*, Jan. 1st, 1937, p. 6.
27. The Modulation Response and Selectivity Curves of a Resonant Circuit loaded by a Diode Rectifier. F. C. Williams, *Wireless Engineer*, April 1938, p. 189.
28. The Properties of a Resonant Circuit loaded by a Complex Diode Rectifier. F. C. Williams, *Wireless Engineer*, Nov. 1938, p. 600.
29. The Diode Detector with Positive Bias. K. R. Sturley, *Wireless World*, March 9th, 1939, p. 220.
30. Diode Operating Conditions. W. P. N. Court, *Wireless Engineer*, Nov. 1939, p. 548.

APPENDIX 1A

"j" NOTATION

THE use of "j" notation to represent complex quantities is a most useful aid in the solution of A.C. problems. In Fig. 1A.1 are shown two vectors of equal length "a" units, one OA is horizontal and the other OB is vertical.

Let us assume that the vectors are represented respectively by "a" and "ja," the factor "j" denoting that the original vector OA has been rotated through 90° in an anti-clockwise direction. By taking the process a step further we may multiply vector OB by j and obtain vector OC, which is a vector equal in length but opposite in direction to OA, thus

$$OC = j\,OB = j.(jOA) = j^2a = -a$$

or

$$j = \sqrt{-1} \quad . \quad . \quad . \quad . \quad 1A.1.$$

In the same manner dividing by "j" can denote rotation of a vector through 90° in a clockwise direction; hence $OD = \frac{OA}{j} = -j.OA$. The result is in agreement with normal co-ordinate axes for

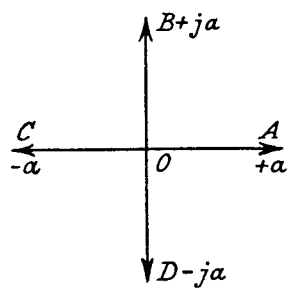


FIG. 1A.1.—Real and Imaginary Axes.

$OD = -j.OA = -OB$. The horizontal vectors OA and OC are known as real, and the vertical vectors as imaginary quantities. This method of vector representation is particularly useful in dealing with circuits containing inductance and capacitance as well as resistance. Taking as an example an inductance L and resistance R connected in series across a generator supplying a voltage E at a frequency f c.p.s., the voltages across L and R cannot simply be

added because that across L is advanced in time by 90° with respect to the current as shown in Fig. 1A.2. The voltage across the resistance R is in phase with the current I. By the theorem of Pythagoras the total voltage E is equal to $\sqrt{E_R^2+E_L^2}$. The voltage E_R across R is IR, and that E_L across L is IX_L.
The impedance Z of the whole circuit is

$$Z = \frac{E}{I} = \sqrt{R^2+(\omega L)^2} \quad . \quad . \quad . \quad \text{1A.2}$$

where $$\omega = 2\pi f.$$

Since the reactance ωL is 90° ahead of R we may write the impedance Z in terms of "j" notation as

$$Z = R+j\omega L \quad . \quad . \quad . \quad . \quad \text{1A.3.}$$

Similarly a series circuit of resistance R and capacitance C gives an impedance

$$Z = \sqrt{R^2+\left(\frac{1}{\omega C}\right)^2} = R+\frac{1}{j\omega C} = R-\frac{j}{\omega C} \quad . \quad \text{1A.4}$$

for the voltage across the capacitance lags 90° behind that across the resistance, and the capacitive reactance vector must therefore

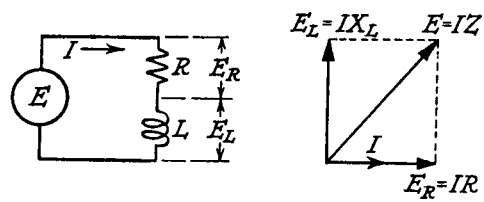


FIG. 1A.2.—The Vector Representation of the Voltages and Current in a Series Circuit.

be divided by "j." Before demonstrating the advantages of "j" notation (in the above elementary circuits it is difficult to see that simplification has resulted) certain points with regard to its use should be considered.

1. A vector impedance $Z = R+jX$ has an amplitude $|Z| = \sqrt{R^2+X^2}$, and a phase angle $\phi = \tan^{-1}\frac{X}{R}$; for a vector impedance $Z = R-jX$, the amplitude $|Z|$ is still unchanged and equal to $\sqrt{R^2+X^2}$, but its phase angle is $\phi = \tan^{-1}\frac{-X}{R}$; i.e., the reactance term is capacitive. Thus the amplitude is inde-

pendent of the sign of the real and imaginary terms in the "j" representation. Only the phase angle is affected.

