RADIO

KEITH HENNEY

AVERAGE CHARACTERISTICS—RECEIVING TUBES

AMPLIFIERS, DETECTORS, OSCILLATORS, ETC.

Tube Type	Purpose	Cathode Type	Filament Volts	Filament Amp	Plate Volts	Plate Ma	Screen Volts	Grid Volts	Plate Resistance, Ohms	Transconductance	Mu	Power Output, Watts	Load Resistance, Ohms
1C5	Power amp. pentode	Filament	1.4	0.1	90	7.5	90	− 7.5	115,000	1550	.....	0.240	8000
1S4	Power amp. pentode	Filament	1.4	0.1	45	3.8	45	− 4.5	0.1 meg.	1250	.....	0.065	8000
2A3	Power amp. triode	Filament	2.5	2.5	250	60	...	−45	800	5250	4.2	3.5	2500
6C5	General purpose	Heater	6.3	0.3	250	8	...	− 8.0	10,000	2000	20	.....	....
6F6	Power amp. pentode	Heater	6.3	0.7	250	35	250	−16.5	80,000	2500	.....	3.2	7000
6K7	Super-control amp.	Heater	6.3	0.3	250	7	100	− 3.0	0.8 meg.	1450	.....	.....	....
6L6	Beam power tube	Heater	6.3	0.9	250	76	250	−14.0	22,500	6000	.....	6.5	2500
6SF5	High-mu triode	Heater	6.3	0.3	250	0.9	...	− 2.0	66,000	1500	100	.....	....
6S7	Super-control amp.	Heater	6.3	0.15	250	8.5	100	− 3.0	1 meg.	1750	.....	.....	....
6T7	Diode-triode	Heater	6.3	0.15	250	1.2	...	− 3.0	62,000	1050	65	.....	....
35L6	Beam power tube	Heater	35	0.15	110	40	110	− 7.5	13,800	5800	.....	1.5	2500

RECTIFYING TUBES

Tube Type	Purpose	Filament Volts	Amp	Typical Operating Conditions: A-c voltage per plate, rms	Typical Operating Conditions: D-c output ma
1-V	Half-wave	6.3	0.3	325	45
5U4-G 5Z3 5X4-G	Full-wave	5.0	3	450	225
5V4-g	Full-wave	5.0	2	500	175
5Y3-G 80	Full-wave	5.0	2	500	125
12Z3	Half-wave	12.6	0.3	235	55
25Z5	Rectifier-doubler	25.0	0.3	235	75 per plate

CONVERTER TUBES

Type	Filament Volts	Filament Amp	Plate Volts	Plate Ma	Screen Volts	Screen Ma	Control Grid, Volts	Plate Resistance, Megohms	Conversion Conductance, Micromhos
1A7-G	1.4	0.05	90	0.055	45	0.6	0	0.6	250
1C7-G	2.0	0.12	135	1.3	67.5	2.5	−3	0.6	300
1D7-G	2.0	0.06	135	1.2	67.5	2.5	−3	0.4	275
1R5	1.4	0.05	45	0.7	45	1.9	0	0.6	235
6A8	6.3	0.3	250	3.5	100	2.7	−3	0.36	550
6D8-G	6.3	0.15	250	3.5	100	2.6	−3	0.4	550
6K8	6.3	0.3	250 *	2.5	100	6.0	−3	0.6	350
			100 †	3.8					
6SA7	6.3	0.3	100	3.3	100	8.5	0	0.5	425

* Mixer plate.
† Oscillator plate.

Trigonometric Functions

Angle α	Sin α	Cos α	Tan α	Angle α	Sin α	Cos α	Tan α
0	.000	1.000	.000				
1	.017	1.000	.017	46	.719	.695	1.04
2	.035	.999	.035	47	.731	.682	1.07
3	.052	.999	.052	48	.743	.669	1.11
4	.070	.998	.070	49	.755	.656	1.15
5	.087	.996	.088	50	.766	.643	1.19
6	.105	.995	.105	51	.777	.629	1.23
7	.122	.993	.123	52	.788	.616	1.28
8	.139	.990	.141	53	.799	.602	1.33
9	.156	.988	.158	54	.809	.588	1.38
10	.174	.985	.176	55	.819	.574	1.43
11	.191	.982	.194	56	.829	.559	1.48
12	.208	.978	.213	57	.839	.545	1.54
13	.225	.974	.231	58	.848	.530	1.60
14	.242	.970	.249	59	.857	.515	1.66
15	.259	.966	.268	60	.866	.500	1.73
16	.276	.961	.287	61	.875	.485	1.80
17	.292	.956	.306	62	.883	.469	1.88
18	.309	.951	.325	63	.891	.454	1.96
19	.326	.946	.344	64	.899	.438	2.05
20	.342	.940	.364	65	.906	.423	2.14
21	.358	.934	.384	66	.914	.407	2.25
22	.375	.927	.404	67	.920	.391	2.36
23	.391	.920	.424	68	.927	.375	2.48
24	.407	.914	.445	69	.934	.358	2.61
25	.423	.906	.466	70	.940	.342	2.75
26	.438	.899	.488	71	.946	.326	2.90
27	.454	.891	.510	72	.951	.309	3.08
28	.469	.883	.532	73	.956	.292	3.27
29	.485	.875	.554	74	.961	.276	3.49
30	.500	.866	.577	75	.966	.259	3.73
31	.515	.857	.601	76	.970	.242	4.01
32	.530	.848	.625	77	.974	.225	4.33
33	.545	.839	.649	78	.978	.208	4.70
34	.559	.829	.675	79	.982	.191	5.14
35	.574	.819	.700	80	.985	.174	5.67
36	.588	.809	.727	81	.988	.156	6.31
37	.602	.799	.754	82	.990	.139	7.12
38	.616	.788	.781	83	.993	.122	8.14
39	.629	.777	.810	84	.995	.105	9.51
40	.643	.766	.839	85	.996	.087	11.43
41	.656	.755	.869	86	.998	.070	14.30
42	.669	.743	.900	87	.999	.052	19.08
43	.682	.731	.933	88	.999	.035	28.64
44	.695	.719	.966	89	1.000	.017	57.29
45	.707	.707	1.000	90	1.000	.000	Infinity

Reprinted by permission from Timbie, Basic Electricity for Communications.

Relation between Wave Length in Meters, Frequency in Kilocycles, and the Product of Inductance (in Microhenries) and Capacity (in Microfarads)

Meters	f in Kc.	$L \times C$	Meters	f in Kc.	$L \times C$	Meters	f in Kc.	$L \times C$
1	300,000	0.0000003	450	667	0.0570	740	405	0.1541
2	150,000	0.0000011	460	652	0.0596	745	403	0.1562
3	100,000	0.0000018	470	639	0.0622	750	400	0.1583
4	75,000	0.0000045	480	625	0.0649	755	397	0.1604
5	60,000	0.0000057	490	612	0.0676	760	395	0.1626
6	50,000	0.0000101	500	600	0.0704	765	392	0.1647
7	42,900	0.0000138	505	594	0.0718	770	390	0.1669
8	37,500	0.0000180	510	588	0.0732	775	387	0.1690
9	33,333	0.0000228	515	583	0.0747	780	385	0.1712
10	30,000	0.0000282	520	577	0.0761	785	382	0.1734
20	15,000	0.0001129	525	572	0.0776	790	380	0.1756
30	10,000	0.0002530	530	566	0.0791	795	377	0.1779
40	7,500	0.0004500	535	561	0.0806	800	375	0.1801
50	6,000	0.0007040	540	556	0.0821	805	373	0.1824
60	5,000	0.0010140	545	551	0.0836	810	370	0.1847
70	4,290	0.0013780	550	546	0.0852	815	368	0.1870
80	3,750	0.0018010	555	541	0.0867	820	366	0.1893
90	3,333	0.0022800	560	536	0.0883	825	364	0.1916
100	3,000	0.00282	565	531	0.0899	830	361	0.1939
110	2,727	0.00341	570	527	0.0915	835	359	0.1962
120	2,500	0.00405	575	522	0.0931	840	357	0.1986
130	2,308	0.00476	580	517	0.0947	845	355	0.201
140	2,143	0.00552	585	513	0.0963	850	353	0.203
150	2,000	0.00633	590	509	0.0980	855	351	0.206
160	1,875	0.00721	595	504	0.0996	860	349	0.208
170	1,764	0.00813	600	500	0.1013	865	347	0.211
180	1,667	0.00912	605	496	0.1030	870	345	0.213
190	1,579	0.01015	610	492	0.1047	875	343	0.216
200	1,500	0.01126	615	488	0.1065	880	341	0.218
210	1,429	0.01241	620	484	0.1082	885	339	0.220
220	1,364	0.01362	625	480	0.1100	890	337	0.223
230	1,304	0.01489	630	476	0.1117	895	335	0.225
240	1,250	0.01621	635	472	0.1135	900	333	0.228
250	1,200	0.01759	640	469	0.1153	905	331	0.231
260	1,154	0.01903	645	465	0.1171	910	330	0.233
270	1,111	0.0205	650	462	0.1189	915	328	0.236
280	1,071	0.0221	655	458	0.1208	920	326	0.238
290	1,034	0.0237	660	455	0.1226	925	324	0.241
300	1,000	0.0253	665	451	0.1245	930	323	0.243
310	968	0.0270	670	448	0.1264	935	321	0.246
320	938	0.0288	675	444	0.1283	940	319	0.249
330	909	0.0306	680	441	0.1302	945	317	0.251
340	883	0.0325	685	438	0.1321	950	316	0.254
350	857	0.0345	690	435	0.1340	955	314	0.257
360	834	0.0365	695	432	0.1360	960	313	0.259
370	811	0.0385	700	429	0.1379	965	311	0.262
380	790	0.0406	705	426	0.1399	970	309	0.265
390	769	0.0428	710	423	0.1419	975	308	0.268
400	750	0.0450	715	420	0.1439	980	306	0.270
410	732	0.0473	720	417	0.1459	985	305	0.273
420	715	0.0496	725	414	0.1479	990	303	0.276
430	698	0.0520	730	411	0.1500	995	302	0.279
440	682	0.0545	735	408	0.1521	1000	300	0.282

PRINCIPLES OF RADIO

KEITH HENNEY

Editor . . . . Electronics

Fellow . . . . Institute of Radio Engineers

FIFTH EDITION

Third Printing

WITH A CHAPTER ON RADAR

New York—JOHN WILEY & SONS, INC.

London—CHAPMAN & HALL, LIMITED

FIFTH EDITION
Third Printing, July, 1946

PRINTED IN THE UNITED STATES OF AMERICA

PRINCIPLES OF

RADIO

PREFACE

It has now been fifteen years since the first printing of this elementary textbook on radio appeared. The book was originally written for the student who had little background in radio upon which to build and yet who wanted to know the basis upon which radio communication existed. To make it useful to a man who had to study without benefit of a teacher, every attempt was made to keep the text clear and practical. This viewpoint has been maintained through subsequent editions. In the belief that problems are of tremendous aid in enabling the reader to make sure that he has really understood the text, many of the old problems have been retained and many new ones added.

Much new text material has been added. As always, the radio art forges ahead into new and adventurous ground. Wave guides, microwaves, Klystron tubes, frequency-modulation broadcasting, measuring distances by radio, transients—some of these matters are new and had not been heard of when the first edition first appeared; others are old in the art but now assume fresh importance.

I wish to acknowledge the guidance, in preparing this Fifth Edition, given me by Hollis Baird, F. E. Christianson, and H. S. Renne. Each of these teachers of pre-radar and radio courses went over previous editions and made many very helpful suggestions for improvement.

KEITH HENNEY

January, 1945

CONTENTS

CHAPTER 1

FUNDAMENTALS

It is essential for anyone studying radio engineering to know something about electronics, for radio apparatus of today is built almost entirely around the electron tube. Consequently, a question of fundamental importance is: what is the electron?

Electrons. Since the earliest days, man has tried to find out what his universe is made of; to find out the smallest possible unit out of which matter is made. He first discovered that there were very few different substances (92), and that these basic **elements** existed in the form of molecules. Then he found that molecules could be broken down into smaller units called **atoms.** He discovered that all known substances could be made out of these 92 elements, which differed from one another in chemical, physical, and electrical characteristics.

But it was a long time before man was able to find out what the individual atoms were made of. He later discovered that the simplest atom, hydrogen, was comprised of two fundamental particles, one of which was very heavy and carried positive electricity, and the other, exceedingly small and light in weight, which carried negative electricity. The positive particle is called a **proton;** the negative particle, an **electron.** The electron is 1800 times lighter than the simplest atom, hydrogen. From this it is known that the positive portion of the hydrogen atom carries almost the entire weight. The same is true of all atoms; the core or nucleus bears the weight or mass and contains positive electricity in sufficient quantity to balance the negative electricity carried by one or more electrons which move about the nucleus. Other basic particles, with and without electrical charge, are found in the nucleus, too, but the nucleus is essentially positive.

Atoms of the 92 elements (hydrogen, oxygen, iron, sulphur, gold, etc.) differ from one another only in the number of electrons and the make-up of the nucleus.

Atoms. The atoms of matter are inconceivably small. Oil falling on a wet street tends to spread out until the oil film may be only half a ten-millionth of an inch thick. Yet atoms of which the oil film is made must be smaller than this, for there is no way to spread the film so that it is only half (let us say) an atom thick.

Radio men are, in general, concerned with two kinds of substances: (1) conductors, such as metals; and (2) insulators, such as glass or mica. In the metals, the atoms are rather closely packed together. The electrons of these atoms are arranged in elliptical orbits about the nucleus. The electrons in the outermost orbits are not so tightly held to the nucleus and can be detached from it. In fact, the atoms of a metal are continually exchanging electrons, so that it is possible for an individual electron to travel from atom to atom, from one end of a wire to the other, if proper conditions exist.

Electrical charges. A particle, or an atom, carrying equal amounts of positive and negative electricity is said to be in equilibrium and tends to stay in this condition. It is possible, however, to make the particle lose or gain one or more electrons, and then it is said to be **charged** or **ionized.** If a particle has too few electrons, it has too little negative electricity to balance its positive charge and will pick up enough electrons to get back into equilibrium at the first opportunity. A particle lacking electrons is said to be positively charged. On the other hand, if it has too many electrons it is negatively charged and will tend to lose an electron when it gets a chance. A charged body in a liquid or gas is called an **ion.**

Two electrical charges which are alike, that is, two positive or two negative charges, repel each other; two unlike charges, a positive and a negative, attract each other. The force tending to make the charges move toward or away from each other depends upon the amount of electricity the particles carry and varies inversely with the square of the distance between the particles.* Within the atom itself, where the distances between electrons or between electrons and nucleus are exceedingly small, the attractive force holding the atom together may be very great.

If one could collect and place a kilogram of electrons at each of the poles of the earth (8000 miles apart), the two masses of electrons would repel each other with a force of about 2 million tons.

Electrons in motion. The electric current. Electrons, therefore, together with their positive counterparts, the protons, which exist in the nucleus, are the basic building blocks of the atoms. Everything is made up of these elementary blocks. The electron is essentially an electrical charge which we arbitrarily call negative. It will be attracted to any positively charged body, and when it moves it carries its negative charge with it. If sufficient electrons move from one place to another, a definite "flow" of electricity takes place.

Note that the word flow does not indicate that electricity is something like water. For many years this was the belief, and so the word flow has come to be used for the transport of charges from one place to another. We say that an electric current flows. Actually, electric energy can be made to move from one place to another with no real transfer of electrons between these places. Radio communication is an example of such a transfer of energy. An electric current, whether passing through a wire or through the vacuum of an electron tube, is merely a movement of electric charges.

In a piece of copper wire the electrons of the individual atoms are moving about in a haphazard fashion at the speed of approximately 35 miles per second. As long as the wire is not attached to a source of electrical charge, the electrons have no preference as to the direction in which they move, and in the wire there is no tendency for the electrons to move from one end

* The mathematical expression for this force is known as **Coulomb's law:**

$$F = \frac{q_1 q_2}{d^2}$$

where F is the force in dynes, q_1 and q_2 are the respective charges in coulombs, and d is the distance in centimeters between the particles.

to the other. If, however, the wire is attached to a battery and the circuit is completed, then the electrons will move along the wire, each transporting its elemental quantity of electricity. This quantity is so small that engineers are generally concerned with the motion of electrons only when large numbers of them move from one place to another. It has been estimated that all the inhabitants of the earth, counting day and night at the highest rate of speed possible, would need two years to count the number of electrons which pass through an ordinary electric light in a second.

Units. Just as a carpenter orders a board so many feet long, if he is in this country, or so many meters long if he is in France, the electrical man must have units for the quantities he uses. The amount of electricity each electron carries is now well established; it is exceedingly small. The unit of quantity is known as the **coulomb.** There are 6.28 million million million electrons in 1 coulomb of electricity. If this number of electrons moves through an electric lamp each second, 1 **ampere** of current will flow, and if this lamp is operated from ordinary house-wiring circuits, 115 watts of energy will be used in it.

Here, then, are two units: the coulomb representing a *quantity* of electricity, and the ampere representing the *rate of flow* of this quantity. The coulomb corresponds to the gallon in fluid measure, and the ampere to gallons per minute. A current of 1 ampere means a rate of flow of 1 coulomb of electricity per second.

The current flowing through a radio receiver tube is of the order of a few thousands of an ampere. A small town power plant may have thousands of amperes flowing from its generators. Approximate currents flowing through some common devices are shown below.

Apparatus	Approximate Current in Amperes
50-watt lamp	0.5
250-watt lamp	2.5
1-hp motor	10.
Electric iron	5.
Filament of battery-type radio tube	0.05
Plate circuit of electron tube	0.005

A meter to measure the flow of electricity is called an **ammeter.**

Engineers' shorthand. Engineers have a simple shorthand method of working with large numbers. For example, the number 6.28 million million million is expressed as 6.28×10^{18}. This many electrons flowing past a given point per second constitutes the electric current known as 1 ampere. As we shall have occasion to use this shorthand system many times in this book, the reader is encouraged to master it as soon as possible. The table below will be helpful.

$$1 = 10^0 = \text{one}$$

$$10 = 10^1 = \text{ten}$$

$$100 = 10^2 = \text{hundred}$$

$$1000 = 10^3 = \text{thousand, etc.}$$

$$1 = 10^0 = \text{one}$$

$$0.1 = 10^{-1} = \tfrac{1}{10} = \text{one-tenth}$$

$$0.01 = 10^{-2} = \tfrac{1}{100} = \text{one-hundredth}$$

$$0.001 = 10^{-3} = \tfrac{1}{1000} = \text{one-thousandth, etc.}$$

The small number above the figure 10 is called the **exponent.** Numbers less than 1 have negative exponents. Thus three-thousandths may be expressed in these several ways:

$$0.003 = 3 \times 10^{-3} = \frac{3}{1000} = \frac{3}{10^3}$$

When numbers are multiplied, their exponents are added; when the numbers are divided, the exponents are subtracted. Thus 100 multiplied by four-tenths may be done in shorthand as follows:

$$\begin{aligned} 100 \times 0.4 &= 10^2 \times 4 \times 10^{-1} \\ &= 4 \times 10^1 \\ &= 4 \times 10 \\ &= 40 \end{aligned}$$

Similarly, let us divide 3000 by 150.

$$3000 \div 150 = (3 \times 10^3) \div (1.5 \times 10^2)$$
$$= \frac{3}{1.5} \times 10^3 \times 10^{-2}$$
$$= 2 \times 10$$
$$= 20$$

The rules are few and simple:

1. To multiply, add exponents.
2. To divide, subtract exponents.
3. When any number crosses the line, change the sign of the exponent.

Example 1-1. Multiply 20,000 by 1200 and divide the result by 6000.

$$20{,}000 = 2 \times 10^4$$
$$1200 = 12 \times 10^2$$
$$6000 = 6 \times 10^3$$
$$\frac{20{,}000 \times 1200}{6000} = \frac{2 \times 10^4 \times 12 \times 10^2}{6 \times 10^3}$$
$$= \frac{2 \times 12 \times 10^4 \times 10^2 \times 10^{-3}}{6}$$
$$= \tfrac{24}{6} \times 10^3$$
$$= 4000$$

Problem 1-1. How many electrons flow past a given point per second when the number of amperes is 6? 60? 600? 0.1? 0.003?

Problem 2-1. The sun is roughly 90 million miles from the earth. Express this in "shorthand."

Problem 3-1. At 100 miles per hour, how many months would it take to reach the sun?

Problem 4-1. If light travels at 300 million meters per second and if a meter equals 3.3 ft, how long does it take the sun's rays to reach the earth?

Problem 5-1. How many amperes of current flow when 31.4×10^{15} electrons per second flow past a point?

In connection with such shorthand methods the following table of prefixes commonly used will be important.

PREFIX	SYMBOL	MEANING	SHORTHAND
micro	μ	one-millionth	10^{-6}
milli	m	one-thousandth	10^{-3}
centi	c	one-hundredth	10^{-2}
deci	d	one-tenth	10^{-1}
deka	dk	ten	10
hekto	h	one hundred	10^{2}
kilo	k	one thousand	10^{3}
mega	M	one million	10^{6}

Thus a thousandth of an ampere is known as a milliampere, a million ohms is called a megohm, etc.; or, expressed in numbers, 1 ma = 10^{-3} or 0.001 amp; 1 megohm = 1,000,000 ohms.

Example 2–1. How many milliamperes are there in 2 amperes? Since 1 milliampere is equal to one-thousandth ampere, 1 ampere is equal to 1000 milliamperes. Thus, 1 amp = 1000 ma = 10^3 ma. Therefore 2 amp is equal to 2×10^3 ma or 2000 ma.

How many amperes are in 2 ma? Here one must remember that there are fewer amperes in 1 ma than there are milliamperes in 1 amp because the ampere is the larger unit. Thus, 1 ma = 0.001 amp = 10^{-3} amp. Therefore 2 ma = 2×10^{-3} amp.

Problem 6–1. How many cycles are 1000 kilocycles? How many kilocycles in 500 cycles?

Problem 7–1. How many megacycles are in 1000 kilocycles? How many cycles in 30 megacycles?

Curve plotting. Many radio problems can be solved without any mathematics at all if one understands the technique of plotting curves. A curve is a visual means of portraying what happens to one quantity when another is varied. For example, the curve in Fig. 1–1 shows the distance traveled by a train moving at a fixed rate of speed. One makes such a curve in the following manner. At zero time, the train is zero distance from its starting point. This is at the lower left-hand corner of the plot. Now let us represent time along the horizontal part of the plot, distance along the vertical. If the train goes 50 miles per hour, at the end of the first hour it will be 50 miles from the starting point (called the **origin** in curve plotting). At the end of the second hour it will be 2×50 or 100 miles away; at the end of the third hour it will be 150 miles away; and so on. All we need to do to make a plot of this kind is to put a mark at the vertical value corresponding to each horizontal value we may choose and then draw a smooth curve which best fits the points.

This curve is a visual picture of the position of the train for each portion of the time we may be interested in. The curve correlates two variables—time and distance.

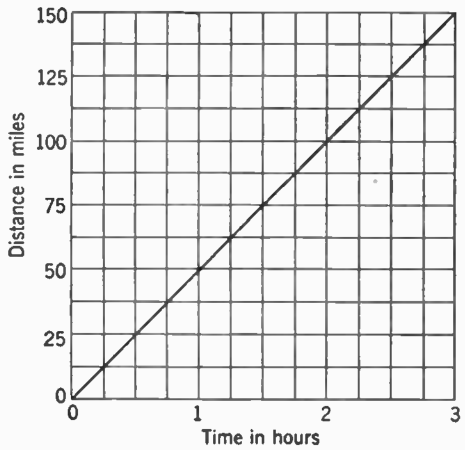


FIG. 1-1. A simple curve or graph showing the distance traveled by a train in any given time.

If the train travels at a uniform speed, the curve correlating distance and time will be a straight line. The slope of the curve, that is, a given vertical length of the curve divided by the corresponding horizontal length, expresses the rate of change of distance with respect to time. This quantity is stated in miles per hour and is known as velocity or speed.

The two factors in this simple graph are known as the variables, one being dependent upon the other. Here the independent variable is time, and the dependent variable is distance since the distance traveled depends upon the elapsed time.

A curve such as we have been discussing has two coordinates, horizontal and vertical, which represent the independent and dependent variables. These reference lines are called the **axes.** The vertical axis is often called the ordinate or Y-axis, and the horizontal the abscissa or X-axis. Horizontal distances to the left of the ordinate are negative; those to the right are positive. Similarly, vertical distances below the abscissa are negative and

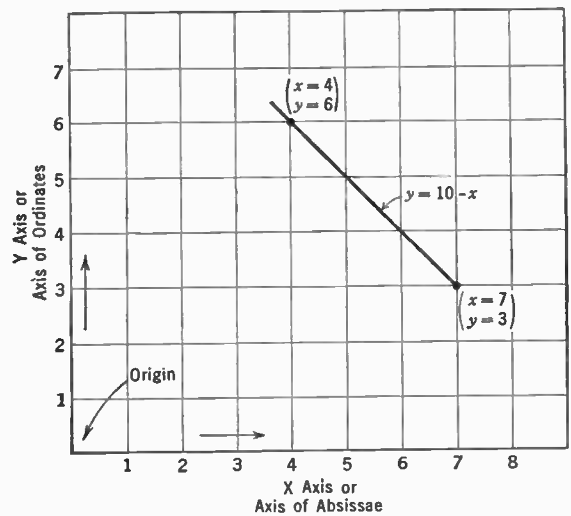


FIG. 2-1. Another type of curve. Here one variable, y, decreases as the other, x, increases. It might represent the amount of money one has as a function of the amount he spends.

those above are positive. Thus we can plot both positive and negative quantities on such a chart.

Slope. The change in vertical units with a given change in horizontal units is called the **slope** of the curve. This is an important factor since it shows the rate at which one quantity varies with respect to the other. The actual appearance of the curve will change depending upon the units employed, but the numerical value of the slope will not change. For example, if the vertical units are doubled in value the curve will appear flattened, and if they are halved it will appear steeper, but for both curves the actual slope as defined by the ratio between the vertical change for a given horizontal change will be the same. The manner in which the slope is calculated is shown in Fig. 3–1.

Problem 8–1. A power supply device has a terminal voltage of 150 when 130 ma of current is drawn from it. Other values of voltage and current are as follows: 200 volts at 80 ma; 250 volts at 30 ma. Plot a

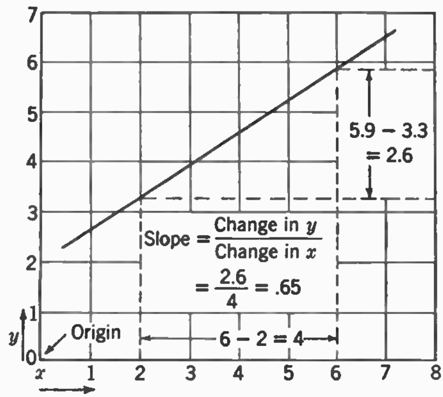


Fig. 3–1. How to calculate the slope, which is the rate at which y changes as x changes. In this case, note that y has a value even though x is zero.

curve showing these relations, and determine the change in voltage per milliampere change. For this curve the slope is negative since an increase in one of the variables causes a decrease in the other.

Symbols. In technical literature a number of abbreviations are used to represent parts of circuits. The symbols used in this

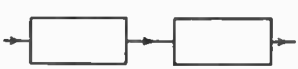


Fig. 4–1. Two devices connected in series.