2. A term containing a complex denominator must be rationalized before the phase angle is determined, for example, a circuit having an impedance $Z = R+jX$ has an admittance $Y=\frac{1}{Z} = \frac{1}{R+jX}$. The amplitude of the admittance $|Y| = \frac{1}{\sqrt{R^2+X^2}}$, but its phase angle is not $\phi = \tan^{-1}\frac{X}{R}$ but

$$\phi = \tan^{-1} -\frac{X}{R} \text{ for } Y = \frac{R-jX}{(R+jX)(R-jX)} = \frac{R-jX}{R^2+X^2}.$$

3. Complex vectors may be multiplied together; the product of the real and imaginary components gives an imaginary, whilst the product of two imaginaries gives a real term, for $j^2 = -1$.

4. For the impedance of parallel circuits or the admittance of series circuits, a complex denominator $a+jb$ must always be rationalized by multiplying numerator and denominator by $(a - jb)$. The real and imaginary parts, which result are the resistance and reactance term respectively in the case of impedance, and the conductance and susceptance term in the admittance case.

For example, $$Z = \frac{10-4j}{2+3j} = \frac{(10-4j)(2-3j)}{2^2+3^2}$$
$$= \frac{8-38j}{13} = 0{\cdot}615 - 2{\cdot}925j$$

thus Z is equivalent to a resistance of 0·615 ohms in series with a capacitive reactance of 2·925 ohms. Similarly $Y = \frac{10-4j}{2+3j}$ is equivalent to a conductance of 0·615 mhos in parallel with an inductive susceptance of 2·925 mhos.

5. An inductive impedance is always represented by $a+jb$ and a capacitive impedance by $a - jb$, whilst the reverse is true for admittance; the $+j$ term is in this case a capacitive susceptance. This may be seen by inverting an inductive reactance to obtain an inductive susceptance.

$$j\omega L \text{ reactance} = \frac{1}{j\omega L} \text{ susceptance}$$
$$= \frac{-j}{\omega L}.$$

The advantage of "j" notation can be demonstrated when it is

desired to convert a parallel circuit such as that of Fig. 1A.3 into its equivalent series circuit.

$$Z = \frac{\frac{Rj\omega L}{j\omega C}}{Rj\omega L + \frac{R}{j\omega C} + \frac{j\omega L}{j\omega C}} = \frac{j\omega LR}{R(1 - \omega^2 LC) + j\omega L}.$$

Rationalizing

$$Z = \frac{j\omega LR[R(1 - \omega^2 LC) - j\omega L]}{[R(1 - \omega^2 LC)]^2 + (\omega L)^2}.$$

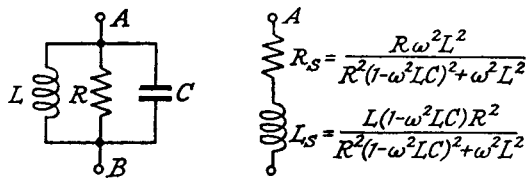


FIG. 1A.3.—A Parallel Circuit and its Series Equivalent.

The equivalent series circuit is obtained by separating the real or resistive term from the imaginary or reactive term, thus

$$Z = R_s + jX_s.$$

$$= \frac{R\omega^2 L^2}{R^2(1 - \omega^2 LC)^2 + (\omega L)^2} + \frac{j\omega L(1 - \omega^2 LC)R^2}{R^2(1 - \omega^2 LC)^2 + (\omega L)^2} \quad . \quad 1A.5.$$

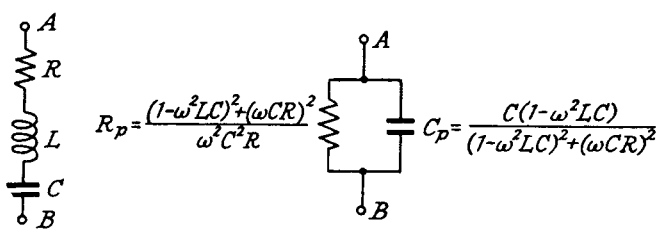


FIG. 1A.4.—A Series Circuit and its Parallel Equivalent.

In a similar manner the series circuit L, C and R (Fig. 1A.4) may be converted into an equivalent parallel circuit.

$$Y = \frac{1}{Z} = \frac{1}{R + j\omega L + \frac{1}{j\omega C}} = \frac{j\omega C}{(1 - \omega^2 LC) + j\omega CR}$$

$$= \frac{j\omega C[(1 - \omega^2 LC) - j\omega CR]}{(1 - \omega^2 LC)^2 + (\omega CR)^2}$$

$$= \frac{R\omega^2 C^2}{(1 - \omega^2 LC)^2 + (\omega CR)^2} + \frac{j\omega C(1 - \omega^2 LC)}{(1 - \omega^2 LC)^2 + (\omega CR)^2} \quad . \quad 1A.6$$

$$= \frac{1}{R_p} + \frac{j}{X_p}$$

where R_p is the equivalent resistance in parallel with X_p the equivalent reactance.

The simplification resulting from the use of "j" notation is even more clearly demonstrated when the impedance or admittance of a complex circuit such as that of Fig. 7.6 is required and its importance cannot be over emphasised.

APPENDIX 2A

FOURIER SERIES

THE Fourier method of analysing a complex wave shape into a series of harmonically related sinusoidal waves has proved of paramount importance to the radio engineer. By the principle of superposition * it is possible to consider separately the action of each frequency component of a complex current or voltage wave applied to a circuit, and all effects may later be added together to give the final result at the output of the circuit. The complex wave is generally considered as periodic, but a transient wave can be treated by the Fourier Integral Method. We shall, however, limit ourselves to an examination of the periodic complex wave.

The Fourier theorem states that any periodic complex wave

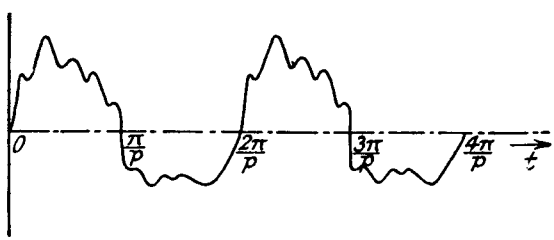


FIG. 2A.1.—An Example of a Periodic Wave which can be resolved into its Components by Fourier Analysis.

shape, such as that of Fig. 2A.1, may be represented by a series of sines and cosines of possibly infinite number, i.e.,

$$f(t) = a_1 \sin pt + a_2 \sin 2pt + \ldots a_n \sin npt$$
$$+\frac{b_0}{2} + b_1 \cos pt + b_2 \cos 2pt + \ldots b_n \cos npt \quad . \quad 2A.1.$$

The significance of writing $\frac{b_0}{2}$ as the first term of the cosine series is considered later. Certain conditions must be fulfilled if the above equation is to hold.

* The superposition theorem states that in any network consisting of generators and linear bilateral impedances the current flowing at any point is the sum of the currents which would flow if each generator were considered separately, all other generators being replaced at the time by impedances equal to their internal impedances.

1. The function $f(t)$ must have only one value for a given value of "t", except where there is a discontinuity, e.g., a circle cannot be treated, but the instantaneous fall from A to B in Fig. 2A.2 can be considered.

2. If the wave is not continuous the number of discontinuities must be finite. The value of $f(t)$ at the discontinuity must also be finite. At a point of discontinuity the Fourier series gives the average value of the wave.

Two terms, $a_1 \sin pt$ and $b_1 \cos pt$, make up the fundamental frequency which is the reciprocal of the periodic time of the complex wave. Two terms are necessary to define each frequency component since correct positioning on the time scale is as important as correct amplitude.

The double fundamental term may be rewritten

$$\sqrt{a_1^2+b_1^2}\left[\frac{a_1}{\sqrt{a_1^2+b_1^2}} \sin pt + \frac{b_1}{\sqrt{a_1^2+b_1^2}} \cos pt\right]$$

which can be simplified to

$$c_1 \sin (pt+\phi_1) \quad \text{where} \quad c_1 = \sqrt{a_1^2+b_1^2}$$
$$\text{and} \quad \phi_1 = \tan^{-1}\frac{b_1}{a_1}.$$

Similarly, every frequency component may be written as $c_n \sin (npt+\phi_n)$. The angle ϕ_n determines the position of each frequency component on the time scale since it fixes the value of any frequency term when $t = 0$.

The values of a_n and b_n may be obtained by multiplying $f(t)$ in equation 2A.1 by the appropriate sine and cosine function and integrating this product over the periodic interval from 0 to $\frac{2\pi}{p}$. Thus a_n may be found by integrating equation 2A.1 multiplied by $\sin npt$.

$$\int_0^{\frac{2\pi}{p}} f(t) \sin npt\, dt = \int_0^{\frac{2\pi}{p}} [a_1 \sin pt + \ldots a_m \sin mpt + \ldots + \frac{b_0}{2} + b_1 \cos pt + \ldots b_m \cos mpt + \ldots] \sin npt\, dt \quad \ldots \quad \text{2A.2.}$$