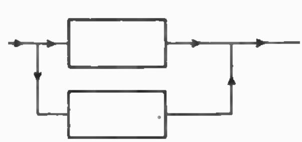
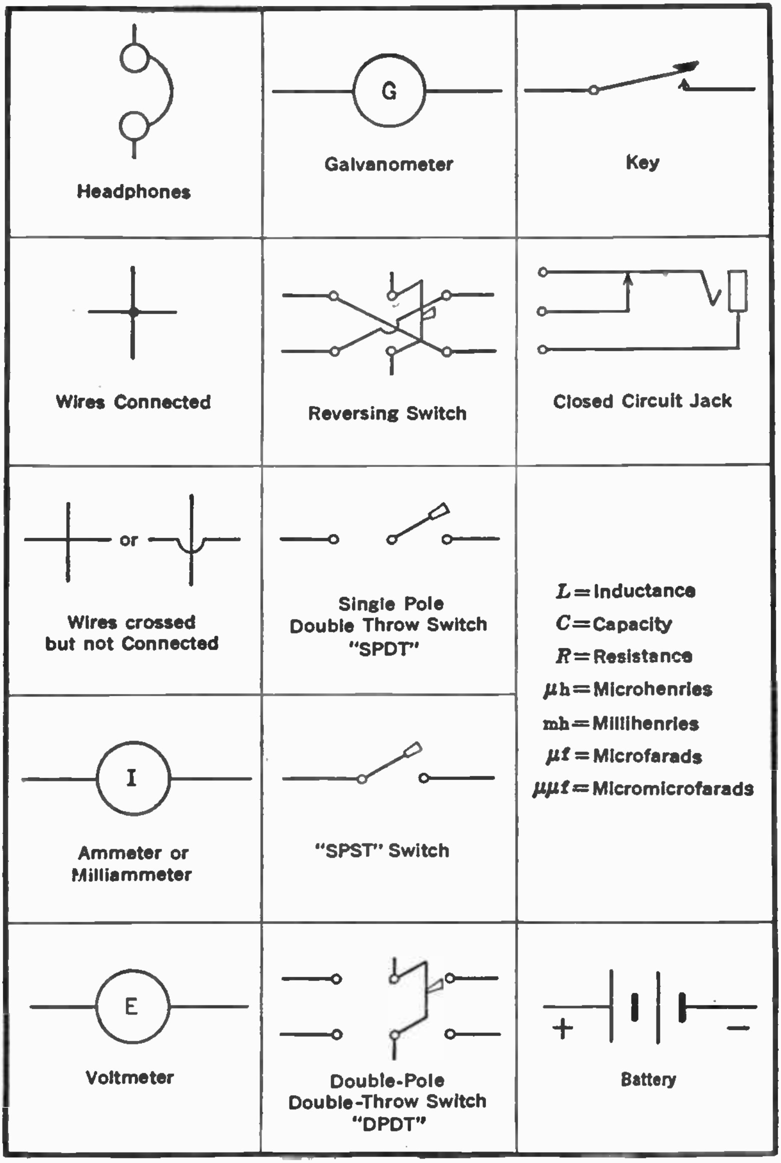
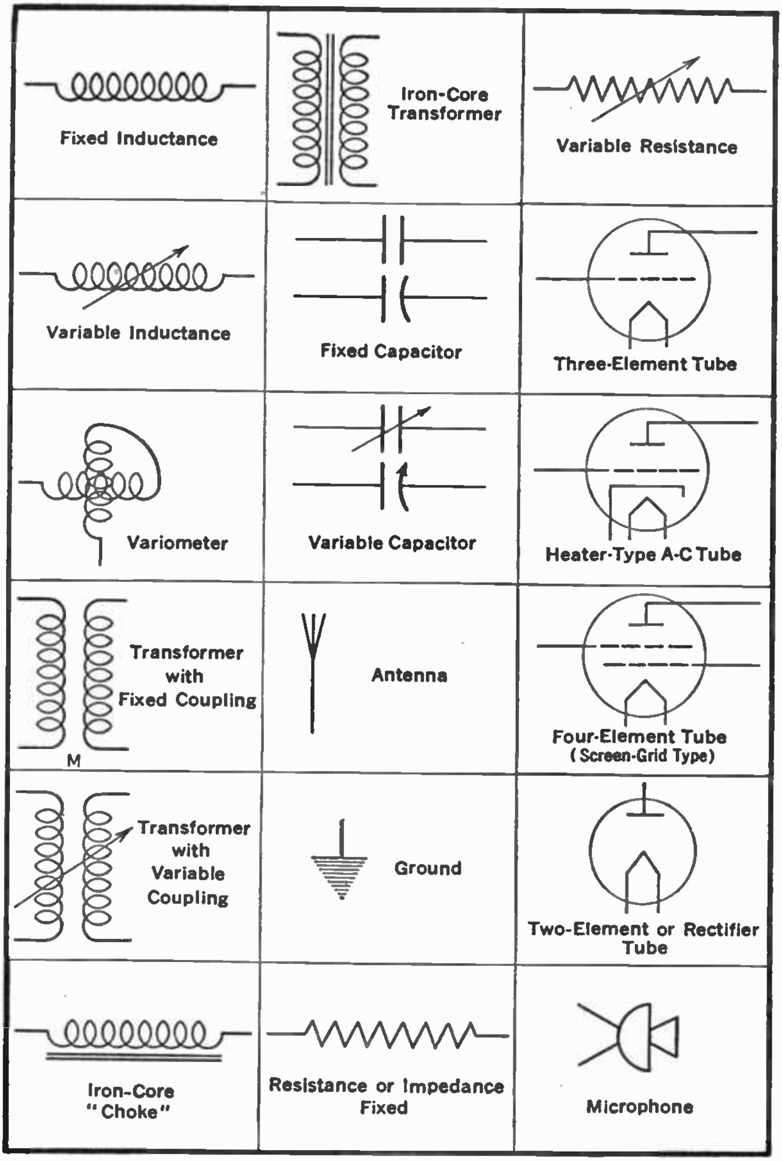


Fig. 5–1. Two devices connected in parallel or shunt.

book are shown on pages 11 and 12. A circuit is built up by connecting together several of these symbols as shown in Figs. 4–1 and 5–1 in one of two ways since each piece of apparatus used in radio circuits has at least two terminals.


Headphones
G
Galvanometer
Key
Wires Connected
Reversing Switch
Closed Circuit Jack
or
Wires crossed
but not Connected
Single Pole
Double Throw Switch
"SPDT"
L = Inductance
C = Capacity
R = Resistance
μh = Microhenries
mh = Millihenries
μf = Microfarads
$\mu\mu f$ = Micromicrofarads
I
Ammeter or
Milliammeter
"SPST" Switch
E
Voltmeter
Double-Pole
Double-Throw Switch
"DPDT"
+
−
Battery


Fixed Inductance
Iron-Core Transformer
Variable Resistance
Variable Inductance
Fixed Capacitor
Three-Element Tube
Variometer
Variable Capacitor
Heater-Type A-C Tube
Transformer with Fixed Coupling
M
Antenna
Four-Element Tube (Screen-Grid Type)
Transformer with Variable Coupling
Ground
Two-Element or Rectifier Tube
Iron-Core "Choke"
Resistance or Impedance Fixed
Microphone

CHAPTER 2

DIRECT-CURRENT CIRCUITS

In the previous chapter it was stated that an electric current was a motion of electrons. Now two questions naturally arise: what makes the electrons move, and in what direction does the electric current flow? Let us answer the second question first.

Direction of current flow. Long before engineers and scientists knew anything about the electron the convention was established that current flowed from the positive terminal of a battery, or other source, toward the negative terminal through the external circuit. Now, of course, we know that electrons move from a more negative point toward a less negative point, or toward a point of positive polarity. Therefore we realize that the plus-to-minus direction of current flow is merely a manner of speaking—that actually the electrons flow from negative to positive through the external circuit. Engineers usually maintain the convention that current flows from positive to negative while the electrons move in the opposite direction. It really makes little difference as long as one is consistent. One should, however, be able to think in terms of both systems.

The reader should remember that in a load circuit the point *from* which the electrons flow is always negative with respect to the point *toward* which they flow.

Electromotive force. Now what makes the electrons move from one place to another? Electrons are driven through the wires and apparatus composing an electrical circuit by a force called an **electromotive force** (emf). The unit of force is called the **volt.** An instrument used to measure voltage is known as a **voltmeter.** The table shows voltages of commonly used sources of emf.

Apparatus	Voltage (approximate)
Dry cell	1.5
Storage battery	6
B battery	45
House-lighting circuit	115
"Third rail"	500

Sources of emf. The oldest source of man-made voltage is friction. Anyone who has rubbed a cat's fur on a cold winter day (or who has worn silk clothes!) will remember the crackling noise and the tendency of the fur to stand up and follow the hand. A pocket comb rubbed on the coat will pick up bits of paper and other light insulated objects. These phenomena are manifestations of frictional electricity, which causes electrical charges to be removed from a substance, leaving it electrically charged. Machines for generating very high voltages have been devised on the friction principle. Charges are removed from a source and stored on an enormous sphere to be released all at one time when it is desired to use the accumulated charge.

Frictional electricity is not a very reliable source of emf since moisture is fatal to the continued production or storage of the necessary charges.

The most common source of voltage is the battery. The common dry cell used to operate door bells is one form of battery; the lead-sulphuric acid storage battery is another.

The generator is a means of converting mechanical energy into electrical energy. It is used to produce very large amounts of electrical energy; its principles will be described later. Another source of electrical energy is the thermocouple. This device is composed of two dissimilar metals joined together mechanically. When the junction is placed in a flame, an emf will be produced which is proportional to the temperature. Still another source of emf is the barrier type of photoelectric cell. In this device, light shines on a surface coated with material which emits electrons under stimulation of the illumination. A barrier photocell will produce 5 ma or more current directly from sunlight.

Resistance. The next question that arises is: how much current will flow?

The electrons in their motion through a conductor are not unimpeded. They constantly run into atoms and other electrons. Since there are no perfect conductors, all materials are said to have a certain **resistance.** This is the measure of the trouble electrons have in moving freely about among the atoms making up the material. Metals have less resistance than insulators; they are better conductors. Some metals have lower resistance than others. In general, pure metals have lower resistances than alloys.

Thus, copper has low resistance, whereas some of the combinations of copper, nickel, and iron-manganese, for example, have resistances many times that of copper. The fact that copper has such low resistance and at the same time is plentiful explains why most conductors are made of this element. Silver has still lower resistance than copper, but it is not so plentiful.

Factors that govern resistance. The comparative resistance of two wires of the same material and at the same temperature depends upon the length of the wires and the area of their cross sections. Naturally, the longer the wire the fewer electrons can pass through it in a given time; similarly, the smaller the diameter of a wire the greater the resistance. You can get more gallons of water per second from a 3-in. fire hose than from a 1-in. garden hose, although they may be attached to the same hydrant.

A wire 2 ft long has twice the resistance of a wire 1 ft long but of the same diameter. Of two wires the same length, the one having the smaller diameter will have the greater resistance. The resistance is inversely proportional to the area of the wire or to the square of the wire diameter. The copper wire table on page 19 shows that a No. 10 wire has a diameter of 102 mils and a resistance of approximately 1 ohm per 1000 ft, whereas No. 16 wire, with one-half the diameter, has four times the resistance.

The absolute value of the resistivity of a substance may be indicated in several ways. The most useful to electrical engineers, since they use so much of their resistance material in the form of wires, is the **ohm per mil-foot.** This is the resistance in ohms of a wire 1 mil in diameter and 1 ft long. A mil is a thousandth of an inch (0.001 in.). A circular mil is a unit of area. A wire having a diameter of d mils will have an area of

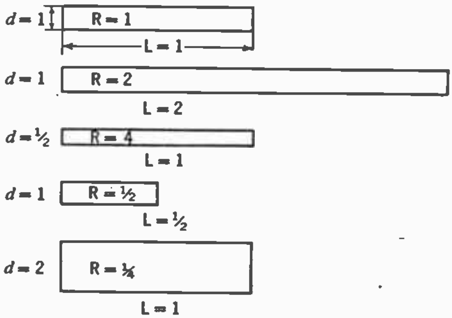


FIG. 1-2. Resistance depends upon the length and the diameter of a conductor.

d^2 circular mils. A mil-foot of copper (1 circular mil in cross section area and 1 ft long) will have a resistance of 10.4 ohms. The resistance of any copper wire therefore will be $10.4L \div A$ ohms or $10.4L \div d^2$ ohms, where A is area in circular mils, L is length in feet, and d is diameter in mils. Tables showing the resistivity of many materials will be found in handbooks used by electrical and radio engineers. In general, however, wire tables showing the actual resistance, in ohms, of wire of various sizes are most practical.

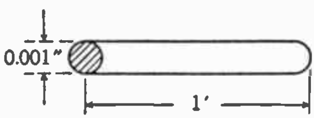


FIG. 2-2. A mil-foot of wire—1 ft long and 0.001 in. (1 mil) in diameter.

The resistances of several metals compared to silver are as follows:

Silver	1.00	Platinum	6.15
Copper	1.06	German silver	20
Aluminum	1.74	Constantan	27
Nickel	4.25	Mercury	59
Soft iron	6.00	Carbon	215

Problem 1–2. How many times higher in resistance is mercury than silver? than copper?

Problem 2–2. Two wires of the same length and diameter have resistances in the ratio of 5.65 to 1. If the lower-resistance wire is copper, could you identify the other wire material from the above table?

Problem 3–2. Two wires, one of soft iron and the other of platinum, are to have the same resistance. They have the same diameter. The platinum wire is 1 ft long. What is the length of the soft iron wire?

Problem 4–2. A spool of 0.008-in. wire has a gross weight of 3.59 lb; the spool weighs 0.41 lb. What is the length of wire on the spool?

Problem 5–2. What is the resistance of 500 ft of copper wire having a diameter of 40 mils?

Problem 6–2. A carbon rod 0.3 in. in diameter is 6 in. long. What is its resistance?

The ohm. The unit of resistance is the **ohm.** It is arbitrarily defined by international agreement as the resistance of a column of mercury weighing 14.4521 grams, having a uniform cross section and a height of 106.3 cm at 0° centigrade. A 9.35-ft length of No. 30 copper wire has a resistance of about 1 ohm. The table on page 19 gives sizes and resistance per 1000 ft of copper wire. The resistance per foot may be obtained from such a table by dividing the resistance per 1000 ft by 1000. In this table will be found the size of wires according to the B. & S. gage, the diameter in thousandths of an inch (mils), the resistance in ohms per 1000 ft, the weight, and the number of turns of the wire that can be got into an inch of winding space when the wire is covered with various insulations. "Scc" means single cotton covered; "Dcc" means double cotton covered, indicating that two layers of cotton thread are wound about the wire as insulation. Similarly "Ssc" refers to silk thread insulation.

Note that decreasing the size of the wire by three numbers, from No. 20 to No. 23, doubles the resistance of the wire, from 10.15 to 20.36 ohms; going from No. 30 to No. 27 lowers the resistance from 103.2 to 51.5 ohms per 1000 ft.

Copper is used in electrical and radio circuits because of its high conductivity compared to other metals and its low cost compared to metals of higher conductivity. It is readily obtainable and easily worked.

A circuit having 1 ohm resistance will pass a current of 1 amp if the emf is 1 volt.

The term **megohm** is frequently used in radio literature. It is equal to 1 million ohms.

Problem 7–2. What size of soft iron wire will have approximately the same resistance as No. 32 copper?

Problem 8–2. What is the resistance of 1 ft of No. 20 copper? of No. 24 aluminum?

Problem 9–2. A two-wire telegraph line is to be run a distance of 2 miles. If the total resistance of the line must be kept below 20 ohms, what size of copper wire must be used? If iron wire must be used because of expense, what size will be required? If copper and iron wire have about the relative weights of 555 to 480, what will the iron line weigh?

Conductance. The inverse of resistance is **conductance.** Conductance expresses the ability of a substance to pass a current of electricity, just as resistance expresses the ability of a substance to interfere with the passage of electricity. The unit of conductance is the **mho.** If the resistance is known, the conductance can be found by dividing the resistance into 1. Thus

$$\text{Conductance} = \frac{1}{\text{Resistance}}$$

The effect of temperature on resistance. The resistance of all pure metals rises with increase in temperature because of the greater molecular agitation at higher temperatures, making it more difficult for the electrons to drift in their progressive motion around the circuit.

At absolute zero, 273° below 0° centigrade, all molecular motion is supposed to stop, making the resistances of metals practically zero. Scientists have approached within a fraction of a degree of absolute zero.

Temperature coefficient of resistance. The amount the resistance of pure metals increases for each degree rise in temperature for each ohm at the original temperature is known as the **temperature coefficient of resistance.** This figure lies between 0.003 and 0.006 for pure metals, being 0.00426 for copper. This is the reason that the resistances of wires in the wire tables are

COPPER WIRE TABLES

Resistance at 68° F (20° C)

Mils = 0.001 in.

Size of Wire B. & S. Gage	Diameter of Wire Mils	Ohms per 1000 Ft	Pounds per 1000 Ft	Turns per Linear Inch			
				Scc	Dcc	Ssc	Dsc
0000	460	0.0490	640.5	2.14	2.10		
000	409.6	0.0618	508.0	2.39			
00	364.8	0.0779	402.8	2.68	2.62		
0	325	0.0983	319.5	3.00			
1	289.3	0.1239	253.3	3.33	3.25		
2	257.6	0.1563	200.9	3.75			
3	229.4	0.1970	159.3	4.18	4.03		
4	204.3	0.2485	126.4	4.67			
5	181.9	0.3133	100.2	5.21	5.00		
6	162	0.3951	79.5	5.88			
7	144.3	0.4982	63.0	6.54	6.25		
8	128.5	0.6282	49.98	7.35			
9	114.4	0.7921	39.63	8.26	7.87		
10	101.9	0.9989	31.43	9.25			
11	90.7	1.260	24.92	10.3	9.80		
12	80.8	1.588	19.77	11.5			
13	72	2.003	15.68	12.8	12.2		
14	64.1	2.525	12.43	14.3			
15	57.1	3.184	9.858	15.9	14.9		
16	50.8	4.016	7.818	17.9	16.7	18.9	18.3
17	45.3	5.064	6.200	20.0			
18	40.3	6.385	4.917	22.2	20.4	23.6	22.7
19	35.9	8.051	3.899	24.4			
20	32	10.15	3.092	27.0	24.4	29.4	28.0
21	28.5	12.80	2.452	29.9			
22	25.3	16.14	1.945	33.9	30.0	36.6	34.4
23	22.6	20.36	1.542	37.6			
24	20.1	25.67	1.223	41.5	35.6	45.3	41.8
25	17.9	32.37	0.97	45.7			
26	15.9	40.81	0.769	50.2	41.8	55.9	50.8
27	14.2	51.47	0.610	55.0			
28	12.6	64.90	0.484	60.2	48.6	68.5	61.0
29	11.3	81.83	0.384	65.4			
30	10.0	103.2	0.304	71.4	55.6	83.3	72.5
31	8.9	130.1	0.241	77.5			
32	8.0	164.1	0.191	83.4	62.9	101	84.8
33	7.1	206.9	0.152	90.0			
34	6.3	260.9	0.120	97.1	70.0	121	99.0
35	5.6	329.0	0.0954	104			
36	5.0	414.8	0.0757	111	77.0	143	114
37	4.5	523.1	0.0600	118			
38	4.0	659.6	0.0476	125	83.3	167	128
39	3.5	831.8	0.0377	135			
40	3.1	1049	0.0299	141	90.9	196	145

indicated as being at a given temperature; the value chosen is 20° C.

The resistance at any temperature when the resistance at a known temperature is available and when the temperature coefficient is known may be found from the formula

$$R_{t2} = R_{t1}[1 + \alpha(t_2 - t_1)]$$

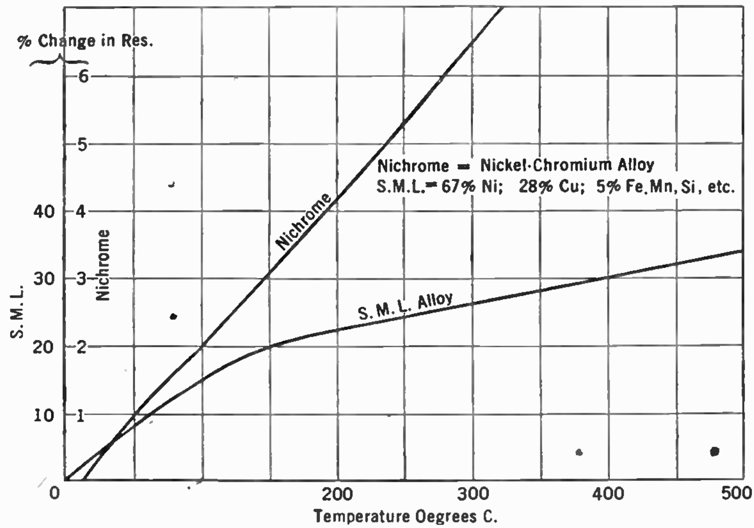


FIG. 3-2. How resistance of alloyed conductors varies with temperature.

where R_{t1} and R_{t2} are the conductor resistances at temperatures t_1 and t_2, and α is the temperature coefficient.

Example 1-2. A copper wire with a temperature coefficient of 0.00426 has a resistance of 80 ohms at 0° C. What will be the resistance at 50° C? The resistance will be increased by 80 × 0.00426 for each degree rise in temperature. At 50° C the resistance rise would be 80 × 0.00426 × 50 or 17.04 ohms, and the resistance would then be 80 + 17.1 ohms or 97.04 ohms. Using the formula this works out as follows:

$$\begin{aligned} R_{t2} &= 80[1 + 0.00426(50 - 0)] \\ &= 80(1 + 0.213) \\ &= 80(1.213) \\ &= 97.04 \text{ ohms} \end{aligned}$$

Typical temperature coefficients of resistance for several metals and alloys used in the radio profession are as follows:

MATERIAL	COEFFICIENT
Constantan (an alloy)	0.000002
Copper	0.00426
Copper-manganese-iron	0.00012
Iron	0.006
Nickel	0.006
Platinum	0.0037
Silver	0.0041
Tantalum	0.0033

Problem 10–2. A transformer supplying power to a load has a temperature of 70° C after being in operation for some time. If the resistance of one of its windings is 210 ohms at room temperature (20° C), what is its resistance at operating temperature?

Problem 11–2. A copper shunt to be placed across a current meter has a resistance of 0.78 ohm at room temperature, 20° C. If the meter is taken outdoors where the temperature is freezing (0° C), what is the resistance of the shunt?

Problem 12–2. Part of a government specification for a certain power inductor included the statement that the inductor should not get hotter than 90° C when running on full load. At room temperature the resistance of the winding was found to be 8.25 ohms, and after a prolonged run the resistance was 10.4 ohms. Did the inductor pass the specification?

Ohm's law. The law which governs all simple and many complex electrical phenomena is known as **Ohm's law.** This law states: Current in amperes equals emf in volts divided by resistance in ohms, or, in electrical abbreviations,

$$I \text{ (current)} = \frac{E \text{ (voltage)}}{R \text{ (resistance)}}$$

Ways of stating Ohm's law. There are three ways of stating this fundamental law.

$$[1]\quad I = E \div R \qquad [2]\quad E = I \times R \qquad [3]\quad R = E \div I$$

These three ways of stating the same law, determined from the first statement of Ohm's law by simple mathematical transformation, make problem-solving less difficult.

From these three expressions of Ohm's law, any one of the quantities can be obtained if the other two are known. Thus from equation 1, the current in a circuit can be determined if the voltage and the resistance of the circuit are known. From 2 the voltage required to force a desired current through a given resistance can be determined. Finally, from 3 the resistance of a circuit can be found if we can measure the current flowing in it under the force of a known voltage. Many circuits and apparatus follow much more complex laws than Ohm's law.

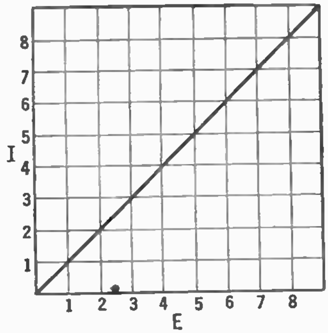


FIG. 4–2. In an Ohm's law circuit, a straight-line curve results when current is plotted against voltage.

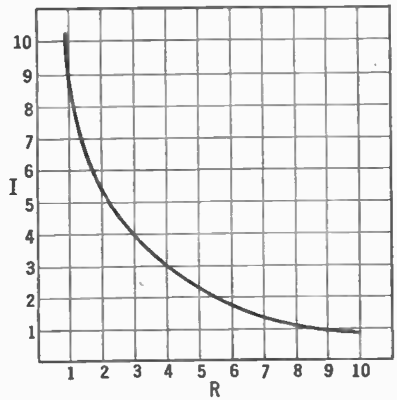


FIG. 5–2. Plotting current against resistance does not result in a straight line because current is proportional to $1/R$.

Example 2–2. A radio tube filament is heated by current flowing through it. The current comes from a battery to which the filament is connected. The battery has a voltage of 5, and the resistance of the filament is 20 ohms. How much current will flow through the filament?

We use equation 1, dividing the voltage, 5, by the resistance, 20, to obtain $5 \div 20$ amp or 0.25 amp.

Now suppose we know that the resistance is 20 ohms and that 0.25 amp is required to heat the filament sufficiently to cause it to produce the proper number of electrons. How many volts must be connected to the filament?

Here we use equation 2, multiplying the current, 0.25, by the resistance, 20, to get 5 volts required.

Finally, if we know from experience that the proper voltage is 5 and that with this voltage 0.25 amp will flow through the filament, what is its resistance? Here we use equation 3, and the reader should work out this relation for himself.

NOTE. The fundamental units are amperes, volts, and ohms. We cannot use volts, milliamperes, and ohms without getting into trouble. First the milliamperes must be converted into amperes and then used in the formulas expressing Ohm's law.

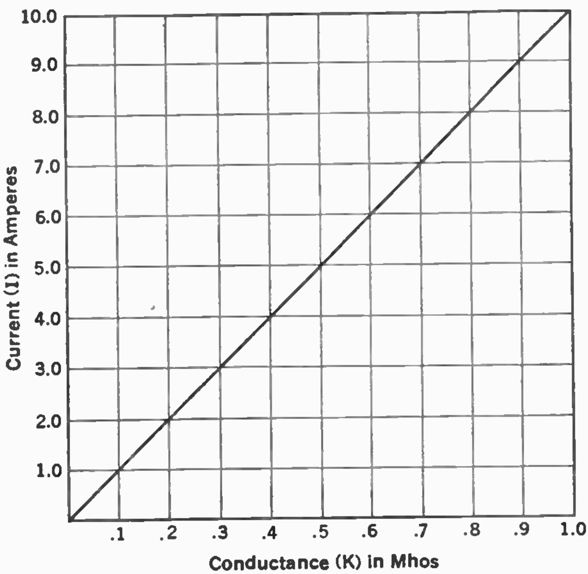


FIG. 6-2. Current plotted against K produces a "linear" curve indicating that current is directly proportional to conductance.

Voltage drop. The second way of stating Ohm's law indicates that, whenever a current flows through a resistance, a difference of potential exists at the two ends of that resistance. For every ampere of current that flows through an ohm of resistance, a volt is lost. In other words, 1 volt is required to force 1 ampere through 1 ohm of resistance.

Consider Fig. 7-2, which shows three resistors * connected in

* A **resistor** is a device for adding resistance to a circuit. Note that resistance is a property of a material or circuit, whereas a resistor is a

series and placed across a source of voltage. The emf causes current to flow through the three resistors. The current flowing through each resistor produces a voltage drop across each resistor; and the sum of the three voltage drops must equal the total voltage impressed across the entire series circuit. This is true since there is no other source of emf in the circuit; all the impressed voltage must be accounted for by the voltage drops. Thus, if 180 volts is impressed across the entire assembly, the voltages that will be measured by a voltmeter at other points in the series circuit are shown.

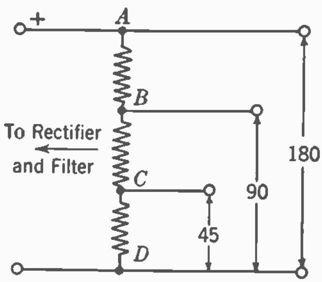


FIG. 7–2. A voltage-dividing network of resistances.

Example 3–2. In Fig. 7–2 suppose that 10 ma of current is forced through the three resistors in series. What is the resistance of each of the three resistors?

Here we use the third form of Ohm's law: $R = E/I$. Since 180 volts appears across the circuit and since 10 ma (0.01 amp) flows through it, the total resistance will be equal to

$$R = E \div I = 180 \div 10 \times 10^{-3} = \frac{180}{10 \times 10^{-3}} = \frac{180}{10^{-2}} =$$

$$180 \times 10^2 = 18{,}000 \text{ ohms}$$

Now, half the total voltage appears across the top resistor. Therefore the resistance between A and B is one-half the total resistance or 9000 ohms. Across 9000 ohms appears a voltage drop of 90 volts, and in the middle of this resistance is a voltage drop of 45 volts or half the total voltage of 90. Therefore the point C is half way between B and C, and the resistance between B and C must be 4500 ohms. The resistance from C to D is also 4500 ohms.

Often in laboratories a voltage is needed which is so small that it cannot be measured with available instruments. A larger voltage, however, can be measured easily, and, if it is impressed across a series of resistors like those in Fig. 7–2 (known as a

device possessing resistance. Engineers use the terms interchangeably in everyday language.

voltage divider), any desired part of the total voltage may be utilized by means of proper taps or connections or a sliding contact. A **potentiometer** is a continuous resistance with a sliding arm arranged so that any value of resistance between zero and the maximum value can be obtained.

The voltage appearing across a resistance because of current flowing through that resistance is known as a **voltage** or *IR* **drop.** It may be calculated by multiplying the resistance in ohms by the current in amperes.

Problem 13–2. A piece of electrical apparatus having a resistance of 8 ohms is plugged into a 32-volt farm lighting system. How much current will flow through the apparatus?

Problem 14–2. Suppose we desire to limit to 1 ma the flow of current in a circuit attached to a 45-volt battery. What must be the total resistance of the circuit?

Problem 15–2. What is the resistance of a 1C7-G tube filament taking 0.120 amp from 2-volt batteries?

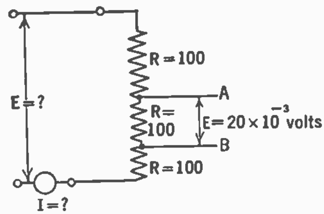


FIG. 8-2. Use of an *IR* drop as a source of voltage.

Problem 16–2. Consider Fig. 8-2. How many milliamperes of current must be forced through the circuit in order to get 20 mv across the resistor *A–B?* How many volts in all will be needed?

Power and energy. A battery converts chemical energy into electrical energy; a generator transforms mechanical energy into electrical energy. What do we mean by energy? What is power? These two terms are used rather loosely by most people, but each has a very definite meaning.

Energy is the *ability to do work.* A body may have one of two kinds of mechanical energy, either **potential** or **kinetic.** Potential energy is due to the position of the body; kinetic energy is due to its motion. A heavy ball on top of a flag pole has potential energy because if it falls it can do work, useful or not. It may heat the ground where it falls, or it may be used to drive a post into the ground. A cannon ball speeding through the air has energy because it can do work, useful or otherwise, if it is stopped suddenly. The target may be heated thereby, the kinetic energy possessed by the ball being converted into heat energy. The amount of damage done gives the eye a certain measure by which to judge the energy possessed by the cannon ball. This energy was originally possessed by the powder and was imparted to the ball when the powder exploded.

Whereas energy is the *ability* to do work, power is the *rate* of doing work. The horsepower, for example, is a unit of mechanical power. It is the power required to raise 33,000 lb of material 1 ft in 1 min; or 1 hp is 33,000 ft-lb per min.

All expressions for power involve the factor of time. It requires more power to accomplish a certain amount of work in a short time than in a longer time. For example, a ton of material raised a foot in the air represents 2000 ft-lb of work. If it is accomplished by a crane in 1 sec of time it represents an expenditure of 2000×60 or 120,000 ft-lb per min of power. Since 1 hp is equal to 33,000 ft-lb per min, the crane has a power of $120{,}000 \div 33{,}000$ or about 3.65 hp.

Now if a man raises the ton of material 1 ft in the air in an hour's time by going up a very long and gradual incline, his power is $2000 \div 60$ or 33.2 ft-lb per min, or roughly one-thousandth horsepower (0.001 hp). The amount of work done in the two cases is the same—the ton of material has been raised 1 ft in the air. The *rate of doing work* has changed.

Since power is the rate of doing work, the amount of work done in a given time is the rate of doing work multiplied by the time. Thus if 1 lb is raised 1 ft per hr and the work goes on for 2 hr, 2 ft-lb of work have been done. The same amount of work would be done if 1 lb were raised 1 ft per min and if the work went on for 2 min. In this case, however, the work would have been accomplished at a faster rate, requiring more power.

Three units are involved: energy, which is the ability to do work; power, which is the rate of doing work; and the work done. Energy and work are rated in the same units, horsepower-hours for example, or kilowatt-hours in electrical language.

In an electric circuit, the amount of power required to force a certain number of electrons per second through a certain resistance with a voltage of E volts is the product of the voltage and the amperage (since amperes are a measure of "rate of flow of electrons"). The unit of electrical power is the **watt.** Thus

Power in watts = Current in amperes × Emf in volts

or

$$P = I \times E$$

A kilowatt is 1000 watts. The kilowatt-hour is a measure of energy or work; a smaller unit is the watt-second or **joule,** and a still smaller one used in scientific circles is the **erg.** It takes 10 million ergs to make 1 watt-second.

Power lost in resistance. According to the law called the conservation of energy, energy can be neither created nor destroyed. It comes from somewhere and goes somewhere. Similarly, all *power*, which is the rate at which energy is used, must be accounted for. The energy required to force current through a resistor must do some work. It cannot disappear. This work results in heating the resistor. The heat appears because of the greater molecular activity which results from the flow of electricity through the material. Whenever current flows through a resistor, heat is generated, and the greater the current the greater the heat. As a matter of fact, the heat is proportional to the square of the current. If the wire is heated faster than the heat can be dissipated by heating the surrounding air, the wire melts. Energy has been supplied to the unit at too great a rate.

A resistor used in a radio circuit is often rated at so many ohms and as capable of dissipating so many watts. Thus a 1000-ohm, 20-watt resistor means that the resistance of the unit is 1000 ohms, and that 20 watts of electrical power can be put into it without danger of burn-out.

Problem 17-2. A voltmeter across a resistor measures 50 volts. Through the resistor flows 100 ma. What power is expended in the resistor?

Problem 18-2. Assuming that 50 ma direct current flows through a resistor which has a d-c resistance of 750 ohms, what amount of heat in watts must be dissipated? What supplies this power?

Expressions for power. Just as there are three ways of stating Ohm's law, so there are three ways of stating the relation between watts, volts, amperes, and ohms. Thus,

$$[1]\quad P = I \times E \qquad [2]\quad P = I^2 \times R \qquad [3]\quad P = E^2 \div R$$

A useful expression is $I = \sqrt{P/R}$. It may be employed in calculating the current safely passed by a resistor of a given wattage and resistance.

Example 4–2. A plate voltage power supply system supplies 180 volts to a power tube which consumes 20 ma. How much power is taken? What is the resistance of the power tube?

The power supplied is $E \times I = 180 \times 0.02 = 3.6$ watts. The resistance into which this power is fed is equal to $P \div I^2 = 3.6 \div 0.0004 = 9000$ ohms, or $E^2 \div P = 180^2 \div 3.6 = 32{,}400 \div 3.6 = 9000$ ohms.

The maximum current that can pass through one's body without serious results is 0.01 amp. The resistance of the body varies with one's health, the surface in contact, etc. If the finger tips of the two hands are dry, the resistance from one hand to the other is about 50,000 ohms.

Problem 19–2. A man touches with his dry finger tips a 500-volt railway "third rail." If the resistance of his body is 50,000 ohms, how much power is used up in heating the body? Will the dangerous current of 0.01 amp be exceeded?

Problem 20–2. A voltage of 110 is to be placed across a circuit whose resistance must be such that 220 watts can be delivered. What is the resistance of the circuit?

Problem 21–2. How much power is taken from a storage battery which supplies five radio receiver tubes each requiring 0.3 amp at 6.3 volts?

Problem 22–2. One milliampere of current flows through a 100,000-ohm resistor. How much power in heat must the resistor be capable of dissipating?

Problem 23–2. How much current can be sent through a 1000-ohm 20-watt resistor without danger of burning it up? What voltage is required?

Problem 24–2. In a voltage supply system the voltage divider has a total resistance of 5000 ohms. The receiver requires a maximum current of 30 ma. What must be the wattage rating of the resistor if all this current flows through it?

Problem 25–2. A mobile radio receiver of five 6.3-volt tubes consumes a plate current of 50 ma at 180 volts. Assume that the car battery supplies this load; what power is required?

Problem 26–2. Electric power is bought by kilowatt-hour. Suppose that your rate is 10 cents per kilowatt-hour. How much does it cost to run a soldering iron on a 110-volt circuit if it consumes 6 amp?

Efficiency. Efficiency is a term that is loosely employed by nearly everybody. Anything which works is said to be efficient, and one's efficiency is often confused with his energy—his ability to do work whether the work is actually carried out or not. The term, however, has a very exact meaning when one uses it in speaking of mechanical or electrical systems of any kind. Efficiency is a relative term. It is a ratio showing how much *useful* work one gets out of a *total* amount of work done.

Let us consider a steam engine connected to a dynamo, a combination of machines for transforming mechanical energy into electrical energy. If the steam engine consumes 1 hp (746 watts) and delivers 500 watts of electrical energy, it is said to be more efficient than if it delivered only 250 watts. Let us consider two men, one of whom gets a lot of work done in a small amount of time and with an expenditure of little effort. The other gets the same amount of work done but with great effort, perhaps flurrying about from one thing to another instead of tackling his problem in a straightforward manner. The first man is more efficient. He *wastes* less time and energy.

Efficiency, then, is the ratio between useful work or energy or effect got out of a machine and the total energy or power or effort put into it. It is expressed in percentage. A machine that is 100 per cent efficient has no losses; there is no friction in its bearings, or, if it is an electrical device, no resistance in its wires. There are no such machines in use today. Efficiency is the ratio of useful power one gets out of a device to the power put into it.

$$\text{Efficiency} = \frac{\text{Useful output}}{\text{Input}} = \frac{\text{Useful output}}{\text{Output plus losses}}$$

Problem 27–2. Seventy-five per cent of the ampere-hours put into a battery are returned by it on discharge. How many hours must a 100-amp-hr battery be charged at a 1-amp rate?

Problem 28–2. A motor generator or vibrator system supplies the plate voltage to a police-car radio which uses eight 6.3-volt tubes requiring a plate current of 80 ma at 180 volts. If its efficiency is 40 per cent, what power and current are taken from the battery for the plate circuits?

Problem 29–2. If a 5-hp motor operated from a 110-volt d-c line is 85 per cent efficient, what current does it take from the power line?

Problem 30–2. A 200-amp-hr battery is taken into the field to operate a portable radio transmitter which requires 40 amp-hr. It is keyed so that power is taken from the battery 75 per cent of the time. The transmitter consumes 50 watts, and in addition to this load there is a continuous load of 1.5 amp for a light. How long will the battery supply the transmitter and lamp?

Problem 31–2. A broadcast transmitter radiates 10 kw in the daytime but is permitted to radiate only 5 kw at night. If the daytime antenna current is 15 amp, what is the night-time current? Assume constant resistance.

Problem 32–2. A transmission line made up of two wires transmits power from a radio transmitter to an antenna. At the transmitter end the voltage across the line is 2185 volts; at the antenna end the current is 4.65 amp. What power is lost in the line if 10 kw is delivered to the antenna?

Problem 33–2. If the strength of a radio signal at a distant point is proportional to the square root of the power at the transmitter, by how much must the transmitter power be increased to double the strength of the signal at the receiving end? How much should the transmitter antenna current be increased?

Problem 34–2. A generator delivers 42 amp at 440 volts at an efficiency of 82 per cent. What power is lost in the generator?

Protective devices. If the wattage rating of a resistor is exceeded, that is, if energy is fed into it too rapidly, heat is generated faster than it can be lost. The amount of heat that any electric device can dissipate without damage has a definite limit.

The simplest means of protecting a device against injury is a fuse. A fuse is merely a piece of wire which will melt and open the circuit before damage is done to the device it is protecting. The fuse must be so constructed that it will not "blow" on sudden and temporary overloads caused, for example, by a motor starting from rest. During the fraction of a second that the speed of the motor is increasing, the current required is much greater than when the motor is running at its normal speed. Therefore a large amount of current will be taken from the line, and it is possible that this current, if it continued to flow, would damage the motor or the line. A fuse in series with the motor will melt if the current passed through it is too great for too long a time. This opens the circuit. Someone, then, must replace the fuse before the circuit will function again.

Replacing a fuse is a nuisance. A more common device in a factory to protect machinery against overload is a circuit breaker or an overload relay. These are mechanical switches which open the circuit automatically when the current becomes too great. They may be arranged to close again automatically after a second or two, or they may be so made that a maintenance man must close the circuit by hand after he has cleared up the cause of the overload.

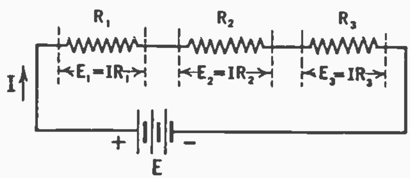


Fig. 9–2. A simple series circuit.

Series and parallel circuits. When two or more pieces of equipment are connected as in Fig. 9–2 they are said to be in **series.** The same current flows through each unit. The voltage drop across each unit is controlled by its resistance, and if one of these units has twice the resistance of the other, the voltage drop across it will be twice as great. The sum of the voltage drops across the three resistances must be equal to the voltage of the battery, for there is no other source of voltage in the circuit.

In a series circuit the total resistance is the sum of the individual resistances. The current in each unit is the same as in all other units. The current is obtained from Ohm's law, equation 1.

If any of the units becomes "open" the current ceases to flow. If, however, any unit becomes "shorted" the current will increase because the total resistance of the circuit has decreased.

Example 5–2. In Fig. 10–2 is a typical series circuit composed of a vacuum tube R_2, a 6-volt battery, a current meter, and a rheostat or variable resistor whose purpose is to limit the flow of current through the filament of the tube. The arrow through R_1 indicates that it can be adjusted in value.

The question is, what current will flow through the circuit as the resistance of R_1 is varied? Suppose it is 4 ohms. We know the same current will flow through both the filament and the rheostat. The resistance, then, in the circuit is equal to 20 plus 4 or 24 ohms, and by Ohm's law we know that the current will be the voltage divided by the total resistance, or

$$I = \frac{E}{R_1 + R_2} = \frac{6}{4 + 20} = \frac{6}{24} = 0.25 \text{ amp}$$

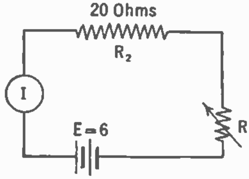


FIG. 10–2. Typical circuit made up of a battery, a rheostat, a tube (R_2), and a current meter.

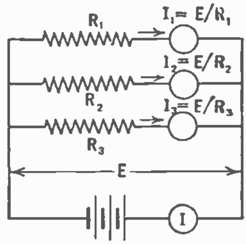


FIG. 11–2. A parallel circuit.

There are two resistances in this circuit. Current flows through them. There must then be two voltage drops. Let us calculate what they are. By equation 2 we multiply the resistance by the current.

$$\text{Voltage drop} = IR_1 = 0.25 \text{ amp} \times 4 \text{ ohms} = 1 \text{ volt}$$

$$\text{Voltage drop} = IR_2 = 0.25 \text{ amp} \times 20 \text{ ohms} = 5 \text{ volts}$$

In other words, of the 6 volts available at the terminals of the battery, 5 have been used up across the 20-ohm resistance and 1 volt has been used to drive 0.25 amp through the 4-ohm resistance.

Problem 35–2. In an ac-dc radio receiver there are four tubes of the 6.3-volt type connected in series. What resistance must be put in series with the filament of those tubes if they are to be put directly across a 115-volt line? Each tube requires 0.3 amp.

Problem 36–2. Suppose you were going to use five 6.3-volt tubes in a series filament circuit. How many volts will be necessary?

Problem 37–2. What would be the resistance of the above tubes in series?

Problem 38–2. On a 32-volt farm system how many tubes of the 6.3-volt type can be run in series? Six of the 2-volt, 0.06-amp type are to be run from this system. What series resistance is necessary?

Problem 39–2. How much resistance would be necessary if one each 1R5, 1T4, 1S5, and 1S4 tubes are to be run from a 2-volt storage battery?

Problem 40–2. An incandescent lamp has a resistance, when hot, of about 55 ohms, and requires 1 amp to light at full brilliancy. How many could be run in series on a 110-volt circuit?

Problem 41–2. How many volts are required to force 1 ma through a circuit composed of a vacuum tube and a resistance, if the latter has 100,000 ohms and if 90 volts are required across the tube?

Characteristics of parallel circuits. A parallel circuit is represented in Fig. 11–2. It consists of several branches. The voltage across each branch is the same as that across every other branch and is equal to the voltage of the battery. The total current supplied by the battery is the sum of the currents taken by the branches. The resistance of the group may be found by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

where R is the resultant or total resistance, and R_1, R_2, etc., are the individual resistances.

The resultant resistance of several units in parallel is less than the individual resistance of any of the components. If two equal resistances are in parallel, the resultant is one-half the resistance of one. This fact follows logically from the following reasoning. If two identical resistances are placed across a given voltage, it will be found that they pass twice the current taken by either of the resistances. If, therefore, a new resistance is selected, which passes as much current as the two resistances taken together, it will have half the resistance of either of the identical resistances. Working this out by Ohm's law will verify the reasoning. Thus, if two 10-ohm resistances are connected in parallel, the resultant resistance is 5 ohms. What would it be if they were connected in series?

If any number of equal resistances are in parallel, the resultant resistance is the individual resistance divided by the number of units.

If only two resistances are in parallel the resultant may be calculated by dividing their product by their sum:

$$R = \frac{R_1 \times R_2}{R_1 + R_2}$$

This simplified formula comes directly from the one above it by simple algebra, and the reader should prove it.

Example 6–2. What is the parallel resistance of two units which have resistances of 4 and 5 ohms?

This can be solved by either of the formulas given.

$$\frac{1}{R} = \frac{1}{4} + \frac{1}{5}$$

$$= 0.25 + 0.20$$

$$= 0.45$$

$$R = 1 \div 0.45 = 2.22 \text{ ohms}$$

Or

$$R = \frac{R_1 \times R_2}{R_1 + R_2}$$

$$= \frac{4 \times 5}{4 + 5}$$

$$= \frac{20}{9} = 2.22 \text{ ohms}$$

Example 7–2. Suppose that, as in Fig. 12–2, two resistances in parallel are placed in series with a resistance of 1 ohm and across a battery of

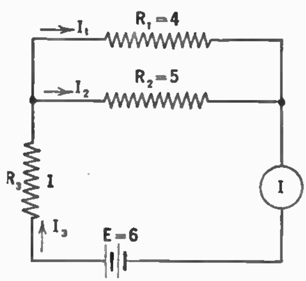


FIG. 12–2. A problem involving both series and parallel resistances.

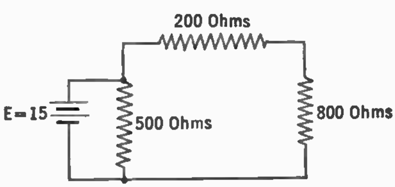


FIG. 13–2. Another problem. What is the voltage across the 800-ohm resistor?

6 volts. What current would flow out of the battery and through each resistance?

The total resistance is 2.22 + 1 = 3.22 ohms. The current flowing, then, is 6 ÷ 3.22 = 1.86 amp. This current through the combined resistance of the 4- and 5-ohm units produces a voltage drop of $I \times R$ or 1.86×2.22 or 4.14 volts. This voltage across 4 ohms produces a current of $4.14 \div 4$ or 1.035 amp, and across 5 ohms produces a current of 0.827 amp. These two currents added together are 1.862 amp, which checks our calculation above.

Problem 42–2. A radio receiver has five tubes of the a-c heater type, each taking 1.75 amp. What is their combined resistance, and how much current do they take from a 2.5-volt transformer secondary? If another load of the same voltage but half the total current is added, what current is required?

Problem 43–2. A circuit has three branches of 4, 6, and 8 ohms. A current of 4 amp flows through the 6-ohm branch. What current flows through the other branches?

Problem 44–2. Consider the circuit of Fig. 13–2. What is the voltage drop across the 800-ohm resistor?

Use of conductances. Remembering that conductance (K) is the reciprocal of resistance ($K = 1/R$), solving for the effective resistance of several parallel resistances really involves calculating the conductance corresponding to each of the resistances, adding these conductances together, and then taking the reciprocal of the effective conductance to find the resultant resistance. Thus

$$\frac{1}{R_0} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

But

$$K = \frac{1}{R}$$

$$K_0 = K_1 + K_2 + K_3$$

and

$$R_0 = \frac{1}{K_0}$$

where R_0 and K_0 are the resultant resistance and conductance.

Conductances are useful in parallel circuits, but, since the conductance of a piece of apparatus is seldom given, conductances are not often used, as such. The fact that the actual values of conductances are usually very small (if $R = 100$ ohms, $K = 0.01$ mho; if $R = 10{,}000$ ohms, $K = 0.0001$ mho) adds to the difficulty of using this unit.

Ohm's law in complex circuits. It must be remembered that Ohm's law applies to a circuit as a whole, or to any of its parts. If the voltage and resistance of any part of a circuit are known, the current through that part can be calculated without regard to any other part of the circuit. In the following section, the solution of complex circuits is described.

Kirchhoff's laws. When circuits are made up of series and parallel elements, solving for the individual currents or voltages becomes somewhat complex. All such circuits can be reduced to simpler circuits which can be solved more easily. Two rules known as Kirchhoff's laws are useful when such complicated circuits as that shown in Fig. 14–2 are to be solved.

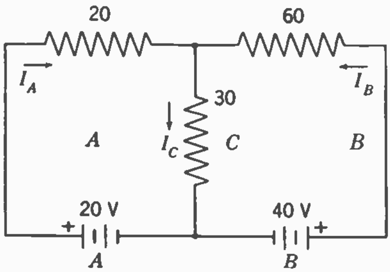


FIG. 14–2. An example of the use of Kirchhoff's laws to solve complex circuits.

Kirchhoff's laws are: (1) The current flowing into any junction in a circuit must be equal to the current flowing away from that junction. (2) The algebraic sum of the sources of emf and voltage drops in any closed circuit must be equal to zero.

The first law states that current flowing toward a point at which it may divide must equal the sum of the portions into which it divides. Otherwise, some current would be left over with no place to go. The second law states that the voltage supplied by the source of emf (a battery, for example) must be equal to the sum of the voltage drops ($I \times R$) appearing across the several circuit components. All the battery voltage must be accounted for, and the voltage drops cannot exceed the emf produced by the battery for there is no other source of emf.

In solving complicated circuits, only one portion need be considered at a time. For example, in Fig. 14–2 one might think that battery *B*, having higher voltage than battery *A*, would force some current down through the *A* circuit and that therefore one ought to subtract the two voltages. This is unnecessary, as will be shown.

Example 8–2. Current through C is made up of two parts, one due to battery A and one to battery B. No other current flows through C since there is no other source of current. Then Kirchhoff's first law states that $I_C = I_A + I_B$. The junction is the point at which the two currents I_A and I_B join.

Applying the second law, the emf of battery A is accounted for by adding up the two voltage drops in circuit A ($20I_A$ and $30I_C$ made up of I_A flowing through 20 ohms and I_C through 30 ohms). Similarly, the emf of battery B produces two voltage drops in circuit B. Thus

$$I_C = I_A + I_B \text{ (Kirchhoff's first law)} \quad [1]$$

$$20 = 20I_A + 30I_C \text{ (Kirchhoff's second law)} \quad [2]$$

$$40 = 60I_B + 30I_C \text{ (Kirchhoff's second law)} \quad [3]$$

Now we have three unknown currents, I_A, I_B, and I_C, and three equations expressing the relations among them. Rewrite equation 1 as

$$I_A = I_C - I_B$$

Substitute this value for I_A in equation 2:

$$20 = 20I_C - 20I_B + 30I_C$$

or

$$20 = 50I_C - 20I_B \quad [4]$$

and from equation 3

$$40 = 30I_C + 60I_B$$

Now multiply both sides of equation 4 by 3:

$$60 = 150I_C - 60I_B \quad [5]$$

and add equations 3 and 5. Thus

$$\begin{aligned} 40 &= 30I_C + 60I_B && [3] \\ 60 &= 150I_C - 60I_B && [5] \\ \hline 100 &= 180I_C + 0 \end{aligned}$$

or $180I_C = 100$.

$$I_C = 100 \div 180 = 0.555 \quad [6]$$

From equation 3

$$40 = 60I_B + 30I_C$$

Substitute for I_C its value 0.555:

$$40 = 60I_B + 30(0.555)$$

$$40 = 60I_B + 16.7$$

$$40 - 16.7 = 60I_B$$

or

$$I_B = 60 \div 23.3 = 0.389 \qquad [7]$$

From equation 1

$$I_A = I_C - I_B$$

$$= 0.555 - 0.389 = 0.166 \qquad [8]$$

Therefore

$$I_A = 0.166$$

$$I_B = 0.389$$

$$I_C = 0.555$$

The current through the middle branch (I_C) is made up of contributions from circuits A and B. To determine the share each contributes, we may make use of another rule called the **superposition theorem,** which states that we can determine the current through C branch due to battery A by shorting out battery B and calculating the current that will flow. With battery B shorted out, battery A forces current through a 20-ohm resistance and through two resistances of 30 and 60 ohms in parallel. Remembering that the resultant of two resistances in parallel is equal to the product of the two divided by their sum we can solve for the current through C due to A as follows:

$$I_A = \frac{20}{20 + \dfrac{30 \times 60}{30 + 60}} = \frac{20}{20 + \dfrac{1800}{90}} = \frac{20}{20 + 20} = \frac{20}{40} = 0.5$$

This is the current flowing out of battery A. But not all of it flows through the 30-ohm resistance in which we are interested. Part flows through the 60-ohm resistance in parallel with the 30-ohm resistance C. The current through the individual portions of the parallel circuit will be inversely pro-

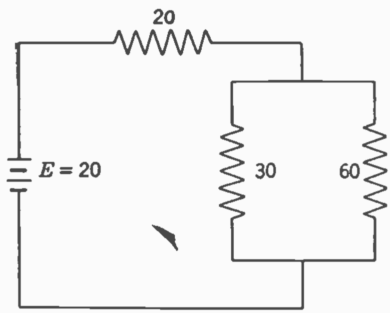


FIG. 15–2.

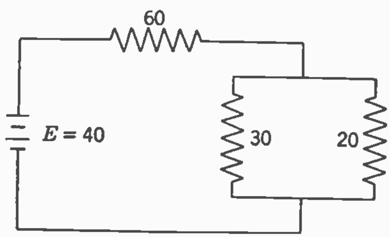


FIG. 16–2.

FIGS. 15–2 and 16–2 show two portions of the circuit of Fig. 14–2 as an example of the use of the superposition theorem.

portional to their resistances and will actually be equal to the resistance in which we are not interested divided by the sum of the two resistances.

Thus the current through the 30-ohm resistance will be equal to the total current flowing (I_A) multiplied by 60 ÷ 90.

$$I_C = I_A \times \tfrac{60}{90} = \tfrac{2}{3}I_A = \tfrac{2}{3} \times 0.5 = 0.333$$

Similarly the current through C due to battery B will be found to be 0.222, and the sum of these currents 0.555 equals the current found by Kirchhoff's laws. (The reader should solve for the current in C due to battery B.)

NOTE. In solving problems of this type the engineer must assign a definite direction of current flow and then stick to it. Thus, if he decides to follow convention, current flowing from positive to negative through a closed circuit, then, in all portions of the circuit, current must flow from positive to negative—considering only the apparatus attached to the battery or emf. Through the battery or emf itself the current will have to flow from negative to positive to complete the circuit.

Problem 45–2. In Example 8–2, reverse the connections of battery A and solve for the currents.

Problem 46–2. In Fig. 17–2, determine the voltage required. If resistors are available in wattages of 0.5, 2, and 10, determine which wattage ratings should be used for each resistor.

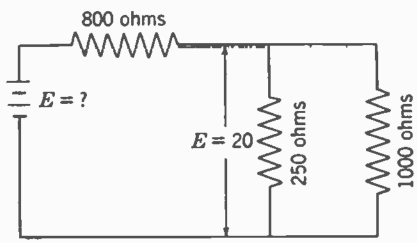


FIG. 17–2. Another typical d-c problem.

CHAPTER 3

ELECTRICAL METERS AND MEASUREMENTS

To utilize electric currents properly we must be able to detect and measure them. We cannot see, hear, or smell the passage of an electric current through a circuit; the current must be made evident to us by its effect upon the circuit. Three kinds of effects may be produced: thermal, chemical, and magnetic. Wire gets hot if too much current flows through it; two dissimilar metals (copper and zinc, for example) placed in a solution of one of them (copper sulphate) give off gas bubbles when a wire connects them externally; a wire carrying an electric current and brought near a compass needle will cause the needle to change from its habitual north-south position.

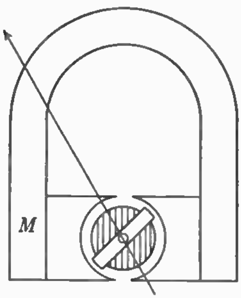


FIG. 1-3. A D'Arsonval meter made up of a permanent magnet, *M*, and a coil of wire free to rotate.

Any one of these three fundamental effects of electricity can be used to detect the presence of a current or even to measure the rate at which the current flows. A hot-wire ammeter, for example, is merely a wire which expands when heated by current flowing through it. A needle is attached to the wire and is pulled across the scale by a spring as the wire gets hot.

By far the greater number of modern meters for use on direct current employ the D'Arsonval movement, which is based on the fact that a force is exerted on a current-bearing wire in a magnetic field. The D'Arsonval movement consists of a permanent magnet to which pole pieces of soft iron are affixed. Between these pole pieces is placed a bobbin or coil of wire to which is attached a pointer. When current flows through the coil, a force is produced which tends to make the coil rotate from its no-current position. Since this force is proportional to the current, a scale can be attached to the meter in such a way that the pointer indicates exactly the current flowing through the coil.

Such current meters may be made sensitive enough to measure a microampere and to detect (without measuring accurately) much smaller current than this.

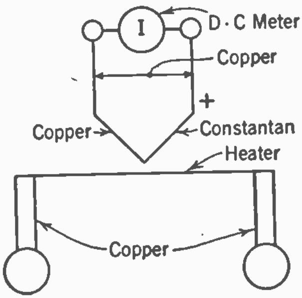


FIG. 2–3. Use of thermoelectric effect as a means of measuring current.