The right-hand side of equation 2A.2 simplifies considerably because for any value of m or n the expression $\int_0^{\frac{2\pi}{p}} \cos mpt \sin npt\, dt$

is zero and except when $m = n$ the expression $\int_0^{\frac{2\pi}{p}} \sin mpt \sin npt\, dt$ is also zero. Hence

$$\int_0^{\frac{2\pi}{p}} f(t) \sin npt\, dt = a_n \int_0^{\frac{2\pi}{p}} \sin^2 npt\, dt.$$

$$= a_n \int_0^{\frac{2\pi}{p}} \frac{(1 - \cos 2npt)}{2} dt$$

$$= a_n \left[\frac{t}{2} - \frac{\sin 2npt}{4np}\right]_0^{\frac{2\pi}{p}}$$

$$= a_n \frac{\pi}{p}$$

or
$$a_n = \frac{p}{\pi} \int_0^{\frac{2\pi}{p}} f(t) \sin npt\, dt \quad . \quad . \quad . \quad 2A.3$$

where n may have any value from 1 to ∞.
A similar expression is obtained for b_n

$$b_n = \frac{p}{\pi} \int_0^{\frac{2\pi}{p}} f(t) \cos npt\, dt \quad . \quad . \quad . \quad 2A.4$$

where n has any value from 0 to ∞.
The factor $\frac{1}{2}$ multiplying b_0 in 2A.1 is necessary to make expression 2A.4 true when $n = 0$. The Fourier expression for the curve represented by $f(t)$ becomes

$$f(t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} a_n \sin npt + \sum_{n=1}^{\infty} b_n \cos npt. \quad . \quad 2A.5.$$

Let us now take certain examples which will make clear the process of Fourier analysis. Let us assume that we wish to know the frequency components of voltage produced across a resistance connected in series with a half-wave rectifier. The voltage wave shape is that of a sine wave with the negative half suppressed. If the pulsance of the applied voltage to the rectifier is p rads/sec.

$$a_n = \frac{p}{\pi} \int_0^{\frac{2\pi}{p}} f(t) \sin npt\, dt.$$

$f(t)$ is made up of two parts; from 0 to $\frac{\pi}{p}$ it is represented by $\hat{E} \sin pt$ and from $\frac{\pi}{p}$ to $\frac{2\pi}{p}$ by 0.

Therefore $$a_n = \frac{p}{\pi}\left[\int_0^{\frac{\pi}{p}} \hat{E} \sin pt \sin npt\, dt + \int_{\frac{\pi}{p}}^{\frac{2\pi}{p}} 0 \sin npt\, dt\right]$$

$$= \frac{p}{\pi}\int_0^{\frac{\pi}{p}} \hat{E} \sin pt \sin npt\, dt.$$

The particular coefficients a_1, a_2, etc., may be evaluated by replacing n in the above expression by 1, 2, etc. As a general rule, however, calculation is very much simplified by proceeding to the integrand before replacing n by its particular value. Thus by noting that $\sin\theta \sin n\theta = \dfrac{\cos(n-1)\theta - \cos(n+1)\theta}{2}$.

$$a_n = \frac{\hat{E}p}{2\pi}\int_0^{\frac{\pi}{p}} [\cos(n-1)pt - \cos(n+1)pt]\, dt$$

$$= \frac{Ep}{2\pi}\left[\frac{\sin(n-1)pt}{(n-1)p} - \frac{\sin(n+1)pt}{(n+1)p}\right]_0^{\frac{\pi}{p}}.$$

It is now only necessary to replace n by its particular value to obtain each coefficient.

Hence $$a_1 = \frac{\hat{E}p}{2\pi}\left[\frac{\sin 0pt}{0p} - \frac{\sin 2pt}{2p}\right]_0^{\frac{\pi}{p}}.$$

Since $\dfrac{\sin 0}{0}$ is indeterminate we must return to the original integral which then gives

$$a_1 = \frac{Ep}{\pi}\int_0^{\frac{\pi}{p}} \sin^2 pt\, dt = \frac{\hat{E}p}{2\pi}\int_0^{\frac{\pi}{p}} (1 - \cos 2pt)\, dt$$

$$= \frac{\hat{E}p}{2\pi}\left[t - \frac{\sin 2pt}{2p}\right]_0^{\frac{\pi}{p}}$$

$$= \frac{\hat{E}p}{2\pi}\cdot\frac{\pi}{p} = \frac{\hat{E}}{2}.$$

$$\left(\text{It is important to note that } \left[\frac{\sin 0pt}{0}\right]_0^{\frac{\alpha}{p}} = \alpha.\right)$$

All other sine function coefficients are zero because $\sin(n+1)pt$ and $\sin(n-1)pt$ are zero when $t = \dfrac{\pi}{p}$ or 0.

$$b_n = \frac{p}{\pi}\int_0^{\frac{\pi}{p}} \hat{E} \sin pt \cos npt\, dt$$

$$= \frac{\hat{E}p}{2\pi}\int_0^{\frac{\pi}{p}} (\sin(n+1)pt - \sin(n-1)pt)\, dt.$$

$$\left[\text{Note that } \sin\theta\cos n\theta = \frac{\sin(n+1)\theta - \sin(n-1)\theta}{2}\right]$$

or
$$b_n = \frac{\hat{E}p}{2\pi}\left[-\frac{\cos(n+1)pt}{(n+1)p} + \frac{\cos(n-1)pt}{(n-1)p}\right]_0^{\frac{\pi}{p}}$$
$$= \frac{\hat{E}}{2\pi}\left[\frac{1-\cos(n+1)\pi}{n+1} - \frac{1-\cos(n-1)\pi}{n-1}\right].$$

When n is odd, $\cos(n+1)\pi = \cos(n-1)\pi = 1$, hence b_1, b_3, b_5, etc., are zero.

When n is even $\cos(n+1)\pi = \cos(n-1)\pi = -1$, and

$$b_n = \frac{\hat{E}}{2\pi}\left[\frac{2}{n+1} - \frac{2}{n-1}\right] = \frac{\hat{E}}{\pi}\,\frac{-2}{(n-1)(n+1)}.$$

Thus
$$b_0 = \frac{\hat{E}}{\pi}.2,\quad b_2 = -\frac{\hat{E}}{\pi}.\frac{2}{3},\quad b_4 = -\frac{\hat{E}}{\pi}.\frac{2}{3.5}$$
$$b_6 = -\frac{\hat{E}}{\pi}.\frac{2}{5.7},\text{ etc.}$$

The expression for the half sine wave is

$$f(t) = \frac{2}{\pi}\left[\frac{1}{2} + \frac{\pi}{4}\sin pt - \frac{1}{3}\cos 2pt - \frac{1}{3.5}\cos 4pt - \frac{1}{5.7}\cos 6pt - \ldots\right]. \quad 2\text{A}.6a.$$

For a half cosine wave

$$f(t) = \frac{2}{\pi}\left[\frac{1}{2} + \frac{\pi}{4}\cos pt + \frac{1}{3}\cos 2pt - \frac{1}{3.5}\cos 4pt + \frac{1}{5.7}\cos 6pt \ldots\right] \quad . 2\text{A}.6b.$$

Any shape of wave may be so analysed and in computing the coefficients a_n and b_n for any discontinuous curve, the integrals may be split up into the appropriate functions and integrated between the limits of the discontinuities. For example, replacing the half-wave by a full-wave rectifier gives

$$a_n = \frac{p}{\pi}\left[\int_0^{\frac{\pi}{p}} \hat{E}\sin pt\sin npt\,dt + \int_{\frac{\pi}{p}}^{\frac{2\pi}{p}} -\hat{E}\sin pt\sin npt\,dt\right].$$

The second function contains $-\hat{E}\sin pt$ because the curve from $\frac{\pi}{p}$ to $\frac{2\pi}{p}$ is part of the sine curve $-\hat{E}\sin pt$. It is interesting to

note that the second integral is the negative of the first so that $a_n = 0$, for b_n we have the second integral equal to the first; hence

$$b_n = \frac{2p}{\pi}\int_0^{\frac{\pi}{p}} \hat{E} \sin pt \cos npt \, dt$$

and the expression for the full-wave rectifier-wave shape is

$$f(t) = \frac{4}{\pi}\left[\frac{1}{2} - \frac{1}{3}\cos 2pt - \frac{1}{3.5}\cos 4pt - \frac{1}{5.7}\cos 6pt - \ldots\right] \quad \text{2A.7}a$$

or for a full-wave rectifier with cosine input.

$$f(t) = \frac{4}{\pi}\left[\frac{1}{2} + \frac{1}{3}\cos 2pt - \frac{1}{3.5}\cos 4pt + \frac{1}{5.7}\cos 6pt - \ldots\right]. \quad \text{2A.7}b.$$

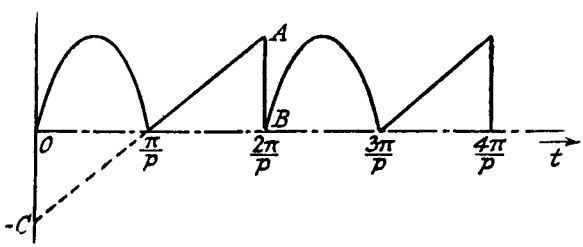


FIG. 2A.2.—A Combined Half Sine and Triangular Wave Shape.

If the wave has the shape shown in Fig. 2A.2, a half sine wave from 0 to $\frac{\pi}{p}$ and a triangular wave from $\frac{\pi}{p}$ to $\frac{2\pi}{p}$ the value of a_n is

$$a_n = \frac{p}{\pi}\left[\int_0^{\frac{\pi}{p}} \hat{E} \sin pt \sin npt \, dt + \int_{\frac{\pi}{p}}^{\frac{2\pi}{p}} (Kt - C) \sin npt \, dt\right]. \quad \text{2A.8}$$

where $(Kt - C)$ is the expression for the wave shape (a straight line) from $\frac{\pi}{p}$ to $\frac{2\pi}{p}$. K is the slope of the line and $-C$ the intercept at $t = 0$.