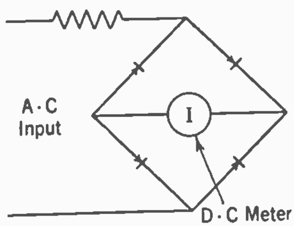


FIG. 3–3. Rectifiers in a bridge circuit for measuring alternating currents by means of a d-c meter.

Thermocouple instruments. In another type of instrument, advantage is taken of the fact that two dissimilar metals heated at the junction between them produce a measurable emf. The basic elements of a thermocouple meter are shown in Fig. 2–3. Here copper and an alloy called constantan are heated by their proximity to a wire through which the current to be measured passes. The emf produced by the couple is measured on a meter of the D'Arsonval type. D'Arsonval meters will measure currents as low as 2 ma through the heater wire. The indicating instrument of such a meter will be a 200-μa, 12-ohm movement.

Rectifier meters. By means of a copper oxide rectifier, alternating current may be transformed into direct current which is then measured on a D'Arsonval movement. The copper oxide units consist of flat plates rather close together, and the capacitance between them may be rather high. At high frequencies, therefore, they will act as a partial short circuit for the currents to be measured. For this reason only the lower frequencies may be measured in this manner. If, instead of the copper

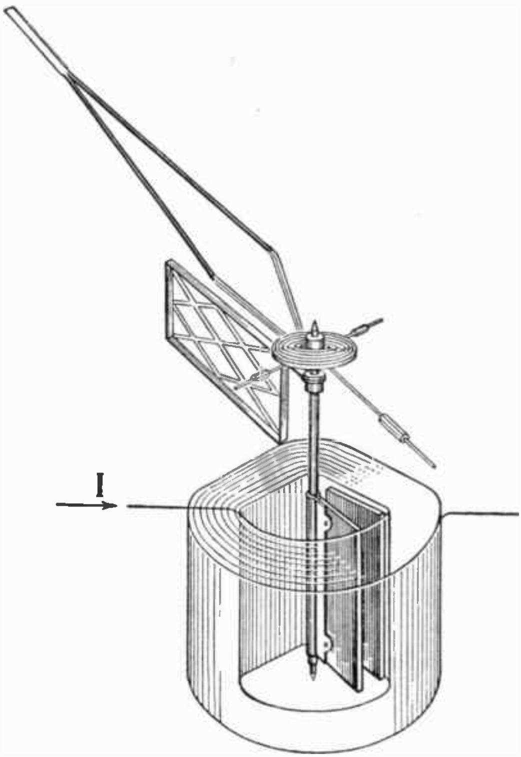


FIG. 4-3. Details of an iron-vane current-measuring instrument.

oxide rectifier, a "crystal detector" is employed, quite high frequencies may be measured. The crystal detector consists of a crystal of some substance like Carborundum or silicon with a small contactor placed on it, and the capacitance between the two portions of the rectifier may be quite low.

Thermocouple instruments are used at the higher frequencies encountered in radio work. Audio frequencies may be measured by the rectifier meters. The old-fashioned hot-wire meter has practically disappeared, although it was in wide use at one time.

Iron-vane meters. In another type of instrument, used mostly on alternating current, the magnet and the coil are reversed in the roles they play. The magnet is made of soft iron and is the movable part while the coil remains stationary. Currents as low as 15 ma may be measured in this way. The meter is limited in its usefulness to frequencies up to 200 cycles per second because the inductance of the coil cuts down the current flow and torque. The copper oxide meter is the common a-c voltmeter; but any ammeter placed in the a-c line will have an iron vane in it. Figure 4–3 shows the elements of a Weston iron-vane meter.

PROPER USE OF METERS

Ammeters. Meters to measure current are called ammeters. They are connected in series with the source of current and the load. Passing too much current through them easily damages them, bending the pointer, injuring the movement mechanically, or actually burning out the moving coil if the current is too great and is allowed to flow too long. Meters may be protected by fuses or by shunting them by conductors which allow most of the current to by-pass the meter itself. After one is certain that the current is within the range of the meter, the shunt may be removed.

For example, the current can be regulated so that the meter reads full scale. Then a rheostat or other variable resistor is placed in parallel with the meter and its value is adjusted until the meter reads one-half or whatever proportion of the total current is to be measured by the combined meter and shunt. Then that value of resistance, so placed across the meter, will decrease the sensitivity of the meter by one-half or whatever ratio is decided upon.

One must be careful to see that the shunt resistance can pass the shunted current without overheating, and one must remember that most conductors have resistances which vary with temperature. Thus, if the shunted current causes overheating of the shunt resistor, then the meter is no longer shunted by the proper resistance but by some new value determined by the temperature coefficient of resistance of the shunt. Manganin is often used in shunts because it has a low temperature coefficient.

The lower the resistance of the shunt compared to the meter resistance, the greater proportion of the total current will flow through the shunt, and the less will go through the meter. Therefore a low resistance is used when the sensitivity of the meter is to be reduced considerably so that much higher currents can be measured than the meter acting alone can handle without damage.

If the resistance of the meter is known, the value of shunting resistance can be calculated. It is a good exercise in parallel-circuit operation to calculate the shunt resistor value in terms of the meter resistance. The value is

$$R_s = R_m\left(\frac{X}{1 - X}\right)$$

where R_s is the shunt resistance.
R_m is the meter resistance.
X is the proportion of the total current that is to flow through the meter, i.e., one-tenth, one-half, etc.

The manner in which ammeters may be adapted to read currents higher than originally intended may be shown by the following example. A given meter reading 1 ma full scale has a resistance of 28 ohms. Now if it is shunted by a resistance of 0.57 ohm the total current taken by the meter and its shunt will be 50 ma, but only 1 ma will go through the meter and 49 ma will go through the shunt, because the shunt has so much less resistance than the meter. If the shunt is properly designed with respect to the resistance of the meter, the original readings of the meter may merely be multiplied by the proper factor to make them a measure of the total current flowing through meter and shunt.

Voltmeters. Ammeters have low resistance. They are connected in series with the apparatus taking current from the source, as shown in Fig. 5–3. Voltmeters, on the other hand, must read the voltage across some part of the circuit. They must not permit much current to flow through them because this current would be taken away from the circuit. Therefore, they are really high-resistance ammeters. An ammeter, a 0- to 1-ma meter, for example, can be made to read volts by putting it in series with a high resistance and placing the combination across the circuit to be measured.

FIG. 5-3. Ammeters are connected in series with the resistance into which the current flows.

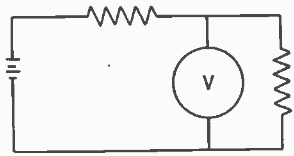


FIG. 6-3. Voltmeters are connected across the point at which voltage is to be measured.

For example, 1 volt is required to give 1 ma of current through 1000 ohms. If we have a 1.0-ma meter and we wish to measure a voltage of the order of 1 volt, we need a total resistance, meter plus series resistance, of 1000 ohms. Then the figures on the meter scale will read volts instead of milliamperes. Such a series resistance is called a *multiplier*. The sensitivity of a voltmeter is often stated as its resistance per volt. Thus a meter to be used on circuits from which it is not permitted to take much current may have a resistance of 1000 ohms per volt. This means that a meter to measure a maximum voltage of 100 will have a resistance of 100,000 ohms. It will require less current for full-scale deflection than a meter with a resistance of only 100 ohms per volt.

Problem 1-3. What voltage is required to produce a full-scale deflection on a 1-ma meter having a resistance of 28 ohms? What current is required to produce a 25-volt reading on another meter having a resistance of 100 ohms per volt?

Sensitivity of meters. A sensitive current-measuring meter is one which will measure very small currents but which has a low resistance. A sensitive voltmeter is one which will give a large needle deflection through a very high resistance. Voltmeters which are used to measure the voltage of high-resistance devices such as plate voltage supply units have high resistance in order that the current taken from the device shall not be great enough to lower appreciably the voltage of the device.

Example 1–3. Suppose that we are to measure the voltage across the circuit at the point X in Fig. 7–3. The voltage at X depends upon the current taken by the meter. What is desired is the open-circuit or no-load voltage across X, that is, the voltage existing there if no current is taken by the meter. If no current flows, there is no voltage drop in the resistance R and hence the voltage at X is the voltage of the battery, or 100 volts. Suppose, however, that the meter has a resistance of 1000 ohms. The current flowing through the meter is given by Ohm's law

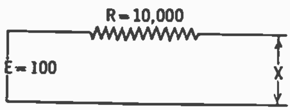


FIG. 7–3. A low-resistance voltmeter placed at X will not read the open-circuit voltage.

$$\begin{aligned} I &= E \div R \\ &= 100 \div (10{,}000 + 1000) \\ &= 100 \div 11{,}000 = 0.0091 \text{ amp or } 9.1 \text{ ma} \end{aligned}$$

This current through the 10,000-ohm resistance R (which may be the internal resistance of the battery E, page 57) causes a voltage drop across this resistance of $I \times R = 0.0091 \times 10{,}000 = 91$ volts.

The voltage actually recorded on the meter, then, is the difference between the battery voltage and the drop across the resistance R, or

$$\text{Voltage at } X = E - (I \times R) = 100 - 91 = 9.0 \text{ volts}$$

If, however, the meter is a high-resistance meter, say 1000 ohms per volt, that is, 100,000 ohms for a meter designed to read 100 volts, the current taken from the battery would be

$$I = E \div R = 0.00091 \text{ amp}$$

and the IR drop across the resistance R would be only

$$\begin{aligned} E = IR &= (0.00091 \times 10{,}000) \\ &= 9.1 \text{ volts} \end{aligned}$$

and the voltage read at X would be $100 - 9.1$ volts or 90.9 volts.

In other words the high-resistance voltmeter gives a reading much nearer the open-circuit or no-load voltage desired.

Wattmeters. Since power in watts is the product of the emf in volts and the current in amperes, a wattmeter must be arranged to read the voltage across a device and the current through it to determine the power taken by the device. Wattmeters, therefore, have two coils, one of which (called the current coil) is in series with the line and the load, and the other (called the potential coil) is in shunt with the load across the line. The current coil measures the current taken, and the potential coil measures the voltage. The interaction of these two coils causes a pointer to indicate the amount of power taken by the device.

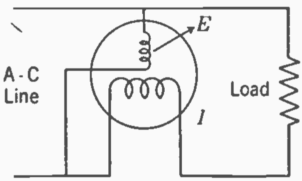


FIG. 8–3. A wattmeter has two windings, one in series with the load to measure current and another across the line to measure the voltage.

RESISTANCE MEASUREMENTS

The most important and most frequent single measurement in radio practice is the measurement of resistance. Many resistors are used in radio receivers and transmitters, and the actual resistance of these units is often important. Furthermore, the ohmic resistance of many units, used for purposes other than as resistors, is an important indication of the condition of the unit, i.e., whether it is normal or whether it needs to be replaced. Methods of measuring resistance, then, are important and useful.

Ammeter-voltmeter method of measuring resistance. The simplest method of measuring resistance follows from Ohm's law. The resistance of a device is the ratio of the voltage across the device to the current through it ($R = E \div I$). If the voltmeter utilized has a much higher resistance than that of the device being measured, its inclusion in the circuit need not be considered. Otherwise it must be remembered that the voltmeter is shunted across the unknown resistance and therefore permits more total current to flow than flows when the device is connected in the circuit by itself. The resistance of the current-measuring meter is usually so low that it will not affect the value of current flowing through the device.

Example 2–3. Consider the circuit in Fig. 9–3. A voltmeter V_1 across the device whose resistance is unknown reads 75 volts, and the current meter I indicates a current of 0.05 amp. What is the unknown resistance?

$$R = E \div I = 75 \div 0.05 = 1500 \text{ ohms}$$

The ohmmeter. A simple instrument very commonly used is the direct-reading ohmmeter consisting of an ammeter and a battery. The circuit is shown in Fig. 10–3, where R is the un-

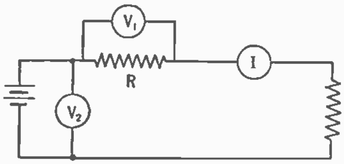


FIG. 9–3. Ammeter-voltmeter method of measuring resistance.

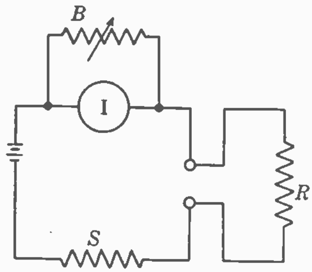


FIG. 10–3. Ohmmeter diagram.

known resistance whose value is to be measured. With R short-circuited, resistor S inside the metal case limits the current taken by the meter from the battery to about full-scale deflection. The deflection is made exactly full scale by means of the variable shunt resistance B across the indicating meter. When R is placed in the circuit, the deflection of the instrument decreases to correspond to the new value of the current flowing. This is, of course, less than it was with the unknown resistance short-circuited, and the meter can be calibrated directly in terms of ohms rather than amperes.

The range of this type of meter is usually considered as 0.1 to 10 times the half-scale reading. Ranges may be 1000, 10,000, and 100,000 ohms. Use of a sensitive instrument or more voltage will permit the measurement of several megohms by this method. The low range on ohmmeters is usually obtained by shunting the unknown resistance across the meter instead of placing it in series with the meter.

Wheatstone bridge. Resistances are often measured by what is known as the comparison method, that is, by comparing them with resistance units whose values are known. For example, we might measure the current through an unknown resistance, R_1, as in Fig. 11–3, and then adjust a variable calibrated re-

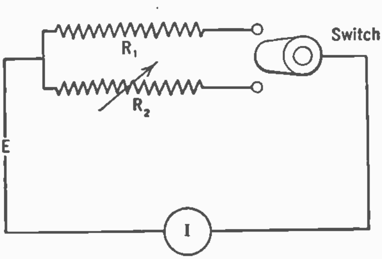


FIG. 11–3. Measuring resistance by comparison.

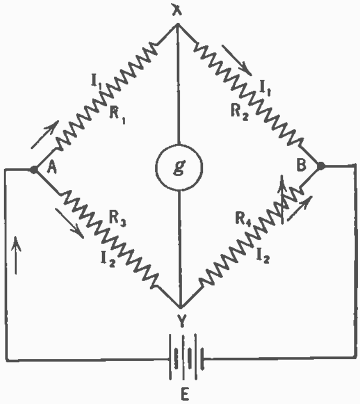


FIG. 12–3. Wheatstone bridge for measuring resistance.

sistance, R_2, until the same current flows under the same emf. The two resistances are then equal in value.

Another method employs a Wheatstone bridge. In diagrammatic form it is represented in Fig. 12–3, in which R_1 and R_2 are fixed resistances whose values are known, R_3 is the unknown resistance whose value is desired, and R_4 is a variable resistance to which the unknown is compared and the values of which are known. The method is as follows. A current is led into the bridge arrangement of resistances at the points A and B, and a sensitive current-indicating meter g is placed between points X and Y. The values of R_1, R_2, and R_4 are adjusted until the meter g shows that no current flows through it; that is, there is no difference in voltage between the two points X and Y which would force current through the meter. In other words X and Y are at the same voltage.

The total current divides at A and flows into the arms of the bridge forming the currents I_1 through R_1 and R_2 and I_2 through R_3 and R_4. If there is no potential difference between X and Y, the voltage drop along R_1 is equal to the voltage drop along R_3.

Thus

$$I_1R_1 = I_2R_3 \qquad [1]$$

Similarly

$$I_1R_2 = I_2R_4 \qquad [2]$$

Dividing equation 1 by equation 2

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} \qquad [3]$$

Suppose that R_1 and R_2 are equal in value. Then equation 3 becomes

$$1 = \frac{R_3}{R_4}$$

or

$$R_3 = R_4$$

and to find the value of the unknown resistance R_3 we need only adjust R_4 (whose values are known) until no current flows through the meter. Then the two resistances are equal. Suppose, however, that the unknown resistance is much larger than any value we can obtain by adjusting R_4. For example, let it be ten times as large. Then it is only necessary to make R_1 ten times as large as R_2, when equation 3 becomes

$$\frac{R_1}{R_2} = \frac{R_3}{R_4} = 10$$

$$R_3 = 10R_4$$

and it is only necessary to adjust R_4 until no current flows through the meter and to multiply the resistance of this standard R_4 by 10 to get the value of the unknown resistance R_3.

Resistances R_1 and R_2 are called the **ratio arms**; R_4, the standard resistance, is usually a **resistance box,** that is, a box in which are a series of resistance units accurately measured and equipped with switch arms so that any value of resistance may be obtained.

Problem 2–3. A soldering iron takes 5 amp when plugged into a 110-volt socket. What is its resistance? What power does it consume?

Problem 3–3. A current of 50 ma flows through a resistance of 2000 ohms. What voltage would be measured if a high-resistance voltmeter were placed across the resistor?

Problem 4–3. A rheostat is a variable resistor, either with a contact arm riding on the resistance wire so that as the arm is moved by a handle any point in the resistance wire can be reached, or with definite connections to one or more points of the resistance. Suppose that the total resistance of a rheostat is 30 ohms and that taps are brought out at the 12-ohm and 20-ohm points. If 1 amp flows through the entire resistance, what voltages are available at the taps?

Problem 5–3. A certain device has a coil wound with wire that will burn up if more than 25 amp flows through it. It has a resistance of 0.06 ohm. What is the highest voltage that can be placed across it without danger?

Problem 6–3. A high-resistance voltmeter is not at hand, and it is desired to measure the voltage output of a power supply system as in Fig. 13–3. When a milliammeter is in series with 20,000 ohms it reads 15 ma; when a voltmeter is shunted across the current meter and the resistance, it reads 270 volts and the milliammeter reads 13.5. What is the voltage output of the device?

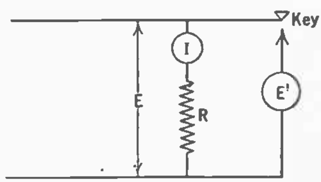


FIG. 13–3. A means of avoiding the use of a high-resistance voltmeter.

Problem 7–3. In Fig. 7–3 consider E as a device (perhaps a rectifier tube) supplying voltage. The device has an internal resistance, R, of 2000 ohms. It is desired to know the actual voltage produced by E. A voltmeter having a resistance of 1000 ohms per volt reading 250 volts full scale measures 240 volts at the point X. What is the value of E?

Problem 8–3. Suppose that R in Problem 7–3 is variable. When the 250-volt meter is placed at X, the terminals of the device, the needle goes off scale indicating that the voltage at X is greater than 250 volts. If, however, R is made equal to 2500 ohms, the needle of the meter indicates that just 250 volts appear across the meter. What is the voltage output of E?

NOTE. This is exactly the function played by a voltmeter multiplier—to reduce the voltage actually appearing at the meter so that the voltage can be read on the meter scale without the needle going off scale.

Problem 9–3. What resistance must be put in series with a 250,000-ohm voltmeter reading 250 volts full scale if it is desired to make it measure 500 volts full scale?

Problem 10–3. A certain microammeter, reading 25 μa full scale, has a resistance of 190 ohms. What voltage is required for full-scale deflection? What value of resistance shunted across it will make it possible to measure 250 μa?

Hint. A good starting point is to remember that the same voltage appears across the shunt resistance as across the meter resistance, and that the sum of the currents taken by these two resistances is equal to 250 μa.

Problem 11–3. A Wheatstone bridge is made up of resistors each having a resistance 10 per cent higher than the value marked on it. Thus, instead of $R_1 = 400$, $R_2 = 100$, and $R_4 = 80$ ohms, the values are actually 440, 110, and 88, respectively. Using these resistors, the bridge is balanced when an unknown resistance is placed at R_3.

What is the value of the unknown resistance?

CHAPTER 4

PRODUCTION OF CURRENT

Electrical energy does not exist in nature in a form useful to man. It must be transformed from some other form of energy. For example, the mechanical energy of a motor or steam engine may be transformed into electrical energy by means of a generator.

The commonest sources of current useful to radio workers are the **battery** and the **generator.** The battery is a device which converts chemical energy into electrical energy; the generator converts mechanical energy to electrical energy.

Batteries. A **battery** is made up of one or more units called **cells.** The essentials of the cell are three: two conductors called **electrodes,** usually of different materials; and a chemical solution known as the **electrolyte,** which acts upon one of the electrodes more than it does upon the other. In this action, one of the electrodes is usually eaten up, and when this conductor, usually a metal, is gone, the battery is exhausted; it must be thrown away or the metal replaced. If the metal can be replaced by sending a current through the cell from some outside source, that is, by reversing the process through which the cell was exhausted, the cell is known as a **secondary** or **storage** cell. A cell that must be thrown away when one of the electrodes is eaten up is called a **primary** cell. The dry cell is a well-known example.

Experiment 1–4. If a plate of copper and a plate of zinc are placed in dilute sulphuric acid and a sensitive meter is placed across the terminals as shown in Fig. 1–4, a voltage of definite polarity will be indicated. The positive terminal of the voltmeter must be placed on the copper plate so that the meter needle will move in the proper direction. The copper plate is therefore positive; the zinc is negative. If an external wire is attached to the plates, a current flows; the zinc is slowly dissolved, hydrogen bubbles appear at the copper plate, and finally the voltage of the cell falls off. Other combinations of metals should be tried.

Another instructive experiment is shown in Fig. 2-4 proving that water is made up of two atoms of hydrogen and one atom of oxygen. Pass a current through two platinum electrodes immersed in sulphuric acid using a battery of 10 or 12 volts. Hydrogen comes off the electrode connected to the negative battery terminal; oxygen, at the other electrode. If the volumes of these gases are measured there will be twice as much hydrogen as there is oxygen.

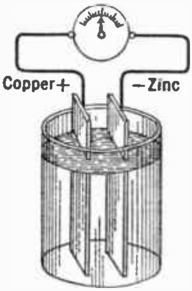


Fig. 1-4. A simple primary cell.

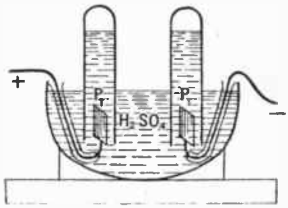


Fig. 2-4. An experiment in electrolysis.

The emf of a primary cell depends upon the nature of the electrolyte and the materials from which the plates or electrodes are made. Copper and zinc plates immersed in a solution of dilute sulphuric acid will give an emf of about 1 volt regardless of the size of plates or their distance apart. Zinc and carbon plates in chromic acid give an emf of about 2 volts.

Before the plates are connected externally by a conductor a difference of electrical potential exists between the two electrodes but no current flows. This voltage is known as the emf of the cell. As soon as the plates are connected, current flows and the destruction of the zinc begins. When all the zinc is destroyed the cell is dead.

Dry cell. The common dry cell is a very familiar source of emf. It is widely used in portable radio equipment and in some field telephone sets, or in any service where the current required and the periods of current drain are not very great.

The dry cell is made up of a zinc container within which is an electrolyte of sal ammoniac mixed with some porous material like manganese dioxide and powdered carbon. In the center of this paste is a carbon rod which forms the positive terminal. The zinc case is the negative terminal. The emf is 1.5 volts, and the cell can be used until the emf falls as low as 1.13 volts or even lower. In use, the zinc case is consumed by the electrolyte, and when the case is badly eaten away the cell must be discarded, as the moisture will dry out of the electrolyte by being exposed to the air.

The life of the dry cell depends, naturally, upon the rate at which power is taken from it. On the shelf, no power being required from the cell, the life is approximately 12 months. The life during use depends upon the current required. Thus a typical small B battery will deliver 54.7 weeks of service at a 3-ma drain 5 hr per day 5 days per week, but if the current is increased to 50 ma the cell is ready for discard at the end of 1.4 weeks.

Cells will not deliver full voltage at low temperatures, but freezing does not harm them. The common practice of determining the condition of a dry cell by placing an ammeter across the terminals shows very little, except the uniformity of a batch of cells.

Dry cells are made in several sizes, including very small units which are built up into sources of fairly high voltage by placing a number in series. A 45-volt B battery has 30 cells in it, each delivering 1.5 volts.

Storage battery. There are two types of storage cells: (1) a cell using sulphuric acid as the electrolyte with a positive plate of lead peroxide and a negative plate of spongy lead; and (2) the nickel-iron alkaline cell (Edison). The lead battery is by far the commonest, every automobile having one as its source of electric power to start the motor. The great virtue of the storage battery compared to the dry battery is that it may be recharged when it is run down. On the other hand, it is bulky, heavy, and expensive.

Alkaline cells stand mechanical shock better than lead cells, but they cost more. They have longer life than lead cells if both are handled properly, and they weigh less and occupy less space than lead cells of the same capacity.

If the lead battery is carefully treated, not over-discharged before it is recharged, it will last from five to ten years. If it is abused or subjected to high current drains, however, its life may be much shorter.

The lead-acid cell delivers, fully charged, a voltage of approximately 2.1 volts; the alkaline cell, about 1.45 volts. A short circuit is harmful to the lead cell but not to the alkaline cell. The lead cell, however, will deliver high currents in emergencies.

Alkaline cells may be stored in a discharged and short-circuited condition. Lead cells should be recharged immediately if they have been accidentally short-circuited, and they should be stored fully charged. A charged lead cell will freeze at $-61°$ F.; a discharged cell, at $+18°$ F.

The condition of charge of a lead cell is best tested by means of a **hydrometer,** a device for measuring the specific gravity (weight per unit volume) of the electrolyte. Since the electrolyte is heavier in a fully charged battery than in a discharged one, the specific gravity is a measure of the amount of electrical energy remaining in the battery. During discharge the sulphur leaves the acid and combines with the lead in both plates to become lead sulphate; during charge the sulphur is driven out of the plates and back into the solution. These reactions account for the variation in specific gravity.

A voltmeter reading under load is an important test on a storage cell, and both the hydrometer and voltmeter test have to be made for conclusive results. A hydrometer test alone will not indicate the presence of a short-circuited cell caused by loose active material touching the plates—a common cause of termination of the cell life.

Since the individual cells of a lead-acid battery have a voltage of 2, common battery voltages of 6, 12, and 24 are made up of the necessary number of cells connected in series.

Battery capacity. The term ampere-hour is used to express the amount of electricity taken from, or put into, a battery. As its name indicates, an ampere-hour represents a current of 1 amp flowing for a period of 1 hr. A 100-amp-hr storage battery will (theoretically, at least) deliver 1 amp for 100 hr or 100 amp for 1 hr. The ampere-hour capacity of batteries decreases as the current drain increases. For example, the B battery mentioned above as having a life of 54.7 hr at a drain of 3 ma and a life of 1.4 hr at a drain of 50 ma delivers 4.1 amp-hr of electricity at the lower drain but only 1.7 at the 50-ma drain.

Internal resistance. One might think that an unlimited current could be secured from a battery if it were short-circuited. This is not true. A low-resistance ammeter placed across a dry cell gives a definite reading—it is not unlimited. Something must be in the circuit which has a resistance greater than that of the ammeter or the connections. For example, a new dry cell will deliver about 30 amp through wires of very low resistance.

This something which restricts the current to a limited value is the **internal resistance** of the cell. This resistance depends upon the construction of the cell, its electrode and electrolyte material, the distance apart of the electrodes, the condition of the cell—whether new or old. The older the cell the smaller the area of electrode in contact with the electrolyte and the greater the resistance. The current delivered by a cell is

$$I = \frac{E}{r + R}$$

when r = internal resistance of cell.
R = external resistance of circuit.

Cells which have a large internal resistance deliver but small currents; low-resistance cells deliver large currents.

When one tests a dry cell with an ammeter, he is actually ascertaining the condition of the cell by measuring the internal resistance. When the cell gets old or has been exhausted because of too heavy currents, its internal resistance becomes high and an ammeter reads only small current when placed across it.

The storage cell is a very low-resistance device. Its terminal voltage when fully charged is about 2.1 volts, and it has a resistance of about 0.005 ohm. Placing an ammeter across such a cell is dangerous. The meter will probably be ruined.

Example 1–4. A dry cell on short circuit delivers 30 amp. Its terminal voltage is 1.5 volts. What is its internal resistance?

By Ohm's law,

$$I = E \div r$$

$$30 = 1.5 \div r$$

$$r = 1.5 \div 30 = 0.05 \text{ ohm}$$

Example 2–4. The emf of a battery is 5 volts. When 100 ohms is placed across it the voltage V falls to 4 volts. What is the internal resistance of the battery?

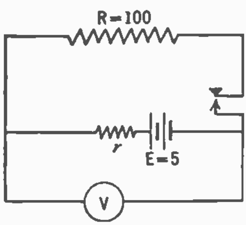


Fig. 3–4. A problem in internal resistance.

In Fig. 3–4, R = 100 ohms, r = the internal resistance of the battery, the voltage drop across the 100 ohms is 4 volts, which leaves a 1-volt drop in the internal resistance of the cell. The current through the external 100-ohm resistance is, by Ohm's law,

$$I = 4 \div 100 = 0.04 \text{ amp}$$

This current must also flow through the internal resistance of the battery, and there it causes a voltage drop of 1 volt.

$$E = I \times r$$

$$1 = 0.04 \times r$$

$$r = 1 \div 0.04 = 25 \text{ ohms}$$

Note that the internal resistance of the cell is represented as being in series with the voltage and the external resistance. This is because the current must actually flow through the internal resistance of all such voltage generators, and hence the resistance of the device is represented in series with the remainder of the circuit. Care must be taken in such representations not to place the voltmeter in the wrong place. In Fig. 3–4 the voltmeter is actually placed across the battery and its internal resistance, which in turn are connected directly to the 100-ohm resistance.