There is often no objection to changing the initial position of the curve as this only alters the relative initial time positions of the frequency components. In so doing the amount of analysis involved may be greatly reduced. In our first example (the half-wave rectifier) we could have moved the origin forward by $\frac{\pi}{2p}$ seconds, thus turning the function from a half sine wave to a half cosine wave. We would then have found the fundamental frequency component changed to a cosine function and all sine

components zero. Expression 2A.6*b* is the Fourier analysis for the changed position. In this connection it is useful to note the following six important cases of symmetry and their significance.

(1) $f(t) = -f(-t)$.

This condition is represented by the curve in Fig. 2A.3.

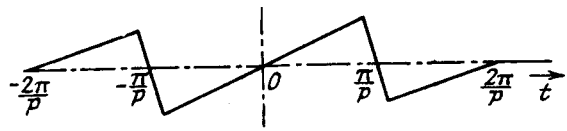


FIG. 2A.3.—A Wave Shape having Sine Functions only.

$$\text{Now } f(t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} a_n \sin npt + \sum_{n=1}^{\infty} b_n \cos npt$$

$$\text{and } f(-t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} a_n \sin np(-t) + \sum_{n=1}^{\infty} b_n \cos np(-t)$$

$$\text{thus } f(t)+f(-t) = 0 = b_0 + \sum_{n=1}^{\infty} a_n[(\sin npt + \sin(-npt)] + \sum_{n=1}^{\infty} b_n[\cos npt + \cos(-npt)]$$

$$\text{but} \qquad \sin(-npt) = -\sin npt.$$

$$\text{and} \qquad \cos(-npt) = \cos npt.$$

$$\therefore\ b_0 + 2\sum_{n=1}^{\infty} b_n \cos npt = 0$$

$$\text{or} \qquad f(t) = \sum_{n=1}^{\infty} a_n \sin npt \qquad . \quad . \quad . \quad \text{2A.9.}$$

Hence the curve contains only sine functions of the fundamental and harmonic frequencies.

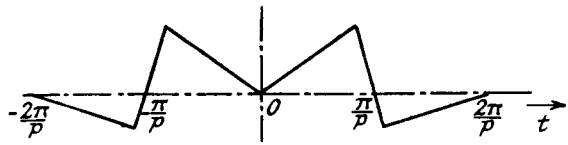


FIG. 2A.4.—A Wave Shape having Cosine Functions only.

(2) $f(t) = f(-t)$.

A curve such as that of Fig. 2A.4 fulfils this equation, which yields the following result.

$$\sum_{n=1}^{\infty} a_n[\sin npt - \sin np(-t)] + \sum_{n=1}^{\infty} b_n[\cos npt - \cos np(-t)] = 0.$$

Hence $$2\sum_{n=1}^{\infty} a_n \sin npt = 0$$

and $$f(t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} b_n \cos npt \quad . \quad . \quad . \quad 2A.10.$$

Moving the origin of the half wave in the first example produces a wave fulfilling this condition and reduces considerably the extent of the analytical investigation.

(3) $f(t) = -f\left(\frac{\pi}{p} + t\right)$ gives

$$f(t) = \sum_{n=1}^{\infty} a_{2n-1} \sin (2n-1)pt + \sum_{n=1}^{\infty} b_{2n-1} \cos (2n-1)pt \quad . \quad 2A.11$$

i.e., b_0 and all even integer sine and cosine coefficients such as a_2, b_2, etc., are zero. A typical curve is shown in Fig. 2A.5.

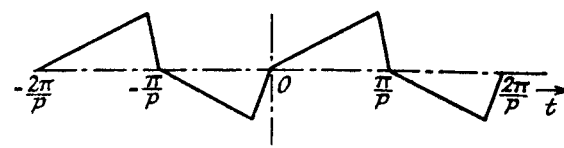


FIG. 2A.5.—A Wave Shape with Odd Sine and Cosine Functions.

(4) $f(t) = f\left(\frac{\pi}{p} + t\right)$.

The expression for the wave shape, which is illustrated in Fig. 2A.6, is

$$f(t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} a_{2n} \sin 2npt + \sum_{n=1}^{\infty} b_{2n} \cos 2npt \quad . \quad 2A.12.$$

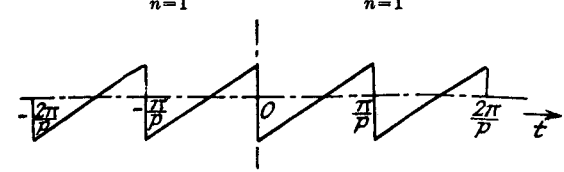


FIG. 2A.6.—A Wave Shape with Even Sine and Cosine Functions.