The voltage of the cell on open circuit is its **emf.** Under load the voltage falls and is then labeled as the **pd** (potential difference). The emf of high-resistance cells can be measured only by high-resistance meters, those that take but little current from the cell. When the voltage of a cell or battery is mentioned its emf is assumed unless otherwise stated.

Polarization. All cells suffer from **polarization.** In polarization the internal resistance of the cell is increased, usually by products of the internal chemical reaction which tend to insulate the electrodes.

For example, a dry cell which has had large currents taken from it becomes polarized and at the moment will not deliver any further large currents.

Various means are taken to overcome the bad effects of polarization. Chemicals may be put into the cell to supply oxygen, which will combine with the hydrogen to form water; shaking the cell may remove hydrogen bubbles which cause the high internal resistance. The manganese dioxide employed in dry cells is a depolarizer.

If a polarized cell is removed from service, the chemicals placed in it do away with the polarization products. The cell "recuperates" and can be employed again. It is for this reason that dry cells are used for intermittent service only; when a continuous current is needed, other kinds of cells are chosen.

Sooner or later, however, the internal resistance of the cell becomes so high that the required current cannot be secured from it, and the cell must be thrown away.

Cells in series. Cells and batteries may be connected in several ways. When the positive terminal

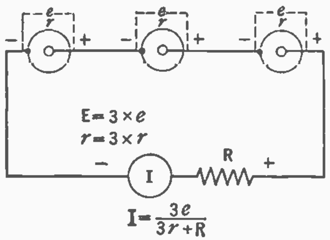


FIG. 4-4. Cells in series. Lower-case letters indicate internal resistance and voltage of individual cells.

of one cell is connected to the negative terminal of the next cell, as in Fig. 4–4, the cells are said to be connected in **series.** Under these conditions, the voltage appearing at the two ends of the series of cells is the sum of the individual cell voltages. For example, if we connect four dry cells in series, each cell having a voltage of 1.5, a voltmeter across the two ends will register 6 volts. At the same time the total internal resistance is the sum of the individual resistances, and whatever current flows must flow through this resistance.

If the ends of the battery are connected with a wire of resistance R the current that will flow may be obtained by Ohm's law as

$$I = \frac{Ne}{Nr + R}$$

where N = number of cells.
e = voltage of each cell.
r = internal resistance of each cell.

Cells in parallel. When the positive terminals of the cells are connected together, and the negative terminals are connected together, as in Fig. 5–4, the cells are said to be connected in **parallel.**

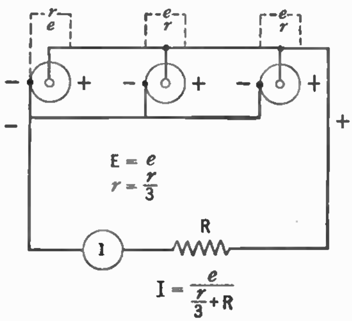


FIG. 5–4. Cells in parallel.

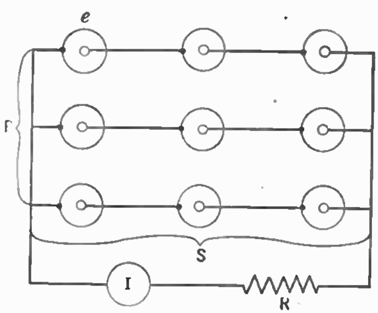


FIG. 6–4. Cells in series-parallel.

Under these conditions the terminal voltage of the combination is the same as the terminal voltage of each cell, but the internal resistance has been divided by the number of cells, N, and has become $r \div N$.

If the ends of the battery are connected with a wire whose resistance is R, the current that will flow is

$$I = \frac{e}{(r \div N) + R}$$

Cells may also be connected in what is called a **series-parallel** arrangement. In Fig. 6–4 are P sets of S cells in series, and the sets themselves are connected in parallel.

If the battery of cells shown in Fig. 6–4 is connected to a wire whose resistance is R, the current that will flow is

$$I = \frac{Ne}{RP + Sr}$$

where $N = P \times S$.

Problem 1–4. A military expedition is to have 100 receivers each requiring 90 volts at 15 ma. Individual B battery cells deliver 1.5 volts, and at the 15-ma drain they have a life of 780 hr. If each 90-volt battery costs $2.00, estimate the cost for energy per kilowatt-hour. If the individual cells have a diameter of 1.25 in. and are 4 in. high, estimate the minimum cargo space required. Each cell weighs 0.366 lb. What total weight must be carried?

Problem 2–4. A radio transmitter requires a minimum filament voltage of 10. Lead-acid cells have a fully charged voltage of 2.05 and a discharge voltage of 1.75; corresponding voltages for alkaline cells are 1.45 and 1.0.

How many lead-acid cells or how many alkaline cells are required?

If the filament current is 3.25 amp, what is the maximum value of a variable resistance to be placed in the battery-filament circuit?

CHAPTER 5

MAGNETISM AND ELECTROMAGNETISM

The second common source of electric power is the generator. Since the generator is made possible by electromagnets, we must look carefully into the phenomenon of magnetism.

Magnetism. Everyone is familiar with the common horseshoe magnet. It is a piece of hardened steel which has been magnetized and, as is well known, will attract bits of iron and steel. One end is identified as the north pole, the other as the south pole. Two magnets will repel each other if their north or south poles are brought together, and they will attract each other if the north pole of one magnet is brought near the south pole of the other. The amount of attraction or repulsion follows the same law as electric charges: the attraction or repulsion is proportional to the strength of the individual magnets and inversely proportional to the square of the distance between them. Thus

$$F = \frac{m_1 m_2}{d^2}$$

Soft iron loses its magnetism easily and is not used for permanent magnets. Instead, hardened steel or certain alloys of nickel and iron, such as Alnico which is composed of aluminum, nickel, cobalt, and iron, are employed when magnetism must be retained over a long period.

Terrestrial magnetism. Everyone knows that the earth is a great magnet, but no one knows precisely why. It is thought that the earth's core is iron and that it might have become magnetized in some manner. But the core is supposed to be very hot, perhaps molten, and iron loses its magnetism when heated. On the other hand, the magnetic poles shift about, a fact that is consistent with a molten core but not with a solid one.

We know that disturbances to wire and radio communication occur during periods of intense sunspot activity. These disturbances are known as magnetic storms; and it is quite likely that the earth's magnetism is partly due to a magnetized core and partly to the effects of streams of electrons and protons produced in the sunspots.

Currents flowing near the earth's surface during these storms produce intense voltages in cables and mask the signals; radio propagation suffers from the same causes.

Electromagnetism. One of the most important discoveries of all science was that a coil of wire carrying an electric current acts like a magnet. This action can be demonstrated easily by distributing some fine iron filings on a sheet of cardboard or glass placed near a magnet. It will be found that the filings orient themselves along lines that are assumed to leave the north pole of the magnet and enter the south pole. Similarly, a coil of wire carrying an electric current will be found to have magnetic lines of force like those of a permanent magnet.

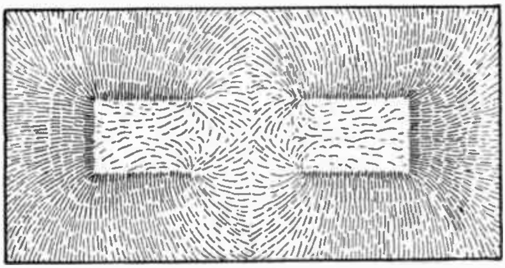


FIG. 1–5. How the lines of force about a bar magnet are located.

The existence of this magnetic field made up of lines of force can be demonstrated by bringing an ordinary mariner's compass near a wire carrying a current. The compass needle will swing from its north-south direction. If the coil of wire is wound on a hollow form, say a cardboard cylinder, and if a steady current flows through the wire, it will be found that the magnetic effect of the coil is increased very appreciably when an iron core is placed within the coil form, for the reason that magnetic lines of force are carried with much less opposition through iron than through air.

Magnetic quantities. What is it that makes a magnet attract pieces of iron? What exists in the space surrounding a magnet? Anyone asking these questions is squarely up against a phenomenon of nature. One might as well ask, "How high is up?" As yet no one knows what magnetism is. The best we can do is to accept it as a fact and to learn all that is necessary about it to make proper use of it.

We say that in the space surrounding a magnet is a magnetic **field** made up of lines of force. These lines take their place as indicated by the iron filings. Their position can be explored by means of the compass needle also, as well as in other ways. In any magnetic circuit, certain quantities must be known in

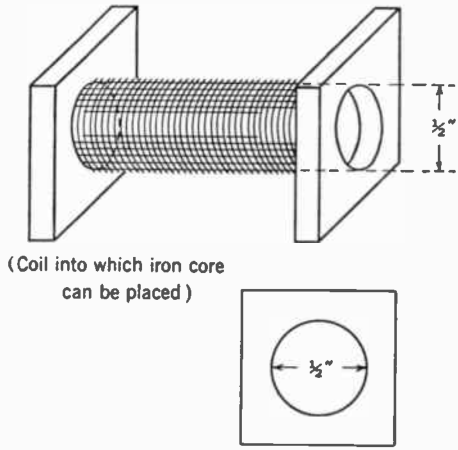


FIG. 2–5. Wire wound on a hollow core to demonstrate that a wire carrying current acts like a magnet.

order to design, properly, apparatus using the circuit. For example, the strength of the magnetic field may be expressed by the total number of lines of force, or by the number of lines per unit of area, that is, the number of lines that go through a square inch or a square centimeter. The total number of lines is known as the **flux**, and the lines per square centimeter as the

flux density. Flux density, therefore, is the total flux divided by the area, in square centimeters, in which the lines exist.

Another quantity that we must know is the magnetic force required to set up a given number of lines in a material having a given magnetic quality, just as, in an electric circuit, we must know the voltage required to produce a given current through a material of a given resistance.

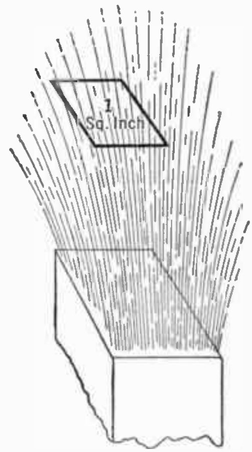


FIG. 3-5. The number of lines of force going through 1 sq in. or other unit of area is known as the flux density.

Magnetomotive force. Corresponding to emf in an electric circuit is magnetomotive force (mmf) in a magnetic circuit; and corresponding to resistance is the magnetic quantity **reluctance.** The unit of mmf is the **gilbert;** it is the magnetic "pressure" required to produce 1 line of force in a circuit having a reluctance of 1 unit. (There is no other name for this unit of reluctance.)

In an electric circuit, the relations between the "cause" and the "effect" are

$$\text{Current } I = \frac{\text{Electromotive force } E}{\text{Resistance } R}$$

In a magnetic circuit, the corresponding relations are

$$\text{Flux } \Phi = \frac{\text{Magnetomotive force } \mathfrak{F}}{\text{Reluctance } \mathcal{R}}$$

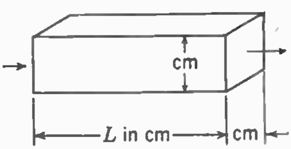


FIG. 4-5. Basic volume for computing magnet circuits. If $L = 1$ cm, the volume is a "centimeter cube."

Consider a centimeter cube of air, that is, a cube of air 1 cm on a side. If a mmf of 1 gilbert is applied to opposite faces, 1 line of force will go through the cube. The reluctance of this cube of air, therefore, is 1 unit. If the cube is replaced by a column of air 1 cm wide, 1 cm deep, but 2 cm long, the reluctance will be 2 units. If, however, the cross-sectional area is increased, the reluctance will be decreased. Similarly, the resistance of a piece of copper wire 1 ft long and having a cross-sectional area of 1 circular mil is 10.4 ohms. A wire 2 ft long will have twice this resistance, and if the cross-sectional area is increased (by using a larger wire) the resistance will be decreased.

Now, in electric circuits, all materials do not have the same resistance as copper: most materials have more resistance; silver has less. How do we express the intrinsic or specific resistance of these materials? We merely state the resistance per circular mil-foot and call this the resistivity.

In a magnetic circuit, all materials do not have the same reluctance. We can measure the reluctance per centimeter cube and call this the **reluctivity**. The reluctance of any larger or smaller piece of the material can be figured from its reluctivity. Thus

$$\text{Reluctance} = \frac{\text{Reluctivity} \times L}{A}$$

where reluctivity = reluctance per centimeter cube;
= 1 for air.
L = length in centimeters of the magnetic circuit.
A = cross-sectional area in square centimeters of magnetic circuit.

Example 1–5. What is the reluctance of an air gap having a square cross section 1.5 cm on a side and a length of 1.6 cm?

Solution.

$$\text{Reluctance} = \frac{\text{Reluctivity of air} \times \text{Length of gap}}{\text{Cross-sectional area of gap}}$$

$$= \frac{1 \times 1.6}{1.5 \times 1.5} = \frac{1.6}{2.25}$$

$$= 0.71 \text{ unit}$$

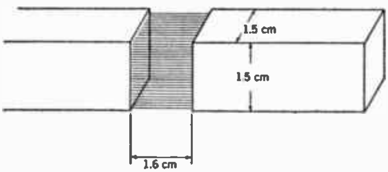


FIG. 5–5. Computation of reluctance of air gap in Example 1–5.

Now how much magnetomotive force is required to produce a given flux in a given air gap? In a magnetic circuit, the relation between mmf, flux, and reluctance is

$$\text{Magnetic flux} = \frac{\text{Magnetomotive force}}{\text{Reluctance}}$$

or, symbolically, $\Phi = \mathfrak{F} \div \mathcal{R}$.

Example 2-5. In the air gap of Example 1-5, how many lines of force will be produced by a mmf of 320 gilberts?

Solution.

$$\begin{aligned} \Phi &= \mathfrak{F} \div \mathcal{R} \\ &= 320 \div 0.71 \\ &= 450 \text{ lines of force (approximately)} \end{aligned}$$

Magnetic calculations are never as accurate as similar electrical calculations, where the current can be very accurately and precisely directed by conductors. The reason is that lines of force cannot be exactly and precisely directed and controlled; some of them escape through paths which cannot be calculated.

Note that reluctivity and reluctance are related but are not the same thing, just as resistivity and resistance are related but are not the same in an electric circuit. Reluctivity always refers to a given geometrical configuration, viz., the centimeter cube. The same material drawn out into a fine wire would have much greater reluctance than it had in the form of the centimeter cube. The analogy here to an electric circuit is clear. The resistivity of a centimeter cube of copper is a fixed and known amount. But the resistance of this amount of copper depends entirely upon how much it is drawn out.

Magnetic field in a solenoid. If a coil of wire is wound on a core of air of such dimensions that the length is greater than the diameter by a factor of 10 or more, and if a current is passed through the wire, it will be found that the number of lines of force (flux) passing through the center of the coil depends upon the number of turns of wire and the magnitude of the current through the wire. It also depends upon the reluctance of the air inside the coil, of course. Thus

$$\text{Flux} = \frac{\text{Magnetomotive force}}{\text{Reluctance}}$$

$$= \frac{0.4\pi NI}{\mathcal{R}} = \frac{1.26NI}{\mathcal{R}}$$

where N is the number of turns in the coil.
I is the current in amperes.
0.4π is a proportionality factor.

The product of N and I is known as the ampere-turns in the coil, and thus

$$\text{Flux } \Phi = \frac{1.26 \times \text{Ampere-turns}}{\text{Reluctance}}$$

Now the reluctance is equal to the reluctivity times the length of the material in centimeters divided by the cross section of the

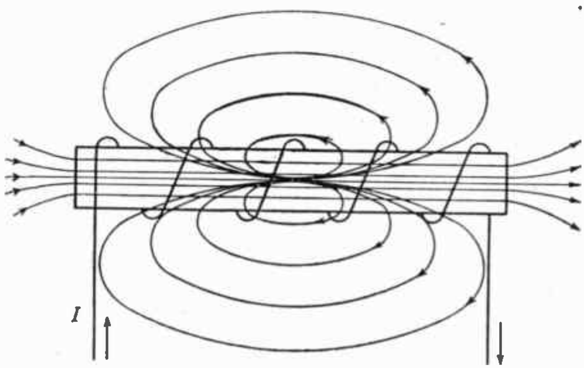


Fig. 6–5. How lines of force make up the field about a solenoid. All lines go through the center of the coil.

material in square centimeters. Here the material on which the coil is wound is air and the reluctivity is equal to 1. The flux at the center of a coil of solenoidal dimensions is

$$\Phi = \frac{1.26NIA}{L}$$

where A = cross-sectional area in square centimeters.
L = length in centimeters.

One-half of the total number of lines of force pass out of the ends of the coil, and the other half constitute the leakage flux. The flux passing out the ends and the sides returns to the coil through the air as though it were passing through a material of zero reluctance.

The mmf produced by a current of I amperes flowing through a coil which is long compared to its diameter and having N turns is equal to $1.26NI$ gilberts.

For the special condition that the coil has circular cross section,

$$A = \pi r^2 = \tfrac{1}{4}\pi d^2 = 0.7854d^2$$

where r = radius of core.
d = diameter of core.

Then

$$\Phi = \frac{0.989NId^2}{L}$$

Problem 1–5. What is the flux produced in the center of a coil of 2000 turns of wire wound on a thin cardboard tube 2 cm in diameter and 100 cm long if 0.5 amp flows through the wire?

Problem 2–5. A coil has a square core 4 cm on a side. It is 64 cm long. How many turns of wire must be put on it if the flux at the center is 1000 lines with a current of 4 amp?

Problem 3–5. How many gilberts are necessary to produce a flux of 2000 lines across an air gap 0.25 cm long and having a cross section of 8 by 4 cm?

Permeability. The reluctivity of a material is seldom used in design work; in its place the term permeability is employed. This is the number of lines of force that will be produced in a centimeter cube of a material when the mmf between opposite faces is 1 gilbert. The statement that the permeability of a material is 2000 means simply that 2000 times as many lines of force will be produced in it by a given mmf as will be produced in air. The permeability of air is 1; that of certain alloys of iron may be as high as many thousand.

Example 3–5. A bar of material having a permeability of 400 has a length of 6 cm and a cross-sectional area of 4 sq cm. Along the longer dimension is a mmf of 0.24 gilbert. How many lines of force (at the center) will be produced in this bar?

Solution. First disregard the fact that the bar is made of iron. Consider an air core of the same dimensions. The flux produced in it will be

$$\Phi = \mathfrak{F} \div \mathcal{R}$$

Now, $\mathcal{R} = L \div A$ for air; therefore

$$\Phi = \frac{\mathfrak{F}}{L \div A}$$

$$= \frac{\mathfrak{F} \times A}{L}$$

$$= \frac{0.24 \times 4}{6}$$

$$= 0.16 \text{ line in air}$$

Now we know that 400 times as many lines will appear in the iron as in air, since the permeability is 400. Therefore

$$\Phi \text{ in iron} = 0.16 \times 400 = 64 \text{ lines}$$

In solving problems in magnetic circuits, one must remember that all the fundamental quantities are referred to a unit cube of air. Thus 1 gilbert of mmf is required to produce 1 line of force through a length of 1 cm and across 1 sq cm of air, as the permeability of air is 1. If a material has a permeability of 400, 1 gilbert will produce 400 lines per square centimeter of cross-sectional area in a piece 1 cm long. The simplest solution, therefore, is to reduce the magnetic circuit to its equivalent in air, that is, to determine the cross-sectional area of the material in square centimeters and its length in centimeters. Determine the reluctance, and either calculate the flux resulting from a given mmf or calculate the mmf required to produce a given flux, depending on what the problem is. Then multiply the resultant flux by the permeability of the material, or divide the mmf required by the permeability. Time can be saved by immediately dividing the reluctance by the permeability.

The Greek letter mu, μ, is often used as a symbol for permeability.

If the circuit is made up of two materials having different reluctances, each portion of the circuit must be solved separately. Thus, if an iron core has an air gap in it, we must determine the number of gilberts necessary for the iron portion, and then the gilberts necessary for the air portion, and add them together to determine the total magnetizing force, $\mathfrak{F}$. If two or more magnetic paths are in series, the total reluctance will be the sum of the individual reluctances.

If two or more magnetic paths are in parallel, the total reluctance will be the reciprocal of the sum of the reciprocals of the individual reluctances.

$$\text{Reluctance in series} = \mathcal{R}_1 + \mathcal{R}_2$$

$$\text{Reluctance in parallel} = \frac{1}{\mathcal{R}} = \frac{1}{\mathcal{R}_1} + \frac{1}{\mathcal{R}_2} + \frac{1}{\mathcal{R}_3}, \text{ etc., or } \frac{\mathcal{R}_1\mathcal{R}_2}{\mathcal{R}_1 + \mathcal{R}_2}$$

when there are only two reluctances.

Here again the reader should note the similarity to the electric circuit where resistances in series and parallel are combined in exactly the same way.

Problem 4–5. A piece of iron with a permeability of 400 has a cross section of 2 sq cm and a length of 10 cm. What is the reluctance?

Problem 5–5. Silicon steel used in the core of a small power transformer has a permeability of 400. What flux will 240 ampere-turns produce in a magnetic circuit made up of this material if the circuit has a circular cross section with a diameter of 3 cm and is 9 cm long?

Problem 6–5. A cylinder of iron with a permeability of 500 is bent into the form of a ring with a small air gap in it. That is, the two ends of the iron do not quite touch. The iron before bending is 40 cm long; its cross section is 4 sq cm. The air gap is 2 mm. If 0.01 amp is to pass through a coil of wire wound on this core, how many turns will be necessary if a flux of 250 lines is to be set up? (See Fig. 7–5.)

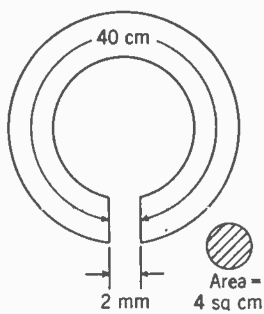


Fig. 7–5. Iron bar bent into an incomplete circle. See Problem 6–5.

Hint. Determine the reluctance of the two parts of the circuit separately. Add to determine the total reluctance, and then determine the mmf, remembering that 1 gilbert = 1/1.26 ampere-turns.

Problem 7–5. In Problem 6–5, what flux would have been produced if there had been no air gap?

Flux density. Since magnetic materials of all shapes and sizes are used, a means of comparing their magnetic properties is necessary. For example, the force necessary to produce a given flux depends upon the length of the circuit, its cross-sectional area, and its permeability. How can one compare two different materials of different dimensions? The method is to reduce everything to unit dimensions. Thus we use the term, **flux density** (B), which is the number of lines per square centimeter. The unit for flux density is the **gauss,** defined as the number of flux lines per square centimeter.

The flux density of any material, therefore, is the total flux Φ, divided by the area

$$B = \Phi \div A$$

where B = flux density in gausses.
Φ = the total number of lines.
A = cross-sectional area in square centimeters.

Another unit is the magnetomotive force required to produce a given number of lines through a 1-cm length of a given material. This is called the **magnetizing force.** The unit is the **oersted,** defined as 1 gilbert per centimeter. It is equal to the total mmf in the circuit divided by the length of the magnetic circuit. Thus

$$\text{Magnetizing force, } H = \frac{\mathfrak{F}}{L}$$

where H is in oersteds,* $\mathfrak{F}$ is in gilberts, and L is in centimeters.

Problem 8–5. What is the flux density in gausses across an air gap which is square in cross section, 2 cm on a side, if the total flux is 2000 lines?

Problem 9–5. A bar of high-permeability ($\mu = 2000$) alloy is 40 cm long, with a cross-sectional area of 4 sq cm. A flux of 4000 lines is produced in it. What magnetizing force is required? What is the flux density?

* Up to 1930, the oersted was used as the unit of reluctance. Older books, therefore, will use the oersted in this manner.

Problem 10–5. In Fig. 8–5 is a drawing of a transformer core. On it are wound 500 turns of wire through which 2 amp passes. If the reluctivity of the core material is 2.0×10^{-3}, find the flux, the flux density, the magnetizing force, and the permeability.

***B-H* curves.** If a given piece of magnetic material is subjected to varying magnetizing forces, the effects produced, expressed in lines per square centimeter (gausses), may be plotted

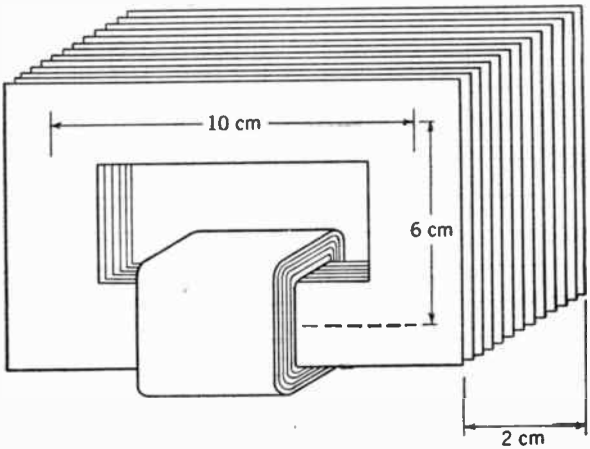


FIG. 8–5. Transformer core and winding—a typical problem in magnetic circuits.

as a curve. Such a curve is called a *B-H* curve, a saturation curve, or simply a magnetization curve.

Since the permeability, μ, of a material is the number of lines of force per unit area set up by 1 unit of magnetizing force, then

$$\mu = B \div H$$

Since B is the flux per square centimeter, the total flux through any area is the product of B and the area. Thus

$$\Phi = BA$$

From a typical *B-H* curve, one can find the permeability of the material at any magnetizing force by dividing a value of B

by the corresponding value of H. It can be seen at once that the permeability is not a constant but varies with the amount of magnetization. At values of B of the order of 100 or less, the permeability is fairly constant, but at higher values of flux density, permeability varies widely. Engineers are supplied by

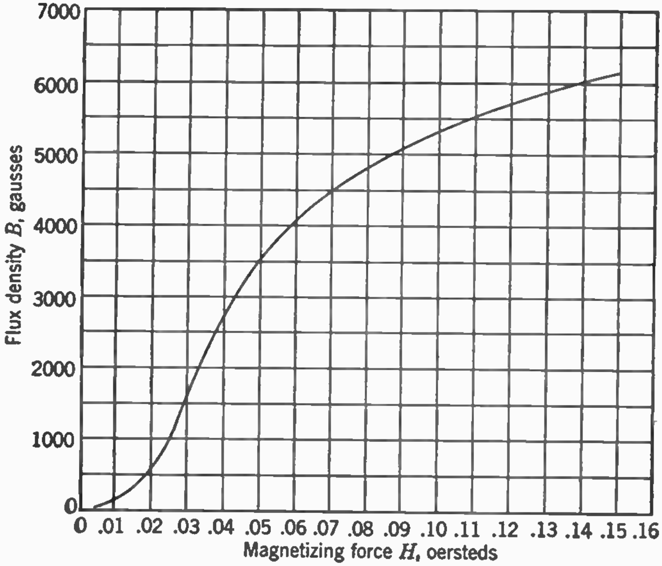


FIG. 9–5. Magnetization curve for Hipernik, a modern magnetic material.

manufacturers of magnetic material with B-H curves so that they can tell the value of H required to produce any value of B.

To calculate the mmf required to produce a given number of lines of force through a given length of material, merely multiply the value of H in oersteds (which is the mmf required to produce a given number of lines through a 1-cm length of the material) by the length in centimeters. Thus

$$\mathfrak{F} = HL$$

where $\mathfrak{F}$ is in gilberts, H is the magnetizing force in oersteds, and L is the length in centimeters.

These several magnetic quantities are summarized in the table.

Quantity	Electrical Analogue	Symbol	Unit	Definition
Magnetomotive force	Electromotive force	$\mathfrak{F}$	Gilbert	1 line through a reluctance of 1 unit
Flux	Amperes	Φ	Lines of force	Total number of lines $= \frac{\mathfrak{F}}{\mathcal{R}}$
Reluctivity	Resistivity			Reluctance per centimeter cube
Reluctance	Resistance	$\mathcal{R}$		Reluctivity $\times L \div A$
Permeability	Conductivity	μ		$1 \div$ Reluctivity $= B \div H$
Magnetizing force	Volts per centimeter	H	Oersted	1 gilbert per centimeter $= \mathfrak{F} \div L$
Flux density	Amperes per square centimeter	B	Gauss	1 line per square centimeter $= \Phi \div A$

Magnetomotive force ($\mathfrak{F}$) and magnetizing force (H) are **causes** producing the **effects** of flux (Φ) and flux density (B) through the opposing quantity reluctance ($\mathcal{R}$).