This is to be expected since the wave repeats itself after every $\frac{\pi}{p}$ interval shown and this interval should really be considered as the complete instead of half interval, viz., $\frac{2\pi}{p}$ instead of $\frac{\pi}{p}$.

(5) $f(t) = -f\left(\frac{\pi}{p} - t\right)$.

A wave shape conforming to this condition (see Fig. 2A.7) gives

$$f(t) = \sum_{n=1}^{\infty} a_{2n} \sin 2npt + \sum_{n=1}^{\infty} b_{2n-1} \cos (2n - 1)pt \quad . \quad 2A.13,$$

i.e., a_1, a_3, etc., b_0, b_2, b_4, etc., are zero.

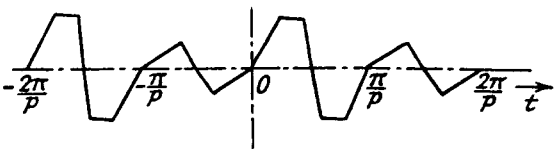


FIG. 2A.7.—A Wave Shape with Even Sine and Odd Cosine Functions.

(6) $f(t) = +f\left(\frac{\pi}{p} - t\right)$.

The curve of Fig. 2A.8 illustrates this and its expression is

$$f(t) = \frac{b_0}{2} + \sum_{n=1}^{\infty} a_{2n-1} \sin (2n - 1)pt + \sum_{n=1}^{\infty} b_{2n} \cos 2npt \quad . \quad 2a.14.$$

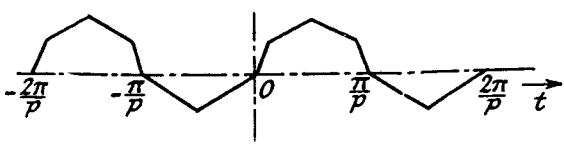


FIG. 2A.8.—A Wave Shape with Odd Sine and Even Cosine Functions.

The special conditions of symmetry are for convenience tabulated below.

Condition.	b_0	Sine coefficients.	Cosine coefficients.
1. $f(t) = -f(-t)$	Zero	All present	All zero
2. $f(t) = f(-t)$	Present	All zero	All present
3. $f(t) = -f\left(\frac{\pi}{p} + t\right)$	Zero	Even integers zero	Even integers zero
4. $f(t) = f\left(\frac{\pi}{p} + t\right)$	Present	Odd integers zero	Odd integers zero
5. $f(t) = -f\left(\frac{\pi}{p} - t\right)$	Zero	Odd integers zero	Even integers zero
6. $f(t) = f\left(\frac{\pi}{p} - t\right)$	Present	Even integers zero	Odd integers zero.

If a wave shape satisfies more than one of the conditions listed above, then $f(t)$ does not comprise any of the terms eliminated.

Taking as an example the curve in Fig. 2A.6, we find that it satisfies

$$f(t) = -f(-t)$$

and

$$f(t) = f\left(\frac{\pi}{p}+t\right).$$

$$\therefore\ f(t) = \sum_{n=1}^{\infty} a_{2n} \sin 2npt \quad . \quad . \quad . \quad \text{2A.15}$$

i.e., the curve contains only sine term coefficients of even suffix.

As a final example of Fourier analysis let us find the distortion present in the output current of a valve having a curved I_aE_g characteristic and an undistorted sinusoidal input voltage. Fig. 2A.9 illustrates a possible set of conditions.

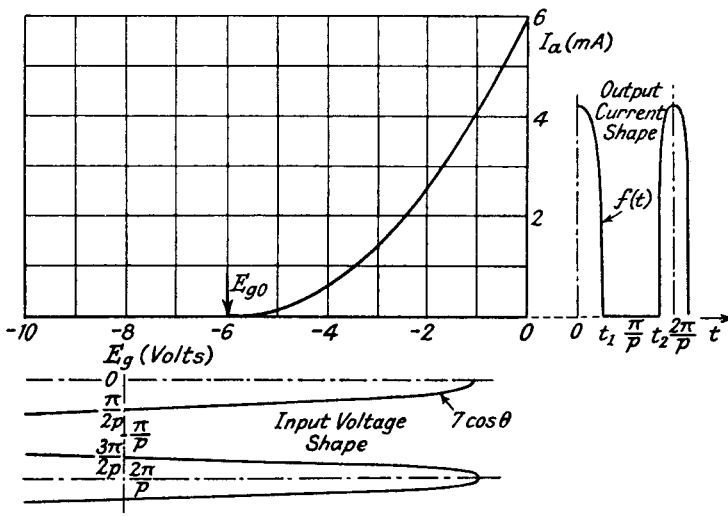


FIG. 2A.9.—An Example of the Application of Fourier Analysis to the Calculation of Distortion produced by the I_aE_g Characteristic of a Valve.

The first step must be to find the equation to the I_aE_g characteristic and this may be achieved by assuming that

$$I_a = c_0+c_1E_g+c_2E_g{}^2+ \ . \ . \ . \ +c_nE_g{}^n \quad . \quad . \quad \text{2A.16.}$$

The curve represented by expression 2A.16 may be made to coincide with the original curve at $(n+1)$ points; $(n+1)$ simultaneous equations are obtained by replacing the particular values of I_a and E_g at these points so that the coefficients c_0, c_1, etc., may be found. Generally it is unnecessary to proceed beyond $n = 4$, and in our particular example we need not consider above $n = 2$. Thus the expression for the I_aE_g curve of Fig. 2A.9 is

$$I_a = 6+2E_g+0{\cdot}166\ E_g{}^2 \quad . \quad . \quad . \quad \text{2A.17}$$

anode current cut-off occurring at $E_{g_0} = -6$. Let us suppose that the input voltage is represented by $7 \cos \theta$ and that the initial bias, $-E_b$, is -8 volts. The output current wave may be drawn out on its time axis to the right of the I_aE_g curve as shown.

The next step is to find the equations to the curve between the discontinuities and the time instants corresponding to the discontinuities. The time instant t_1 is easily determined because $\hat{E} \cos pt_1 = E_b + E_{g_0} = 8 - 6 = 2$.

$$\therefore\ t_1 = \frac{\cos^{-1}\dfrac{2}{\hat{E}}}{p} = \frac{\cos^{-1}\dfrac{2}{7}}{p} = \frac{1{\cdot}28}{p} \text{ radians}$$

and $$t_2 = \frac{2\pi}{p} - t_1 = \frac{5}{p} \text{ radians.}$$

Between t_1 and t_2 the equation to the curve is $f(t) = 0$ so that the Fourier coefficients are

$$b_n = \frac{p}{\pi}\int_0^{\frac{1{\cdot}28}{p}} f(t) \cos npt\, dt + \frac{p}{\pi}\int_{\frac{5}{p}}^{\frac{2\pi}{p}} f(t) \cos npt\, dt$$

$$= \frac{2p}{\pi}\int_0^{\frac{1{\cdot}28}{p}} f(t) \cos npt\, dt$$

where $f(t)$ is the equation to the curve between $t = 0$ and $t = t_1$. a_n is zero because $f(t) = f(-t)$ (see condition 2 above).

The expression for $f(t)$ is obtained by replacing E_g in 2A.17 by $\hat{E} \cos pt - E_b$, i.e., $7 \cos pt - 8$.

Thus $$f(t) = 4{\cdot}75 - 4{\cdot}66 \cos pt + 4{\cdot}083 \cos 2pt.$$

The next step involves finding a general expression for b_n. An attempt to obtain a formula containing only the variable n should always be made as it greatly simplifies the analysis. It can be achieved in this case by separating b_n into its three components, thus :

$$b_n = \frac{2p}{\pi}\Bigg[4{\cdot}75\int_0^{\frac{1{\cdot}28}{p}} \cos npt\, dt - 4{\cdot}66\int_0^{\frac{1{\cdot}28}{p}} \cos pt \cos npt\, dt + 4{\cdot}083\int_0^{\frac{1{\cdot}28}{p}} \cos 2pt \cos npt\, dt\Bigg]$$

$$= \frac{2}{\pi}\Bigg[\frac{4{\cdot}75 \sin 1{\cdot}28n}{n} - \frac{4{\cdot}66}{2}\left(\frac{\sin 1{\cdot}28(n+1)}{n+1} + \frac{\sin 1{\cdot}28(n-1)}{n-1}\right) + \frac{4{\cdot}083}{2}\left(\frac{\sin 1{\cdot}28(n+2)}{n+2} + \frac{\sin 1{\cdot}28(n-2)}{n-2}\right)\Bigg]$$

Where replacing n by a particular value gives $\frac{\sin 1{\cdot}28 \times 0}{0}$, this function becomes 1·28, thus

$$b_0 = 1{\cdot}74,\ b_1 = 1{\cdot}56,\ b_2 = 1{\cdot}09,\ b_3 = 0{\cdot}56,\ b_4 = 0{\cdot}16.$$

The output current therefore consists of the following components :

$$I_a = 0{\cdot}87 + 1{\cdot}56 \cos pt + 1{\cdot}09 \cos 2pt + 0{\cdot}56 \cos 3pt + 0{\cdot}16 \cos 4pt.$$

INDEX

END OF PART ONE.