Saturation. Studying typical B-H curves shows that, for low values of magnetizing force H, the flux density rises rapidly; at higher values of H the flux density rises less rapidly, and a point is reached where further increase in H produces no noticeable increase in flux density. At this point the material is said to be **saturated.** Above this point we can obtain still greater flux densities only by applying much larger values of H.

Hysteresis. When iron or other magnetic material is operated in an a-c circuit, the magnetizing force is continually changing: increasing to a maximum, then reversing its direction and increasing to a maximum in the opposite direction. If we measure the flux density for a complete cycle we will obtain a curve like that in Fig. 10–5. Note that the material initially has no magnetization, and that, as we increase H, B increases steadily, then more slowly up to a final point. Now if the magnetizing current is reversed in direction and decreased in value, the values of flux density will not retrace the upward curve but will trace out a new curve, and when the point is reached at which H is equal to zero some flux remains in the material. To reduce the value of B completely to zero we must apply considerable magnetizing force in a direction opposite to that which produced the original

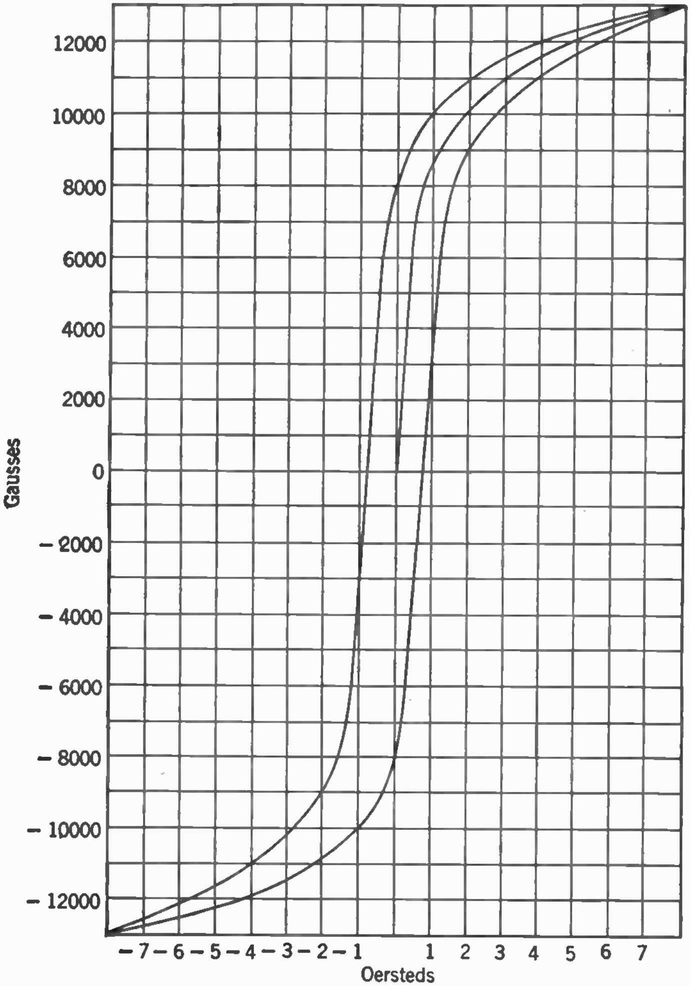


FIG. 10-5. Typical hysteresis curve. Note that as H in oersteds is increased, B in gausses does not increase in proportion but at a slower rate and that as H is decreased, B follows a different curve.

flux density. The flux density B seems to lag behind the magnetizing force H.

This lagging effect is called **hysteresis;** it is a result of the fact that not all the energy used to produce the flux is usefully employed. The area of the hysteresis curve is actually a measure of the work done in overcoming the molecular resistance to

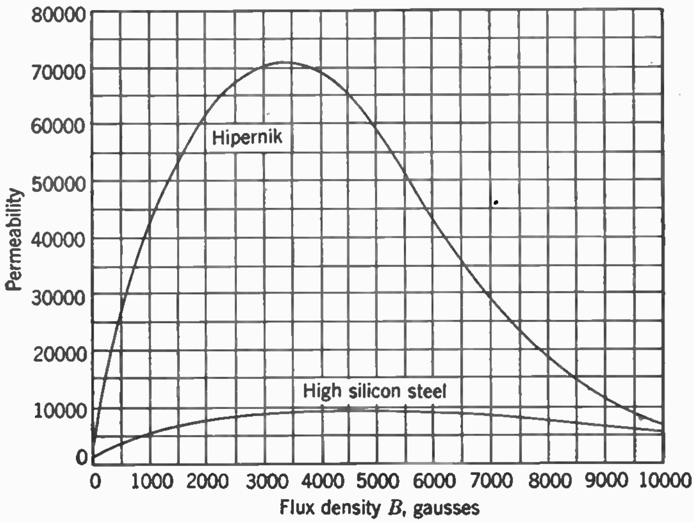


FIG. 11-5. Comparison of two types of magnetic core materials, showing greater permeability of the alloy Hipernik.

magnetization. This power loss depends upon the material, the flux density, and the frequency of the alternating current. Since it increases with frequency, care must be taken to choose magnetic materials properly for a job. Much research has been expended to develop magnetic materials having low hysteresis losses and other desirable properties.

Magnetic alloys. Combinations of iron, nickel, cobalt, and aluminum in various proportions produce magnetic alloys of most interesting and useful properties. Permalloy, for example, has a very high permeability, a magnetizing force of 0.06 oersted producing a field strength of 6000 gauses. Alnico is an alloy widely used for permanent magnets for loud speakers, microphones, and relays. Hipernik is a new material used for power transformers where minimum weight and space are important.

Magnetic shielding. Since lines of force "prefer" to flow through iron rather than through air, it is possible to protect a piece of apparatus from a magnetic field by enclosing it in an iron box. Since the total number of lines of force through the space occupied by the object (a measuring instrument, for example) is fixed, placing an iron box around it merely shifts the path of the lines so that they thread through the iron rather than through the space in which the object lies. Since the air is still a medium in which flux may exist, it is not possible to exclude all the lines of force from the object, but the number of them that go through it can be greatly reduced by the iron box. Such an iron box is called a **screen** or **shield.**

Faraday's discovery. Another very important discovery that helped bring about the development of present-day electrical machinery was made by the celebrated English experimenter Faraday. The experiment may be performed by anyone who has a coil of wire, a bar magnet, and a sensitive current indicator such as a galvanometer or a magnetic compass. The coil should have such dimensions that the bar magnet may be thrust into its center. It will be found that, if the terminals of the coil are connected to the galvanometer, a current will be indicated when the magnet is thrust into the coil; that no current flows when the magnet is stationary; that a current opposite in direction to the first is generated when the magnet is removed from the coil; and that the magnitude of the current depends upon the **rate** at which the magnet is pushed into the coil.

Now what is happening that makes it possible to produce an electric current by a motion of a bar magnet with respect to a coil of wire? The essential phenomenon is the change of position of the magnet and the coil. As we have already discovered, lines of force surround the bar magnet. When the magnet is thrust into the coil, these lines of force move with respect to the individual turns of wire in the coil. The lines of force are said to "cut" the coil of wire; and it is this cutting of lines of force by a conductor that "induces" an electric current in the conductor.

Such was Faraday's discovery. He investigated the matter further, however, and made other discoveries of vast importance. For example, if we use two coils of wire instead of one coil and a bar magnet, and if the two coils are in close proximity but are not electrically connected (perhaps one coil inside the other), then if an electric current from a battery flows through one coil (the "primary"), a current will be produced in the second coil (the "secondary") when the relative position of the two coils is changed. The coil carrying the current acts exactly like the bar magnet, inducing a current in the second coil.

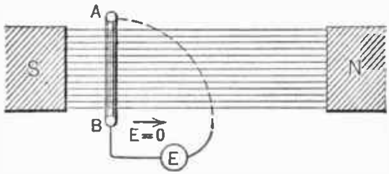


FIG. 12-5. As long as the conductor *AB* moves across the page, no voltage is generated. If it moves perpendicular to the paper, or up and down, a voltage is generated.

Now, instead of actually moving the position of one coil with respect to the other, let us maintain their relative positions but change the value of the current that flows through the primary. As long as this current changes, a current will appear in the secondary coil; but as long as the primary coil is constant in value, no current will be "induced" in the second coil.

The two coils present an exact analogy to the single coil and the bar magnet. The coil carrying the current acts like a magnet; lines of force emanate from it just as they do from the bar magnet. When these lines of force cut the turns of wire of the secondary coil, a current will be produced. If, however, the positions of the coils are not changed, and if an unvarying current flows through the primary, there will be no secondary current. A change in the lines of force with respect to the secondary coil is necessary. This change can be achieved either by moving one coil with respect to the other or by changing the amount of current flowing in the primary.

There is no discovery in all science more important than this discovery of Faraday's. Nearly every electrical application of modern times is dependent upon the phenomenon of **electromagnetic induction** outlined above.

The electric generator. We have described in the paragraphs above the essentials of a crude electric generator, a device for converting mechanical energy to electrical energy. A more workable arrangement is shown in Fig. 13–5. It consists of a turn of wire (a conductor) which is rotated mechanically between two magnets. In so doing this conductor cuts the lines of force between the magnets. Remembering that induced currents are produced only when the lines of force are cut by the conductor, it is easy to see that, as long as the conductor moves parallel with the lines, it will not cut any of the lines and no current will be produced. On the other hand, when the conductor moves at right angles to the lines of force, the greatest number of lines will be cut per unit of time and the greatest current will be produced under these conditions.

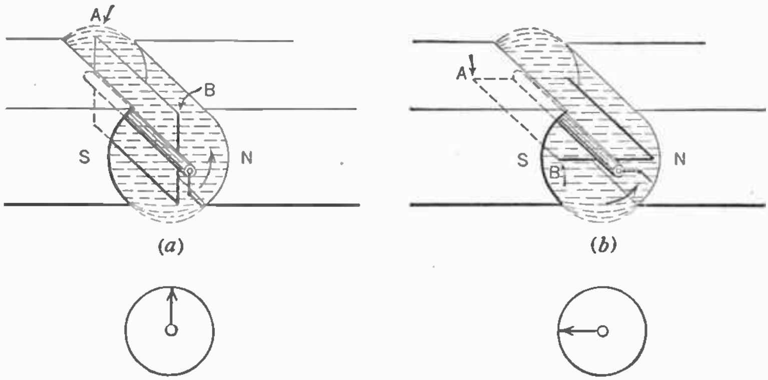


FIG. 13–5. In (*a*) the conductor is moving parallel or along the lines of force. In (*b*) the conductor is moving across or at right angles to the lines. At (*a*) the generated voltage is zero; at (*b*) it is maximum.

Now let us rotate the coil. At position (*a*), Fig. 13–5, the conductor moves parallel with the lines of force. No current will flow in the wire. As the wire rotates, it begins to move more and more at right angles to the field of force, and more and more current will be produced in the coil as at (*b*). Soon, however, the wire *AB* begins to approach the position (*a*) when it is again moving parallel with the lines of force, and the current begins to decrease in value.

At another portion of the complete rotation, the coil will be again cutting lines of force at right angles, but in a direction opposite to the direction in which it was cutting them at (*b*). Now the current in the coil will be found to flow in the direction opposite to its direction in (*b*).

Twice in the cycle of events (one revolution) there will be zero induced current; twice in the cycle there will be a maximum of current; these maxima will be in different directions; and between the zero-current positions, or between the maximum-current directions, the current will have values between zero and maximum. Thus the current rises and falls as the coil is rotated in the electric field.

The elements of an electric generator producing alternating current are described above. An actual generator is much more complex. The magnet is replaced by a heavy iron core covered with wire through which a direct current flows producing a strong magnetic field. The moving coil is wound on an iron form and is made up of many turns of wire wound in slots.

Alternating-current definitions. Since a circle, like a compass, can be divided into 360 degrees, we may describe the position of the rotating member of a generator in terms of the angle of rotation through which it has moved rather than by "position (*a*)" or "position (*b*)," or we may describe its position in terms of the time it has been moving since the start of its rotation. If it makes a complete rotation it has gone through 360 degrees; if it has turned half way through a complete rotation, it has turned through 180 degrees; and so on.

If the current at each position is plotted against the angle of rotation, expressed in degrees, a curve like that in Fig. 14–5 will result. Note that the current rises and falls about its zero-current value.

One complete revolution is called a **cycle.** The number of cycles per second is known as the **frequency** of the induced voltage or current. The time required for a complete revolution is called the **period.**

$$\text{Frequency} = \text{Cycles per second}$$

$$\text{Period} = \text{Time of one cycle} = \frac{1}{\text{Frequency}}$$

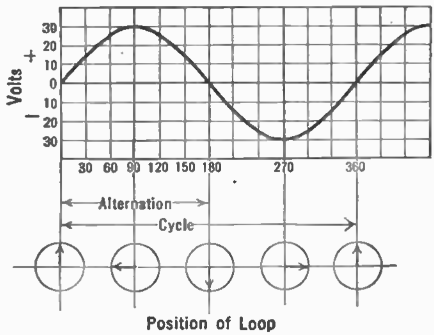


Fig. 14–5. Each position of the conductor as indicated by the arrows corresponds to some voltage as shown on the graph—called a sine wave of voltage.

Although frequency is correctly rated in "cycles per second," abbreviated as cps, engineers seldom use the "per second" part of the expression, the assumption being that it is the number of complete changes of current direction per second that is being considered and not the cycles per minute or other period of time. In this book the simple term cycles is used rather than cycles per second.

Work done by alternating current. Some may wonder whether or not a current that is continually reversing its direction—never getting anywhere, so to speak—is useful. It is.

Consider a paddle wheel in a stream of water. To the paddle wheel are attached millstones. Grain is to be ground between the stones. It matters little whether the water flows continuously turning the millstones in a certain direction, or whether the water flows first one way and then another. So long as the millstones turn against each other, grain will be ground, work will be done.

The alternating current which lights our homes is usually of 60 cycles. In some communities 25-cycle current is supplied. In radio installations for use on shipboard and on aircraft, generators produce 400- or 500-cycle currents. At the high-power radio stations of the Radio Corporation of America at Rocky Point, Long Island, are huge alternators which turn out radio frequencies of 20,000 cycles a second. Smaller generators which produce frequencies as high as 100,000 per second have been built but are not in common use. By means of vacuum tubes, alternating currents having frequencies of many millions of cycles are generated.

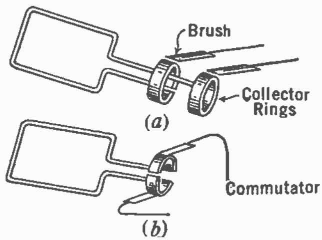


FIG. 15–5. Current is taken from an alternating-current generator by collector rings; a commutator serves the same purpose on a direct-current machine.

D-c generator. Current is taken out of a generator by **collector rings,** illustrated in Fig. 15–5. If current flowing continuously in the same direction is desired, a device called a **commutator** is used instead of the collector rings. A commutator is a switch, or valve, which keeps the output current flowing in the same direction by reversing at the proper time the position of the wire with respect to the rotating wire. In this manner the current flows through the external circuit in a single direction, although the current in the conductor which cuts the lines of force must reverse each time the conductor passes through 90 degrees and reverses its direction with respect to the field.

The commutator serves the same purpose as a valve in a pump which keeps water flowing upward whether the pump handle is worked down or up. A machine that sends out current that flows in a given direction is called a **direct-current** (d-c) generator, and, naturally, the current is known as direct current.

CHAPTER 6

INDUCTANCE

Coupled Circuits. Consider the two coils *P* and *S* in Fig. 1–6. They are said to be "coupled" when lines of force from one go through the other. When *P* is attached to a battery and the

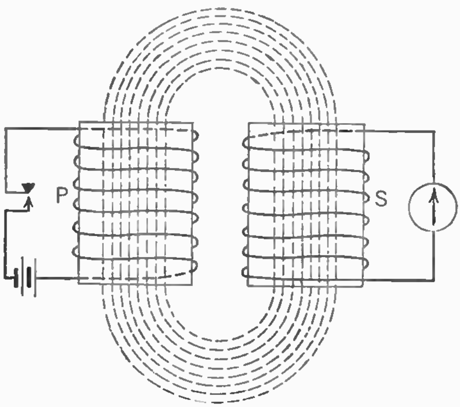


FIG. 1–6. If the coils *P* and *S* are coupled so that lines of force from *P* go through *S*, a current indicator will show a flow of current in *S* when the key in *P* is momentarily closed or open.

switch closed there is a momentary indication of the needle across *S*. When the switch is opened the needle moves again but in the opposite direction. As long as the current in the primary *P* is steady in value and direction, there is no movement of the needle across the secondary coil. A deflection of the needle indicates a momentary flow of current in the secondary coil; this current flows only when the primary current is changing, i.e., starting or stopping, not when the primary current is fixed in value or direction.

Lenz's law. Two fundamental facts about this phenomenon of coupled circuits should be noted. The first is that, when lines of force couple two coils together, and some change in these lines takes place, perhaps because of a change in the relative positions of the two circuits, a voltage is induced in the second circuit. The second fundamental fact is known as **Lenz's law:** the induced current is in such a direction that it opposes the change in position that produced it.

Thus, when the battery is attached to *P*, lines of force thread their way across the turns of wire in *S*. This movement of the lines of force through *S* induces a voltage across the coil, and a current flows in the coil and in the apparatus connected to it. The current in the second coil is in such a direction that its field, that is, its lines of force threading through the primary, induces a countervoltage in the primary opposite in direction to the battery voltage across it.

When the battery voltage is broken, the lines of force from the primary current collapse back on the coil and, in crossing the secondary turns in a direction opposite to that taken when the current in the primary is increasing, induce a voltage in the secondary in such a direction that it tends to keep the primary current flowing.

If it were not for the phenomenon expressed in Lenz's law, electrical equipment would burn up. If a current is induced in the secondary by a change in lines of force in the primary, a current will be induced in the primary by these changing secondary lines of force. This current would, in turn, induce more current in the secondary, which would produce more in the primary, and soon there would be excessive currents flowing and the apparatus would be destroyed.

If a large coil, a battery, and a quick-acting current meter are placed in series with a switch, the meter will be seen to reach its final value slowly and not at the instant the switch is closed. If the voltage drop across the coil is measured and compared to the battery voltage, it will be found that only when the final value of current is reached are the two voltages equal. At all other instants of time, the voltage across the coil will be less than the impressed voltage, but it builds up at the same rate that the current through the coil builds up. The actual difference between the two emf's is the voltage induced in the coil by the changing lines of force produced by the changing current. On the opening of the switch the current through the coil and the *IR* drop across it do not fall to zero instantly; they take as long to decrease to zero as they took to rise to their final "steady-state" values. The actual time of build-up may be quite short, a fraction of a second, but in large iron-core coils, say the field of a large generator, several seconds may be required for the final value of current to be attained.

The fact that current takes longer to reach its final value in a circuit in which there is a coil of wire indicates that something about the coil tends to prevent any change in the current. This phenomenon of induced voltages is of fundamental importance. It is the basis of all modern electrical machinery. Motors, dynamos, and radio signals are the result of our ability to produce changes in one circuit by doing something to another circuit even though the two circuits have no metallic connection whatever.

The property of an electrical circuit which tends to prevent any change in the current flowing through it is called **inductance.** It has a mechanical analogy in **inertia.** A flywheel requires considerable force to get it up to speed; and after it is started it will continue to run for some time after the driving force is removed. It does not stop suddenly. Considerable force is required to stop it, and the more suddenly one wants to stop it, the more force he must apply.

Inertia is evident in a mechanical system only when some change in motion is attempted. It is not the same as friction, which is always present. The inductance of an electrical system manifests itself when currents are changing. It is not to be confused with resistance, which is always present. Current flowing in a circuit containing resistance stops only when the driving force (voltage) is removed. If inductance is added to the circuit, the resistance remaining the same, a longer time will be required for the current to drop to zero when the voltage is removed.

Self-inductance. Inductance is added to a circuit by winding up a wire into a compact coil. For example, 1000 ft of No. 20 copper wire strung up on poles would have a resistance of about 10 ohms and the current into it would reach a final value very

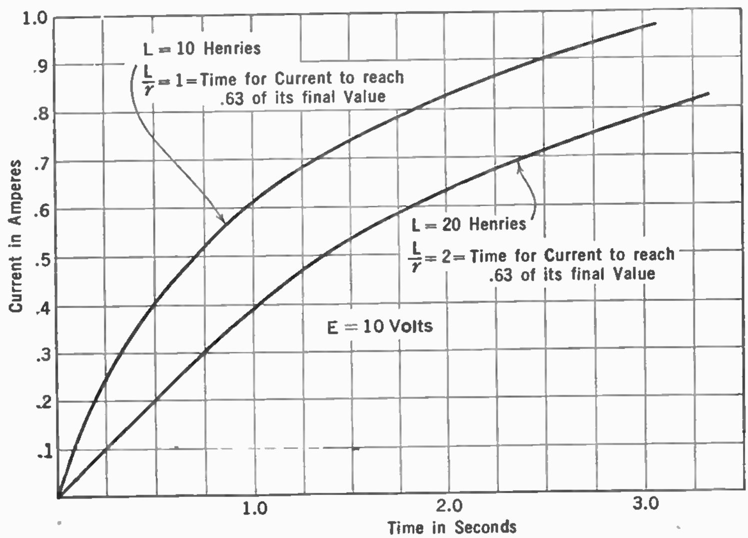


FIG. 2–6. Current in an inductive circuit does not rise instantaneously to its maximum value as these curves show.

soon after a battery was applied. If, however, the wire is wound up on a spool with an iron core, the time required for the current to reach its final value will look like the curve in Fig. 2–6. The resistance has not changed; we have merely added inductance.

When the switch is opened a fat spark occurs; this does not happen if the wire is strung up on poles. The current seems to try to bridge the gap—to keep on flowing. It does not "want" to stop. This is because of the inductance of the coil.

A single coil can have inductance and can have a voltage induced across it just as though it were the secondary coil shown at S in Fig. 1–6.

When the current starts to flow through the coil, lines of force begin to thread their way out through the coil, thereby cutting adjacent turns of wire and, according to Lenz's law, inducing in each turn a voltage in such a direction that it tends to oppose the building up of the current from the battery. When the battery connection is broken, these lines of force collapse, cutting the turns of wire and thereby inducing voltages which tend to keep the battery current flowing.

Magnitude of inductance and induced voltage. The greater the number of turns of wire in a small space, or the higher the permeability of the core on which the wire is wound, the greater will be the inductance of the coil and the longer the time required for the current to reach its final value. The inductance of a coil is proportional to the square of the number of turns on the coil; the induced voltage is proportional to the inductance and the rate of change of the current through the coil. The time required for the current through a coil to reach its final value is directly proportional to the inductance and inversely proportional to the resistance of the coil. Since the current tends to keep on flowing when an inductive circuit is broken the induced voltage across the coil must be in the same direction as the battery voltage. Thus, if a coil has 100 volts from a battery across it, and the current is suddenly broken, the voltage at the instant of break across the coil will be 100 plus the additional induced voltage. Breaking the current in a highly inductive circuit may set up a tremendous voltage across the coil and a severe shock can be felt by holding the ends of the wire at the moment of break. This is a practical demonstration of Lenz's law.

Example 1–6. In Fig. 3–6 is an inductance made up of the field winding of a generator. Across the inductance are placed a flash lamp and a battery. The current is regulated so that it is insufficient to light the lamp. Now when the switch is opened the lamp will suddenly light (and may burn out) because of the tendency of the collapsing field of the coil to maintain the current flowing in the same direction in which it has been flowing.

In radio circuits the coils are usually wound on non-magnetic cores. In audio * and power circuits iron is utilized to build up large inductances in small spaces and with a minimum of copper wire.

If an a-c voltage is placed across a coil, the current through the coil is much less than if a d-c voltage is placed across it. This is because of the countervoltage or back-voltage induced across the coil by the effects just described.

The unit of inductance. When a current change of 1 amp per sec causes an induced voltage of 1 volt, the inductance is said to be 1 **henry.** This unit is named in honor of Joseph Henry, an American experimenter who discovered the phenomenon of electromagnetism at the same time as Michael Faraday.

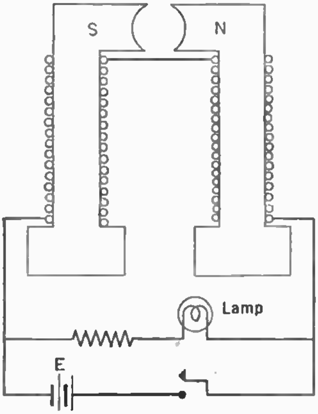


FIG. 3–6. An example of Lenz's law. The lamp will light when the key is opened even though insufficient current flows through it when the key is closed.

The quantitative facts of electromagnetism may be expressed as

$$E_{av} = L\frac{\Delta I}{\Delta t}$$

where E_{av} is the average induced voltage.

Δ is the Greek letter delta indicating "a change in."

I = current in amperes.

t = time in seconds.

The portion of the expression $\Delta I/\Delta t$ may be translated as "the rate of change of current with respect to time." The induced

* **"Audio" circuits** involve alternating currents which can be heard by the human ear, i.e., from 50 to 15,000 cycles. **"Radio" frequencies** are much higher and are inaudible.

voltage, therefore, is proportional to the rate at which the current changes and is equal to amperes change per second multiplied by a factor L known as the inductance.

Coils added to circuits for the purpose of increasing the inductance are properly called **inductors,** although many engineers use the terms inductance (a property of a circuit) and inductor (a piece of apparatus) interchangeably.

The inductance is really a measure of the flux lines linking with the turns of a coil. The unit of inductance, the henry, is that inductance which causes 10^8 interlinkages per ampere change in current. Thus

$$L = \frac{N\Delta\Phi \times 10^{-8}}{\Delta I} \text{ henries} \qquad [1]$$

where N = the number of turns on the coil.
Φ = magnetic flux linking these turns due to I.
I = amperes of current through the coil.
10^{-8} = a constant term to bring L into practical units.

Now we have already learned that the flux in a long coil, or solenoid, is

$$\Phi = \frac{0.4\pi NI}{\mathcal{R}}$$

where $\mathcal{R}$ is the reluctance.

$$\mathcal{R} = l \div \mu A$$

where l = length of the winding in centimeters.
A = area of cross section in square centimeters.
μ = permeability of core on which coil is wound.

Therefore,

$$\Phi = \frac{0.4\pi NIA\mu}{l} = \frac{1.26NIA\mu}{l} \qquad [2]$$

If this value of Φ is substituted in formula 1 an expression for the inductance in henries of a coil whose length is large compared to its diameter will be found. Thus

$$L = \frac{1.26N^2A\mu}{10^8 l} \qquad [3]$$

where μ = permeability = 1 for air.
A = cross section of coil in square centimeters.
l = length of winding in centimeters.

Typical inductances. The coils used in radio apparatus vary from small coils having inductances of the order of microhenries (μh) to very large coils having 100 or more henries inductance. Broadcast-frequency tuning coils are of the order of 300 μh and may be from ½ to 3 in. in diameter wound with No. 30 to No. 20 wire of 50 to 100 turns or so. There are a number of complicated formulas by which one can calculate the inductance of coils of various forms and sizes. The ones in Fig. 4–6 are accurate enough for practical purposes. Coils of the type (*b*) have an inductance of

$$L = \frac{a^2N^2}{9a + 10b} \quad \mu\text{h}$$

when the dimensions are in inches and when b equals or is greater than a.

These formulas show that the inductance increases as the square of the number of turns. Thus, if a coil of 3 units inductance has its number of turns doubled, the inductance will have increased four times or to 12 units. This is true provided that there is good "coupling" between turns; that is, if the coil is on an iron core this rule is strictly true, but if the coil is wound on a core of air the rule is only approximately true. It becomes more nearly a fact the closer together the turns of wire.

Problem 1–6. A coil like that in Fig. 4–6*a* is called a **multilayer** coil. Such a coil is wound to have 1000 turns, in a slot 1 in. square. The distance from the center of the coil to the center of the winding (a in Fig. 4–6) is 2 in. What is the inductance of the coil in microhenries? Remember that the dimensions given in Fig. 4–6 are in centimeters and that 1 in. equals 2.54 cm.

Problem 2–6. A coil like that in Fig. 4–6*b* is called a **solenoid.** Calculate the inductance of such a coil composed of 60 turns of wire in a space of 3 in., the diameter of the coil form being 3 in. Coils used in modern broadcast receivers are wound on cores about 1 in. or less in diameter.

Problem 3–6. A transformer has a core made of silicon steel having a permeability of 4500. The core has a cross section 15 cm square and is 100 cm long. The primary of this transformer has 500 turns. What is the inductance?

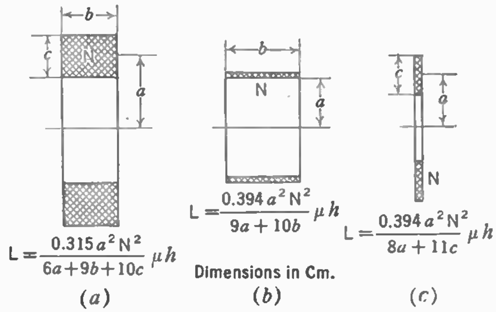


FIG. 4–6. Some typical coil forms and formulas on which the inductance may be calculated.

Problem 4–6. If the current in this primary changes from 20 to 1.5 amp in ¼ sec, what voltage is induced across the primary?

NOTE. $E = L\left(\frac{I_2 - I_1}{t_2 - t_1}\right)$.

Coupling. The closer together the two coils, P and S, Fig. 1–6, the greater the number of the lines of force produced by the primary current that link with the turns of the secondary, and the better the "coupling" is said to be. The better the permeability of the medium in which the lines go, the better the coupling.

The voltage across the secondary of such a two-coil circuit as that shown in Fig. 1–6 depends upon the sizes of both coils, their proximity, the permeability of the medium, and the rate at which the primary current changes. All the factors except the rate of change of the primary current are grouped together and are called the **mutual inductance** of the circuit.

The secondary voltage, then, is equal to

$$M \times \text{Rate of change of primary current}$$

where M is the mutual inductance and is rated in **henries.**

Magnitude of mutual inductance. Formulas in Fig 4–6 show that the inductance of a coil depends upon the square of the number of turns. Doubling the turns increases the inductance four times. Consider two coils built alike and having the same inductance. If they are connected together in series and if the coils are perfectly coupled as in Fig. 5–6*a*, the total inductance will be equal to that of a single coil of double the turns. In other words, the total inductance of two coils connected "series

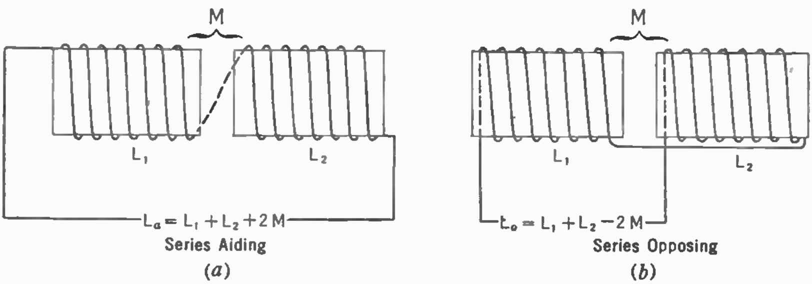


FIG. 5–6. Coils may be connected so that their fields aid or "buck." In one the total inductance is increased; in the other the total inductance is less than either of the component inductances.

aiding" will be four times the inductance of a single coil. If the connections to one coil are reversed, the total inductance will be zero because the lines of force from one coil will encounter the lines of force from the other coil which are in the opposite direction. The coils are now connected "series opposing."

Consider the series-aiding case, Fig. 5–6*a*. The total inductance is made up of the inductance of coil 1, that of coil 2, the mutual inductance due to the lines of force from coil 1 which go through coil 2, and the mutual inductance associated with the lines from coil 2 which go through coil 1; these two inductances are equal because the coils are identical. Thus,

$$L_a = L_1 + L_2 + 2M$$

$$= 2L_1 + 2M \text{ (because } L_1 = L_2\text{)}$$

$$= 4L_1 \text{ (by experiment or measurement)}$$

whence

$$M = L_1$$

Now if the coupling between coils is less than perfect, that is, if, as always happens, some of the lines from one coil do not link the other, the total inductance, L_a, in the series-aiding case is less than four times the inductance of one coil and in the series-opposing case is greater than zero. But in any event the total inductance of two coils of any inductance connected in series aiding will be given by $L_1 + L_2 + 2M = L_a$, and if they are connected series opposing the resultant inductance will be $L_1 + L_2 - 2M = L_o$.

Whatever the values of inductance possessed by the individual coils, L_1 and L_2, if they are connected series aiding so that complete flux linkage results and if the total measured inductance has subtracted from it the individual inductances of the two coils, a final value will be found which is numerically equal to twice the square root of the product of the individual inductances. Thus

$$L_a = L_1 + L_2 + 2\sqrt{L_1 L_2}$$

Therefore

$$M = \sqrt{L_1 L_2} \quad \text{(for complete coupling)}$$

The expression $M/\sqrt{L_1 L_2}$ is called the "coefficient of coupling" (τ) since it is a measure of the completeness of the coupling between the individual coils. Values of this factor may be as low as 5 per cent (0.05) in air-core coils used in radio-frequency circuits or as high as 98 per cent (0.98) in power and audio transformers using iron cores.

The coefficient of coupling depends upon the total inductance in the primary and secondary circuits as well as upon the mutual inductance between inductances L_1 and L_2, Fig. 6–6.

The mutual inductance depends upon the two coils L_1 and L_2 and the coupling between them, or $M \doteq \tau\sqrt{L_1 L_2}$; the coefficient of coupling between two circuits depends upon the total inductance in the circuits. The maximum possible value of τ is 1.0. This value, called unity coupling, is approached in iron-core transformers. Of course, if the two coils are situated so that there is no coupling, either because they are too far apart or are placed in iron boxes, the total inductance in the circuit will merely be the sum of the individual inductances.

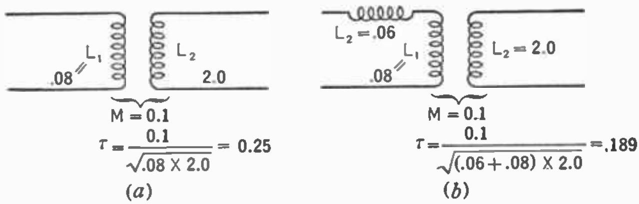


FIG. 6-6. Examples showing dependence of coefficient of coupling on series inductance.

Measurement of inductance. Inductance is usually measured, like resistance, by means of a Wheatstone bridge. The Wheatstone bridge method consists essentially of comparing the unknown inductance to a known inductance. Resistances are used as the ratio arms, as in Fig. 7-6. When there is no sound in the telephones the inductances are equal if the ratio arms are equal, or if the ratio arms are not equal the unknown inductance is given by the equation

$$L_x = L_s \frac{A}{B}$$

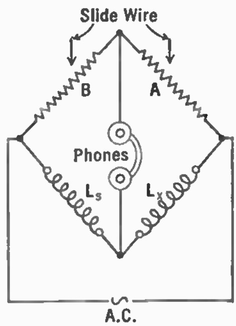


FIG. 7-6. If A and B are so adjusted that there is no sound in the phones, $L_x = L_s \frac{A}{B}$.

Mutual inductance is measured on a bridge, or by the following method: The inductance of the individual coils may be measured first. Then they are connected series aiding and the total inductance is measured. This gives the formula $L_1 + L_2 + 2M$, from which M can be calculated at once. Of course, the same result will be obtained by connecting the coils series opposing. It is not even necessary to measure the individual inductances first, provided that we can measure the inductance both series aiding (L_a) and then series opposing (L_o). Then

$$4M = L_a - L_o$$

$$M = \frac{L_a - L_o}{4}$$

Problem 5–6. In Fig. 8–6, when the coils are connected series aiding, the inductance is measured as 400 μh. What is the mutual inductance?

Problem 6–6. In a screen-grid-tube radio-frequency amplifier transformer the primary inductance is 350 μh, the secondary is 230 μh, and the mutual inductance is 160 μh. Calculate the coefficient of coupling.

Problem 7–6. In Fig. 8–6, L_1 = 100 μh, L_2 = 200 μh, and τ = 0.6. What is the mutual inductance? What is the total inductance ($L_a = L_1 + L_2 + 2M$)? What is it if L_2 is reversed (M is negative)?

The transformer. A transformer is a device for raising or lowering the voltage of an a-c circuit. It transforms one voltage

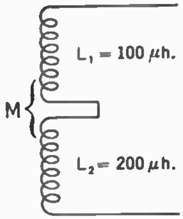


FIG. 8–6. A problem in coupled circuits.

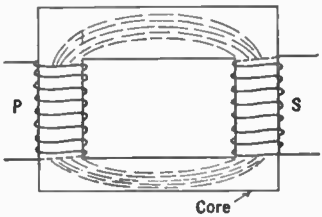


FIG. 9–6. A simple transformer.

into another. It consists of two windings on an iron core, as in Fig. 9–6. The purpose of the iron core is to insure that the magnetic field set up about the primary will flow through the secondary coil without loss. The lines of force that do not link primary and secondary are called **leakage** lines, and the inductance associated with them is called **leakage inductance.**

The primary is attached to an a-c line, the secondary to the load, whether this is a house lighting circuit, a motor, or any other device that requires electricity.

The continually changing lines of force from the primary flow through the secondary and induce voltages in it. The relation between the primary and secondary voltages is simple and definite—it depends upon the relative number of turns. If there are twice as many secondary as primary turns, the voltage developed across the secondary terminals will be double that across the primary. The following formula gives the relation between primary and secondary turns and the respective voltages:

$$\frac{e_p}{e_s} = \frac{n_p}{n_s} = N \quad \text{(turns ratio)}$$

By using the proper ratio of turns, voltages either greater or less than the primary voltages may be secured at the secondary terminals.

Example 2–6. A transformer is to connect a 110-volt motor to a 22,000-volt transmission line. What is the ratio of turns between secondary and primary?

$$\frac{e_p}{e_s} = \frac{n_p}{n_s}$$

$$\frac{22{,}000}{110} = \frac{n_p}{n_s} = 200$$

Note. This does not give the number of turns in either primary or secondary windings. The absolute number of turns depends on several factors; the ratio of turns depends on the voltages to be encountered.

Problem 8–6. In electric welding a very low voltage is used. What would be the turns ratio of a transformer to supply a welding plant with 5 volts if it takes power from the standard 110-volt circuit?

Problem 9–6. In a radio receiver, 5 tubes are operated in parallel from a 2.5-volt secondary winding on a transformer. Each tube requires 1.75 amp. The primary voltage is 110. What are the turns ratio and the minimum primary current? If the transformer is 90 per cent efficient, what power does it take from the 110-volt line?

Problem 10–6. A primary consists of 200 turns of wire and is connected to a 110-volt circuit. The secondary feeds a rectifier circuit requiring 550 volts. How many turns will be on the secondary winding?

Problem 11–6. A certain transformer used in a radio-frequency amplifier has a secondary inductance of 240 μh. When the two coils of the transformer are connected series aiding the total inductance has increased to 350 μh, and when connected series opposing the inductance drops to 230 μh. What is the inductance of the primary? What is the mutual inductance between them in the two cases, and what is the coefficient of coupling?

Power in transformer circuits. Since the transformer does not add any electricity to the circuit but merely changes or transforms from one voltage to another the electricity that already exists, the total amount of energy in the circuit must remain the same. If it were possible to construct a perfect transformer there would be no loss in power when it is transformed from one voltage to another. Since power is the product of volts and amperes, an increase in voltage by means of a transformer must result in a decrease in current, and vice versa. On the secondary side of a transformer there cannot be more power than in the primary, and, if the transformer has high efficiency, the power will be only slightly less than on the primary side. The product of amperes and volts remains the same. Thus the primary power is

$$E_pI_p \qquad [1]$$

and the secondary power is

$$E_sI_s \qquad [2]$$

and since there is no loss or gain in power

$$E_pI_p = E_sI_s \qquad [3]$$

whence

$$\frac{I_p}{I_s} = \frac{E_s}{E_p} = N \quad \text{(turns ratio)}$$

which shows that the secondary voltage increases as N increases, the secondary current decreases when N increases, assuming constant primary voltage and current.

The transformer acts like an electrical lever, enabling us to get high voltage at low current from a low-voltage, high-current source. Since heat losses in a transmission line are proportional to the current squared (I^2R), it is common practice to transmit power at high voltages and low currents. There will be 100 times the power loss in a 110-volt line that there will be in an 1100-volt line if both lines transmit the same power. Transformers make it possible to generate a-c power at comparatively low voltage, to transmit it at high voltage, and to utilize it at low voltage.

The auto-transformer. It is not necessary for proper transformation of voltage that the primary and secondary windings be distinct. Figure 10–6 represents what is known as an auto-transformer, in which the secondary is part of the primary. The voltage across the secondary turns, however, bears the same relation to that across the primary part as though there were two separate windings. The ratio of voltages is the ratio of the respective number of turns possessed by the secondary and primary. Since one side of the a-c line is usually grounded, one primary terminal of the auto-transformer will be grounded and therefore these transformers cannot be used in circuits where a ground is apt to be on the secondary side.

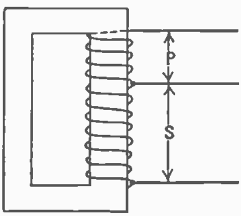


FIG. 10–6. An auto-transformer. The secondary can be the entire winding or part as shown.

Transformer applications. A transformer is often used when both alternating and direct currents flow through a circuit and it is desired to keep the direct current out of the circuit to which the secondary is attached. No direct current can go across the transformer when the two windings are distinct, although the a-c voltage variations occurring in the primary are transferred to the secondary by the effects already described. If no increase or decrease in voltage is desired the turns ratio is made unity; that is, the same number of turns will be on the secondary as on the primary.

An output transformer which couples a loud speaker to a power tube in a power amplifier serves this isolating purpose. The power tube has considerable direct current flowing in its plate circuit where the useful alternating currents also flow. Since the direct current is undesirable in the loud-speaker windings, a transformer is used to isolate the speaker from the direct current of the tube. A good transformer of this type will transmit all frequencies in the audible range with an efficiency of about 80 per cent.

Problem 12–6. The line voltage in a certain locality is only 95 volts, but a radio set is designed to operate from a 115-volt circuit. A transformer is to be used like that in Fig. 10–6. What will be the ratio of turns?

When no current is taken from the secondary of a transformer the primary acts merely as a large inductance across the line. The current will be rather small. The energy associated with this current is used in two ways, one of which is heating the transformer and its core. The other part maintains the magnetic field of the primary. This consumes no energy from the line because at each reversal of the current the energy of this field is given back to the circuit.

When the secondary load is put on, however, it begins to draw current from the secondary and more power is taken from the line leading to the generator. This additional power is that required by the load together with the loss in primary and secondary resistance.

Transformers are used in practically all radio, television, and radar equipment.

Transformer losses. Transformers are not perfect. Not all the power taken from the line by the primary results in power in the secondary load circuit. There are two sources of power loss: (1) losses in the windings (which have resistance and therefore heat up when current passes through them), and (2) core losses. The winding losses are called the copper losses. They are equal to the I^2R loss in the primary plus the similar loss in the secondary. These losses in watts may be found from the resistance of the windings, which is determined by measuring the current flowing through the individual windings when a known voltage (direct current) is across the winding. This calculation will give a value for R from Ohm's law. Then the current taken by the primary and secondary when the load is applied to the secondary may be measured with an a-c ammeter. From these values the I^2R losses may be found. At frequencies above the audio range, the "skin effect," described later, causes another loss.

Example 3–6. A small power transformer is tested with a 1.5-volt battery. The primary passes 3.75 amp, and the secondary, 0.75 amp. The d-c resistance of primary and secondary windings, respectively, is, therefore, $E \div I = 0.4$ and 2.0 ohms.

When the transformer is connected to an a-c line and the secondary feeds power to its load, the primary and secondary currents are 10 and 2 amp, respectively. The total heat loss will be equal to

$$P = I_p^2R_p + I_s^2R_s = 10^2 \times 0.4 + 2^2 \times 2 = 40 + 8 = 48 \text{ watts}$$

The second source of power loss is in the core. The magnetic flux expanding and contracting about the core induces a voltage in the secondary. But the same lines of force which cause voltages to be induced in the secondary also produce voltages in the iron core upon which the windings are placed. These voltages produce currents in the core, and the currents produce heat (I^2R), and a heat loss must be added to the copper losses. In addition to these loss currents, known as **eddy** currents, there is another power loss in the core which is due to the basic principles of magnetism. It is supposed that, when iron is magnetized, the position of the molecules is changed so that all the north poles point in one direction and all the south poles in the opposite direction. When the direction of flux changes every half cycle, the orientation of the molecules must change so that the north poles now point in the opposite direction. The continuous shifting of the positions of the molecules in the core causes the core to heat up, and the heat represents a definite loss of power. This wastage of power is known as **hysteresis** loss.

Since all the losses must be supplied by the power taken from the line, they represent power which is not supplied to the secondary load. Special pains are taken to keep these unwanted losses as small as possible, for example, by using copper wire as large as practicable to keep the resistance low and by making the core of laminations (thin insulated sheets of iron) rather than of solid iron. The currents induced in the core flow across the breadth and depth of the laminations and not very much in the lengthwise direction. Currents through the depth are kept small by insulating the individual laminations; currents across the breadth may be kept low by making the core out of fine insulated wires. Now the currents are restricted to the diameter of the iron wire or the thickness of laminations, both of which are quite small.

Eddy-current and hysteresis losses may be determined by measuring the power input and power output of the transformer. The difference is the total loss. Then the copper losses may be calculated as indicated above. The copper losses subtracted from the total loss give the hysteresis and eddy-current or core losses.

Iron cores of one type or another are universally used in power transformers. In transformers operating at higher than commercial power frequencies (60 cycles) the use of iron becomes much more difficult and the difficulty increases as the frequency increases. Hysteresis effects become very great because of the rapid reversal of the magnetic flux direction. Radio-frequency coils contain either non-magnetic core material or core material of very finely powdered iron which is mixed with a binder and then pressed into the desired form. These iron dust cores have fairly low losses and may be used at broadcast frequencies or somewhat higher. They increase the inductance of a given coil appreciably and enable engineers to get high inductances in small space.

Certain high-grade audio transformers utilize high-permeability iron that saturates with only moderate direct current through the primary. For this reason it is good practice to keep the direct current out of the windings entirely. This is accomplished by feeding power to the circuits allied with such a transformer through a high-inductance single-winding coil (choke coil) and allowing the alternating currents to go through the transformer by means of a large condenser.

Transformers with direct current. It is not necessary that the primary be connected to a source of alternating current. All that is necessary is for the primary current to be interrupted or reversed in direction to realize an interrupted current in the secondary circuit. A battery, for example, furnishing direct current may be placed in series with the primary of the transformer and a vibrator or other current-interrupting device. At each break of the primary battery current, a voltage will be induced in the secondary. The same step-up or step-down in voltage is secured as where alternating current is utilized. This principle is employed in supplying power from vibrator interrupters to mobile radio apparatus.

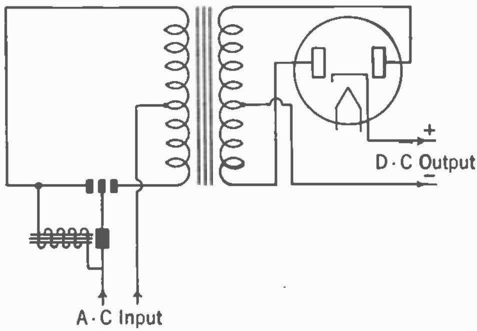


FIG. 11–6. The use of a vibrator, a transformer, and a rectifier to secure high-voltage direct current from a low-voltage source.

Variable inductors. Several means are available for varying the inductance of a coil. One is the simple expedient of having connections brought out to several points in the winding. Transformers for use at several commercial voltages, say at 105, 110, 115, etc., are made in this way. Another method is to scrape the insulation off the wire along the length of the winding and make contact with any point desired by means of a sliding contact. Also, merely sliding an iron core into the coil will vary the inductance. Still another method consists of orienting one coil with respect to another; thus the coupling between the windings may be changed so that the total inductance made up of the primary, secondary, and mutual inductance can be varied from nearly zero to a point appreciably greater than the sum of the two individual winding inductances.

Air gap in iron-core inductors. If the magnetizing force applied to an iron-core coil is increased indefinitely, the magnetic flux will not increase indefinitely but will reach a "saturation"

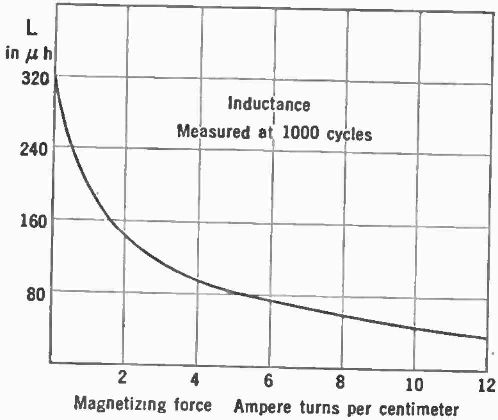


FIG. 12-6. Inductance of iron-core coils varies with direct current because of variation in permeability.

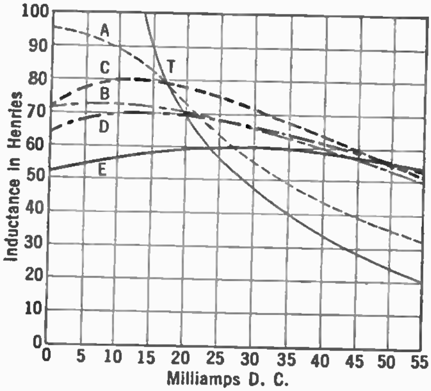


FIG. 13-6. Variation of inductance with air gap.

T no air gap.
A average air gap.
B air gap at one end, 0.01 in.
C air gap at both ends, 0.005 in. each.
D air gap at both ends, 0.0075 in. each.
E air gap at both ends, 0.01 in. each.

value. Further increase in magnetizing force (as by increasing the current through the coil) will not produce any more flux and therefore will not produce any further "back emf," and the coil will overheat. The difficulty may be overcome by inserting a small air gap in the iron core. Even a gap of only a few thousandths of an inch has a very useful effect for the simple reason that air cannot be saturated by flux. Increasing the current in the coil increases the flux in the air gap, even though the flux in the iron does not increase.

The air gap is useful where direct as well as alternating current flows through the coil. The direct current may produce sufficient flux to almost saturate the core, for example. Then, at the peaks of the impressed alternating currents, the flux does not increase proportionately to the increase in current, and the core saturates. The air gap overcomes the trouble. Coils in which both direct and alternating current flow are used in rectifier power supply circuits to filter out hum and noise.

Problem 13–6. A transformer has an input power of 5000 watts. From the secondary 4000 watts is drawn. The total copper loss is 700 watts. What power is being consumed in the core losses?

Problem 14–6. A small power transformer feeds power to a radio set. Normally it delivers 2 amp to the receiver. When it is disconnected from the receiver and from the a-c line and connected in series with a 2-volt battery, the secondary passes 1 amp. What is the heat loss in the secondary?

Problem 15–6. A transformer is used to change voltages in an a-c system. Suppose that we have a water wheel connected to a generator 50 miles from a town which wishes to utilize the output of the generator. The town requires 1000 amp of current. If the same line were to be used in the following two cases, calculate the power lost, the line losses, and the saving accomplished by stepping up the voltage at the generator to 11,000 volts compared to transmitting it at 110 volts.

CHAPTER 7

CAPACITANCE

The two essential electrical quantities in every radio circuit are **inductance** and **capacitance;** they are represented in the circuit by coils and condensers. Upon the relative size of these two essential components depends the frequency or wavelength to which the receiver or transmitter is tuned or is responsive. It is a fact that resistance is always present too, but the effort of radio engineers is to reduce the resistance (where not wanted) and to overcome the loss in power due to its presence, just as mechanical engineers devise ways to reduce friction and the loss in power due to it.

Capacitance. Inductance has been likened to inertia. In an a-c circuit, it acts to prevent changes in the current flowing. If current increases, owing to an increase in voltage, inductance tends to prevent or delay this increase; if the current decreases, owing to a decrease in voltage, inductance tends to prevent or delay this decrease. Inductance is a property of the circuit as is also capacitance.*

When a current flows through a coil of wire, a magnetic field is produced. Energy resides in this field. Similarly, energy can be stored in a condenser. A capacitor tends to prevent any change in the voltage of a circuit in which it is placed. If the voltage tends to decrease or increase for any reason, the capacitance of the circuit tends to prevent this decrease or increase. The proper combination of inductance and capacitance will force the current and voltage of a circuit to remain constant in spite

* Capacitance is the preferred word to describe this property of a circuit; a device having capacitance is a **capacitor.** Engineers nearly always use the terms "condenser" for a capacitor and "capacity" for capacitance.

of varying conditions which, in a d-c circuit, would change the voltage and current.

The charge in a condenser. A condenser is made up of two conductors separated by an insulator; for example, two long sheets of metal foil may be wound up with a sheet of waxed paper between.

The questions naturally asked about such a device are these: Can a condenser pass an electric current? What is the charge in a condenser? How much capacitance exists in a condenser? What is the spark that passes when the condenser is discharged?

When a condenser is connected to the terminals of a battery, there is a momentary rush of electrons to one of the conducting plates, and a corresponding momentary rush of electrons away from the other plate. There is no passage of electrons through the separating non-conductor. Since a movement of electrons constitutes an electric current, a proper kind of voltmeter placed across the condenser plates would be seen to rise from zero to the voltage of the battery during the period in which the condenser is being "charged." If the battery is now removed it will be found that the condenser remains charged; that is, a proper kind of voltmeter will indicate a difference of potential across its two terminals.

This simply means that one plate has too many electrons, and the other has too few. The condenser is in a state of unequilibrium. If a wire is connected from one terminal to the other, a spark will pass, indicating that the maldistribution of electrons on the two plates is being corrected. Now the condenser is discharged.

If the condenser is placed in an a-c circuit, there is a flow of electrons to the plate that is positive and away from the plate that is negative; then, when the polarity of the voltage changes, the electrons again move and charge the condenser in the opposite direction. This to-and-fro motion of the electrons is just as effective in doing work as if the electrons moved through the insulator separating the plates.

So long as the electrons are unevenly distributed on the two plates the condenser is charged. In this condition it possesses energy, which is like the energy possessed by a ball on top of a flag pole. The kind of energy possessed by the ball is potential energy; it is due to the position of the ball. The energy possessed by the condenser when charged is also potential energy, due to the strain existing in the non-conductor. Nothing happens until the condenser discharges; then it may set fire to a piece of paper, may puncture a hole in a sheet of glass, or may give some person a severe shock. Thus the condenser, just like the ball on top of the pole, has the ability to do work—which is our definition of energy. The energy resides in the distorted placement of the charges or electrons, a condition that will equalize itself at the first opportunity.

Quantity of electricity in a condenser. The electricity that rushes into a condenser when it is connected to a battery is a perfectly definite quantity depending upon only two factors, the size of the condenser and the voltage of the battery. Naturally, the larger the condenser the greater the quantity of electricity that can be stored in it; and the greater the electrical pressure tending to produce a state of unequilibrium, the greater the charge that can be forced onto the condenser.

The quantity of electricity in the condenser is a definite number of electrons or charges and is rated in coulombs.

The capacitance of a condenser depends only upon the physical make-up of the unit: (*a*) the size of the conducting plates, (*b*) the nature of the non-conductor between the plates, and (*c*) the distance between the plates.

$$Q \text{ (coulombs)} = C \text{ (capacitance)} \times E \text{ (voltage)} \qquad [1]$$

The unit of capacitance is the farad (f), named after Michael Faraday; 1 farad is the capacitance of a condenser whose voltage is raised 1 volt when 1 coulomb of electricity is added to it, or it is the capacitance of a condenser to which 1 coulomb of electricity can be added by the application of an external emf of 1 volt. This is a very large unit, and in radio circuits the millionth part of a farad (microfarad, μf) or even the micromicrofarad ($\mu\mu$f) is the customary unit.

We may write equation 1 as

$$C \text{ (farads)} = \frac{Q \text{ (coulombs)}}{E \text{ (volts)}} \qquad [2]$$

The capacitance may be considered as the ratio between the quantity of charge in the condenser and the voltage across it. As a third way of stating this relation, we may write

$$E \text{ (volts)} = \frac{Q \text{ (coulombs)}}{C \text{ (farads)}} \qquad [3]$$

A discharged condenser has no energy in it; there are just as many electrons on one set of plates as there are on the other set of plates; there is no strain across the insulating material between the plates; there is no voltage across it.

Time required to charge a condenser. How long does it take to charge a condenser? Let us connect a condenser across a battery. Electrons will leave the plate connected to the positive terminal of the battery, since negative electrons are attracted toward any positive body; and, at the same time, electrons on the negative terminal of the battery will leave this terminal and move to the condenser plate connected to this battery terminal. If the condenser and the connecting wires have little resistance, this movement of electrons will take place practically instantaneously. The voltage across the condenser plates becomes equal to the voltage across the battery terminals very quickly.

Until the voltage across the condenser becomes equal to the battery voltage, however, there is a motion of electrons—that is, there is a flow of current. This current is proportional to the rate at which the voltage across the condenser changes. No current flows when there is no change in voltage. When the battery is first connected to the condenser, the rate of change of voltage is a maximum since the voltage is changing from zero to some finite value. The current that flows, therefore, is a maximum.

As electrons move away from one condenser plate and toward the other plate, the voltage across the condenser rises, and, since the voltage must oppose the battery voltage to satisfy Lenz's law, the rate of change of voltage decreases. The current flowing decreases accordingly, and it reaches zero when the condenser and battery voltages are equal.

If the condenser is perfect, that is if it has no resistance and if there is no resistance in the leads connecting it to the battery, the instantaneous current flowing into the condenser will be infinite and zero time will be required to charge the condenser. All practical condensers, however, and all connecting wires have some resistance. The current, therefore, is not infinite, and finite time is required for the battery emf to move the appropriate number of electrons to and from the condenser plates.

The instantaneous currents and voltages occurring in inductive and capacitive circuits are of great importance in modern electronic applications such as radar, welding timers, and television. The conditions during these short intervals of time, known as "transients," will be discussed in more detail later in this book.

Time constant. The greater the resistance and the greater the condenser capacitance, the longer it takes to charge a condenser. The time required for the voltage of a condenser to reach 63 per cent of its final value, either on charge or discharge, is equal to the product of the capacitance and the resistance, in farads and ohms, respectively. This RC product is often called the time constant of the circuit.

Example 1–7. A condenser of 15 μf is attached to a 220-volt circuit. What quantity of electricity flows into it? If it is charged through a resistance of 10,000 ohms, how long will it take to put 63 per cent of its total charge into it?

$$\begin{aligned} Q &= C \times E \\ &= 15 \times 10^{-6} \times 220 \\ &= 3300 \times 10^{-6} \\ &= 0.0033 \text{ coulomb} \\ t &= RC \\ &= 10{,}000 \times 15 \times 10^{-6} \\ &= 15 \times 10^{-2} \\ &= \tfrac{15}{100} \text{ second} \end{aligned}$$

Problem 1–7. What is the capacitance of a condenser that holds 0.0022 coulomb when attached to 220 volts?

Problem 2–7. In a radio circuit is a 0.0005-μf (500-$\mu\mu$f) condenser across a 500-volt source. What quantity of electricity will flow into it?

Problem 3–7. What voltage will be necessary to put 0.012 coulomb into a condenser of 1 μf?

Problem 4–7. Suppose that the average voltage across a condenser when it is being discharged is one-half the voltage when fully charged. Connect a 10-ohm resistance across a 1-μf condenser charged to 110 volts. What average current flows? What is the average power used to heat the wire?

Energy in a condenser. The amount of energy that can be stored in a condenser in the form of static electricity can be computed from the formula

$$\text{Energy} = \tfrac{1}{2}CE^2$$

where energy is in joules, C in farads, and E in volts. This equation represents the work done in charging the condenser and, naturally, the energy released if the condenser is discharged.

Similarly, the energy in the lines of force about a coil carrying a current is equal to $\frac{1}{2}LI^2$.

The unit of energy or work is the joule. It is the amount of work required to force 1 coulomb of electricity through a 1-ohm resistance. Thus if a 1-μf condenser is charged to a voltage of 500, the energy is

$$\tfrac{1}{2} \times 1 \times 10^{-6} \times 500^2 = 0.125 \text{ joule}$$

Since the power is the rate of doing work, we can find the power required to charge such a condenser in 1 sec of time by dividing the above expression by 1 sec. Thus

$$\text{Power in watts} = \frac{CE^2}{2t}$$

and if we attach the condenser to a source of energy which charges the condenser 120 times a second the power will be $= \frac{1}{2}CE^2 \times 120$ provided that the condenser is permitted to discharge each time it is charged. Thus, the shorter the time, the greater the power required to charge the condenser.

A general expression for such a problem is

$$\text{Power} = \tfrac{1}{2}CE^2N$$

if N is the times per second the condenser is charged and discharged.

A joule is also known as a watt-second. It is a measure of work done, or energy, not power. Power is work per unit of time.

Example 2–7. A condenser in a transmitting station has a capacitance of 0.001 μf and is charged from a 20,000-volt source. What energy goes into the condenser, and what power is required to charge it 120 times a second (60-cycle source)?

Energy or work:

$$W = \tfrac{1}{2}CE^2$$
$$= \tfrac{1}{2} \times 0.001 \times 10^{-6} \times 20{,}000^2$$
$$= 0.2 \text{ joule}$$

Power:

$$P = W \times N$$
$$= \tfrac{1}{2}CE^2N$$
$$= 0.2 \times 120$$
$$= 24 \text{ watts}$$

Example 3–7. Consider the methods of making photographs at very high speeds utilizing a large condenser charged to a high voltage and then discharged through a gas-filled lamp.

The discharge condenser has a capacitance of 112 μf and is charged to 2000 volts. It is then discharged in 1/30,000 second.

The total stored energy, $\tfrac{1}{2}CE^2$, is $\tfrac{1}{2} \times 112 \times 10^{-6} \times 4 \times 10^6 = 224$ joules, or watt-seconds. This work is done in 1/30,000 second, however, and so is equivalent to a power of $3 \times 10^4 \times 224 = 6{,}720{,}000$ watts.

The quantity of charge, CE, is $2 \times 10^3 \times 112 \times 10^{-6} = 0.224$ coulomb. This quantity discharged in 1/30,000 second results in an average current of $3 \times 10^4 \times 0.224 = 6{,}720$ amp.

Problem 5–7. If the source in Example 2–7 charged the capacitor 1000 times per second, what would be the power taken from it?

Problem 6–7. How much energy can be stored in a condenser of 100 μf charged to 1000 volts?

Problem 7–7. How much power is required to charge a 1-μf condenser to a voltage of 220, 120 times a second?

Problem 8–7. A transmitting antenna has a capacitance of 0.0005 μf. It is desired to transmit 1 kw of power. To what voltage must the antenna be charged from a source which charges it 1000 times per second?

Problem 9–7. Suppose that an antenna is to be supplied with 500 watts of power and that between the charging mechanism and the antenna exists an efficiency of 30 per cent. The condenser, which discharges into the antenna, has a capacitance of 0.012 μf, and is charged 1000 times per second. A transformer is used to step up the voltage 110 from the generator to the value required by the condenser. What is the secondary voltage of the transformer, and what is the turns ratio between secondary and primary?

Electrostatic and electromagnetic fields. The energy existing in an inductive circuit is said to exist in the electromagnetic field surrounding the inductance. The field is made up of lines of force, and it can be explored with a compass or by sprinkling iron filings on paper. The energy in an inductive circuit is equal to $\frac{1}{2}LI^2$ joules, where L is henries and I is amperes.

The energy in a condenser is said to exist in the electrostatic field, where the electrical strain exists, that is, in the non-conductors in the vicinity of the conducting surfaces which are charged. The electrostatic field cannot be explored with any magnetic substance, but it can be discovered by any form of charged bodies or any container of static electricity. The energy in a condenser is equal to $\frac{1}{2}CE^2$.

Some circuits, an antenna system, for example, have both capacitance and inductance and, if properly charged, have both a magnetic and a static field. Since the wire can be charged to a very high voltage with respect to earth, considerable energy can be fed into it.

Condensers in a-c circuits. A perfect condenser is one which is an absolute non-conductor to direct currents—that is, it has an infinite d-c resistance—and which has no a-c resistance. All the energy that is put into it is used in setting up an electrostatic field. Unfortunately all condensers have some a-c resistance, and few have infinite d-c resistance. Otherwise a condenser once charged would keep its charge forever.

A perfect condenser placed across an a-c line would take no average energy from the line since for one half cycle it would take energy from the line but during the next half cycle it would deliver this energy back to the line. If the condenser has resistance (and all do), the current taken from the line to charge the condenser also flows through the resistance and in so doing produces a power loss there equal to I^2R. For this reason, a condenser having resistance or a circuit made up of resistance and capacitance will consume energy from an a-c source.

Experiment 1–7. Charge a filter condenser of about 2- to 10-μf capacitance by connecting to a 110-volt d-c circuit. Then discharge by a heavy wire; then charge again and allow to stand for a half hour and discharge. Charge again, and permit to stand for an hour and discharge. The relative sizes of spark give an idea of how poor a condenser it may be from the standpoint of leakage. Then charge and place a 10-megohm resistance across it for a second or so. Then see if it can be further discharged by means of a wire.

A good condenser may retain its charge for many hours after being removed from the source of charging current. The leakage of condensers takes place not only across the insulating material through which its terminals are brought out but also through the wax filling the container and through the dielectric itself. Leakage is caused by an actual movement of electrons between the terminals or through the insulation.

Example 4–7. Since a condenser definitely prevents the flow of d-c current but permits a movement of electrons in an a-c circuit, a condenser may be

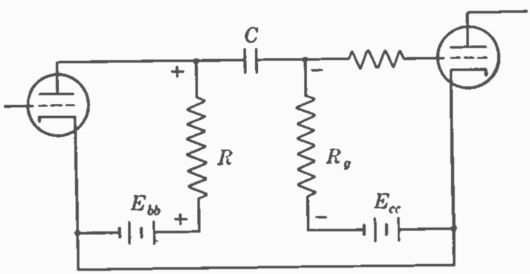


FIG. 1–7. Leakage of current through the condenser C causes a voltage drop across the input of the following tube.

employed to separate direct and alternating currents existing in the same circuit. For example, Fig. 1–7 shows a resistance-capacitance network for coupling the output of one tube to the input of another in an amplifier. The purpose of the condenser C is to prevent E_{bb}, the plate-battery voltage of the preliminary tube, from being impressed upon the grid of the following tube

Any leakage of current through this condenser will impress upon the grid of the second tube a voltage which will be in such a direction as to be highly detrimental to the operation of the circuit.

Figure 2–7 is the equivalent circuit in which r_p is the plate resistance of the first tube. The problem is to find what current flows through the circuit and

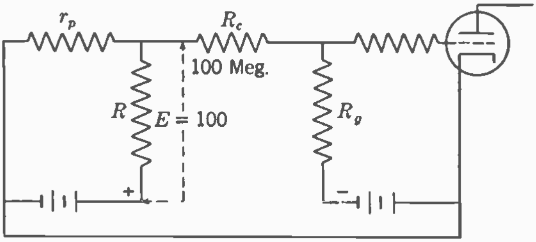


FIG. 2–7. The circuit equivalent of Fig. 1–7.

what voltage drop this causes across the grid leak. Suppose that 100 volts appears across the series circuit composed of R_c and R_g as represented in Fig. 3–7. One hundred volts across 100 megohms (the coupling resistance R and the grid leak resistance are negligible in comparison with the condenser leakage resistance) produces 1 μa of current. This 1 μa flowing through the grid leak of 1 megohm produces a voltage of 1 volt. The polarity of this voltage is opposite to the grid voltage and therefore decreases its value. In addition a certain amount of noise may be developed by the direct current flowing through this condenser and grid leak. Such noise is directly impressed on the input to the amplifier tube and may assume large values when amplified by succeeding stages. Condensers used in resistance-coupled amplifiers must have a high d-c resistance.

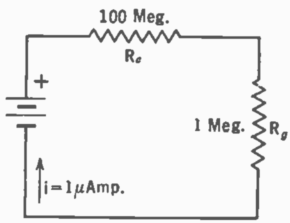


FIG. 3–7. The battery in this figure represents the voltage across R_c and R_g in Fig. 2–7.

The nature of the dielectric. If two square metal plates 10 cm on a side are suspended in air about 1 mm from each other, the capacitance will be about 88.5 μμf. If a sheet of mica is placed between the plates the capacitance will be increased about eight times. Other substances will have different values of the capacitance. Each substance, in fact, will have a certain value of capacitance depending upon what is called the **dielectric**

constant of the substance. The table gives the value of K, the dielectric constant, of several materials.

Material	Dielectric Constant K	Material	Dielectric Constant K
Air	1	Paraffin	2
Bakelite	4 to 8	Porcelain	5 to 6
Celluloid	4 to 16	Rubber	2 to 3½
Glass	4 to 10	Pyrex	5.4
Mica	3 to 7	Shellac	3.5
Oil, castor	4.7	Varnished cambric	4
transformer	2.2	Wood	2 to 8
Paper	2 to 4		

Condenser formulas. Formulas have been worked out by which it is possible to compute the capacitance of condensers of various forms. For example, the capacitance of a flat plate condenser may be computed from the formula

$$C = \frac{8.85 \times A \times K \times (N - 1)}{10^8 \times d}$$

where C = capacitance in microfarads.
A = area of the metallic plate in square centimeters.
K = dielectric constant of the non-conductor.
d = thickness of dielectric in centimeters.
N = number of plates in condenser.

Or

$$C = 0.0885 \frac{KA(N - 1)}{d}$$

where C = capacity in micromicrofarads.
d = thickness in centimeters.
A = area in square centimeters.
K = dielectric constant.

Formulas for other types of condensers may be found in the *Bureau of Standards Bulletin* 74, "Radio Instruments and Measurements," page 235.

The factor K has nothing to do with the ability of an insulating material to withstand high voltages without rupture. Some materials with a high dielectric constant may have low tolerance to high voltages; other materials which can be used in high-voltage condensers may have a low dielectric constant. Mica, with a medium value of K, has a highly prized ability to withstand high voltages. Furthermore, the condition of the dielectric material affects its breakdown voltage. Temperature, percentage of moisture present, and mechanical pressure to which the insulating material is subjected are among the factors that affect voltage breakdown.

Example 5–7. How many plates 16 by 20 cm in area and separated by paraffined paper ($K = 2.1$) 0.005 cm thick are required for a condenser of 24 μf?

$$C = \frac{885AK}{10^{10}d}$$

$$= \frac{885 \times 16 \times 20 \times 2.1}{10^{10} \times 0.005}$$

$$= 0.0119 \ \mu\text{f capacitance per plate}$$

$$\text{Number of plates} = \frac{24}{0.0119} = 2200.$$

Problem 10–7. Express in micro-microfarads: $\frac{1}{1000}$ farad, 1 μf, 0.00025 μf. Express in farads: 500 μf, 0.01 μf, $\frac{1}{10}$ μf. Express in microfarads: 1 farad, 500 $\mu\mu$f, 0.01 $\mu\mu$f.

Problem 11–7. What is the capacitance of two square plates suspended in air 0.1 mm apart, if the plates are 10 cm on a side?

Problem 12–7. A variable tuning condenser has a capacitance of 0.001 μf when air is used as dielectric. Suppose that the container is filled with castor oil. What is the capacitance now?

Problem 13–7. Lead foil plates 2 in. square are separated by mica 0.1 mm thick having a dielectric constant of 6. What is the capacitance of a condenser made up of 200 pairs of such plates?

Problem 14–7. How many joules of energy can be stored in such a condenser as in Problem 13–7 when 100 volts are impressed across it? How much power will it take to charge it to this voltage 120 times a second?

Condensers in series and parallel. When condensers are connected in parallel, the resultant capacitance is the sum of the individual capacitances for the simple reason that the voltage which is applied stores as much charge in each condenser as it would store if it were applied to each condenser separately.

When condensers are connected in series, however, a different condition exists. Figure 5–7 shows two condensers connected across a source of voltage, E. Since the current which charges the condensers must be the same in the two units (the same

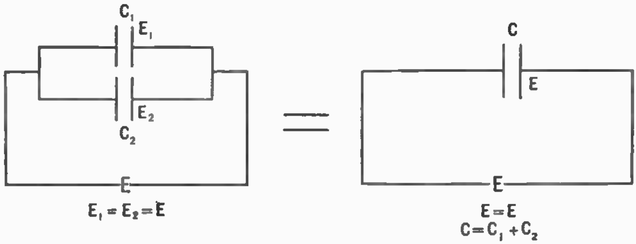


FIG. 4–7. Condensers in parallel.

current must flow through all parts of a series circuit), the same charge (current is the *rate* at which charges flow) exists in the condensers. The voltage across the combination, E, is equal to the charge divided by the resultant capacitance of the series

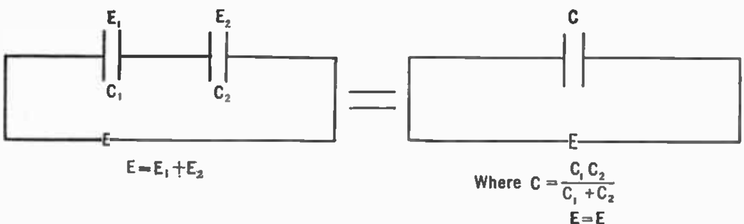


FIG. 5–7. Condensers in series.

circuit. The voltage across each condenser is equal to the charge divided by the capacitance. Thus

$$E = Q \div C$$

$$E_1 = Q \div C_1$$

$$E_2 = Q \div C_2$$

$$E = E_1 + E_2$$

or

$$Q \div C = Q \div C_1 + Q \div C_2$$

or

$$\frac{1}{C} = \frac{1}{C_1} + \frac{1}{C_2}$$

or when only two condensers are involved, the resultant capacitance is equal to $\dfrac{C_1C_2}{C_1 + C_2}$.

Thus it can be seen that condensers in series act like resistances in parallel; condensers in parallel act like resistances in series. When condensers are connected in parallel the voltage is the same across each; when they are connected in series the voltage across each varies inversely as the capacitance. If two equal condensers are in series, half the total a-c voltage is impressed across each condenser. If one condenser has half the capacitance of the other, that one will have two-thirds the total voltage and the other will have one-third the total voltage across it.

The resultant capacitance of several condensers in parallel is always greater than any single capacitance; in series the resultant is less than the capacitance of the smallest of the group. In series the voltage across each condenser varies inversely as its capacitance and is always less than the total voltage across the combination. When a condenser is to be placed in an a-c circuit whose voltage is more than the condenser can tolerate, it is only necessary to use two or more condensers in series so that the voltage across individual members of the group is below the danger limit, and so that the total capacitance in the circuit is the value desired.

When condensers are used across d-c circuits, as in filters, the d-c resistance of the condenser becomes of importance. If two condensers in series across a certain d-c voltage have equal capacitances but different d-c resistances, the voltage drop across the two condensers will differ. The voltage across one of them may be sufficient to destroy it.

Example 6–7. In a radio circuit a 0.0005-μf variable condenser is available but the circuit calls for a 0.00035-μf condenser. What fixed condenser may be used to reduce the maximum capacitance in the circuit to the proper value? How shall it be connected?

Solution. Since the total capacitance is to be reduced, the fixed condenser must be connected in series with the variable condenser. The total capacitance is given as 0.00035 μf. This is equal, from the equation on page 119, to

$$0.00035 = \frac{C_1 \times C_2}{C_1 + C_2}$$

$$= \frac{0.0005 \times C_2}{0.0005 + C_2}$$

whence

$$C_2 = 0.001166 \text{ or } 1166 \ \mu\mu\text{f}$$

Condensers compared to inductors. When voltage is applied to an inductance and a resistance in series, the current through the coil rises slowly from its initial value of zero up to its final value; when a similar voltage is applied to a condenser and resistance in series, the initial current is high and slowly decreases to its final value, zero. The reasons are as follows: The applied voltage across an inductor produces a current in the winding, and the rate of change of this current is greatest at the beginning. The inductance multiplied by the rate of change of the current is numerically equal to the back emf or inductive voltage induced in the coil. By Lenz's law, this voltage is in such a direction that it opposes the change in current. The final value of current, therefore, is delayed.

Voltage applied to a condenser finds no back emf until some current has flowed into it. As current flows into the condenser, charging it, the voltage across the condenser rises, opposing the flow of further current; when the voltage across the condenser due to the charge in it is equal to the applied voltage, no further current will flow into it.

The current flowing into a condenser, therefore, is a function of the rate of change of *voltage* across the condenser; the voltage across an inductance is a function of the rate of change of *current* through the inductance. The higher the frequency of the voltage placed across a condenser, the greater the current that will flow through it; the higher the frequency of the current through an inductance, the higher will be the inductive voltage across it.

Condensers in radio circuits. A condenser or capacitor, then, consists of two or more conducting plates separated by a non-conductor (dielectric). Electric charges can be placed in a condition of non-equilibrium on these plates, say by collecting too many electrons on one plate and producing a corresponding lack of electrons on the other set of plates. As long as this condition exists, energy is stored in the condenser. This energy is due to the *position* of the charges; energy in a magnetic field about an inductance is due to the *motion* of the charges. Either of these funds of energy can be put to use in a manner which will be described later.

Condensers are used for several purposes in radio circuits. They may tune a receiver or transmitter to a certain frequency; they may exclude direct current from a circuit in which alternating current only is desired; they may provide a low-impedance path for certain frequencies; they may be used in power supply systems to smooth out any ripple which remains after rectification.

Condensers take several forms. The capacitance may be fixed or variable; the conducting plates may be exposed to the air or placed within a container and kept under gas pressure, or they may be molded within an insulating compound. The plates may be made of zinc or brass or steel or aluminum. The dielectric may be air, or paraffined paper, or mica or a thin chemical film. The capacitance may be as small as 1 $\mu\mu$f or as great as many microfarads.

The capacitance may vary in a known and desired manner with temperature. Used in connection with inductances which may also have temperature coefficients, circuits which are stable in frequency under varying temperature conditions may be designed.

Electrolytic condensers. When a high capacitance housed in a small space is required, the electrolytic condenser is supreme. In this interesting and valuable condenser, high capacitance is obtained by having large conducting surfaces (the positive plate being a sheet of aluminum foil; the negative being a cloth saturated with a chemical solution) separated by a very thin dielectric (a layer of hydrogen gas). The aluminum plate is often etched chemically or mechanically so that the greatest possible surface is made available. Between the positive aluminum plate and the negative chemical plate is a special conducting fluid which may be more or less solid in form.

These condensers are polarized; that is, they must be connected into a circuit in which there is definite polarity, positive and negative. They have high power losses compared to mica condensers. They are employed in filter circuits in which a d-c voltage exists, the anode being connected to the positive side of the voltage source. The d-c voltage which the condenser can stand without injury depends upon the voltage at which the oxide film was formed in the process of manufacture. The thickness of the film is inversely proportional to the voltage at which the film is formed.

Electrolytic condensers employed in low-voltage circuits may have much greater capacitance in the same space than can be possessed by condensers to be used in high-voltage circuits.

The earth as a condenser. The earth is considered as being at zero potential. All objects not connected to earth have a higher voltage than the earth; they have a capacitance with respect to the earth. The purpose of "grounding" metallic objects in radio receivers or transmitters is to make sure that they are at ground potential so that there will be no currents flowing through the capacitance between these objects and the metal chassis of the receiver or transmitter. The chassis, in turn, is usually grounded either by actually connecting it to earth or by connecting it electrically to one of the power wires that feed energy to operate the equipment.

Condenser tests. If a voltmeter, a battery, and a condenser are connected in series a momentary deflection will be noted if the condenser is good. If the condenser is leaky, a constant deflection will be noted. If the condenser is fairly large, 1 or 2 μf, and no deflection is noted, the condenser is probably open.

If the full battery voltage is read by the meter, the condenser is short-circuited.

If a condenser that has been determined to be not short-circuited is placed across a battery and then a pair of phones is placed across the condenser, a click will indicate that the condenser is good. The click is the discharge of the condenser through the phones.

Problem 15–7. Paper condensers are made up of a thin metallic foil pressed to a dielectric of paper which has the form of a long sheet about 6 in. wide. Suppose that the dielectric constant of such paper is 2 and that it is $3/1000$ in. thick. A condenser is to have a capacitance of 1 μf; the metal foil is 5 in. wide. Two sheets of foil and three of paper are used. How long will each foil sheet be?

Problem 16–7. A "band spread" condenser is one in which many degrees of rotation of the movable elements of the condenser correspond to only a relatively small change in capacitance. One way to make such a band spread is to connect a fixed condenser in series with the variable condenser. Suppose that a variable with a total capacitance of 500 $\mu\mu$f and a minimum of 50 $\mu\mu$f (a change of 10 to 1 in capacitance) is connected in series with a fixed capacitance of 1000 $\mu\mu$f. What capacitance variation is now secured by rotating the variable condenser plates?

Problem 17–7. A condenser of 1 μf has in series with it a resistance of 1 megohm. How long will it take the voltage across the condenser to fall to 37 per cent of its initial value if the circuit is short-circuited?

Problem 18–7. A resistanceless condenser is to be used in a 1000-volt circuit, but the condenser will not stand so high a voltage. Another condenser is placed in series with it, so that the total voltage is divided between the two. One condenser will withstand 200 volts, and the total capacitance is to be 1000 $\mu\mu$f. What will be the capacitances of the two condensers if the voltage across them is inversely proportional to their capacitances?

CHAPTER 8

PROPERTIES OF ALTERNATING-CURRENT CIRCUITS

The two kinds of current in common use are: **direct currents,** which have a more or less constant value and which flow in the same direction all the time; and **alternating currents,** in which not only the magnitude but also the direction are constantly changing.

Definitions used in a-c circuits. When the voltage (or current) has started from zero and risen to its maximum value in one direction, decreased to zero and risen to its maximum value in the opposite direction, and finally come back to its starting value, zero, it is said to have completed a **cycle.** Ordinary house lighting current which has a frequency of 60 cycles per second goes through this cyclic change in magnitude and direction 60 times a second. The **frequency** is the number of times a second a cycle is completed. In a-c circuits the element of time must be considered. In d-c circuits time does not enter; the magnitude of the current is constant.

Alternating currents of nearly all ranges of frequencies exist. Sixty cycles is the common power frequency; tones generated by audio-frequency oscillators for testing purposes may go from almost zero frequency to as high as the human ear can hear, that is, about 15,000 cycles per second, depending on the person. Electric waves of frequencies as low as 15,000 cycles exist. They are generated in the long-wave high-power radio stations carrying on transoceanic communication. Radio frequencies up to about 10,000,000,000 cycles (10,000 megacycles) per second are now used.

The table gives the terminology which has been built up about certain ranges of frequencies now or soon to be in common use.

Frequency Range		Name	Abbreviation
kilocycles	megacycles		
10–30		Very low frequency	VLF
30–300		Low frequency	LF
300–3,000		Medium frequency	MF
3,000–30,000		High frequency	HF
30,000–300,000	30–300	Very high frequency	VHF
300,000–3,000,000	300–3,000	Ultra high frequency	UHF
3,000,000–30,000,000	3,000–30,000	Super high frequency	SHF

In dealing with a-c circuits, we are employing continuously varying currents and voltages. Therefore it is helpful to have somo means of knowing the value of these voltages and currents at any instant.

Let us consider the circle in Fig. 1–8, in which the radius or arm rotates counterclockwise at a constant speed. The position

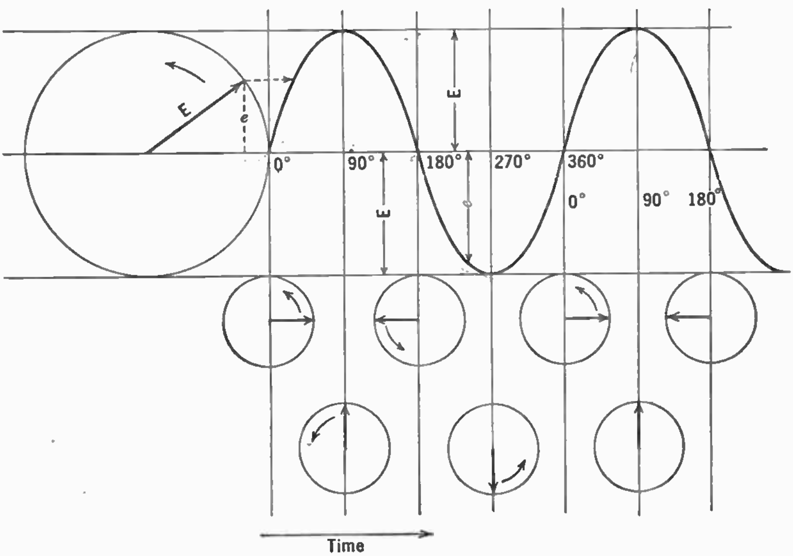


Fig. 1–8. As the vector E rotates, the vertical height, e, of the end of the vector varies according to the curve known as a sine wave. When the vector has the position as shown in the first small circle, the instantaneous value e is zero. At other times the instantaneous voltage has the other values as shown.

of this radius at any instant represents the position of the rotating part of an a-c generator; the length of the radius represents the maximum value of the voltage produced by the a-c generator. The height of the end point of the arm represents the instantaneous value of the voltage at any instant.

Suppose that the generator starts from rest at zero time. The voltage is zero; now the arm begins to rise as the generator rotor moves, and when the rotor has gone through one-quarter of a complete rotation or cycle (90 degrees) the voltage has risen to its maximum value and the end point of the arm has reached its maximum height above the zero axis. At the end of another quarter cycle the arm is again on the zero axis but pointing in the opposite direction and moving down instead of up; another quarter cycle shows the arm pointing down perpendicularly, and its distance below the axis is again a maximum.

If the heights of the end point of the arm are plotted against time expressed in seconds or in degrees, a curve like that shown at the right will result. This curve represents exactly the variation in voltage produced by an a-c generator throughout the cycle of 360 degrees. The angle through which the arm has moved at any instant, reckoned from its zero position, is called the **phase.** Thus, when the arm is vertical, it has gone through 90 degrees of its rotation and the phase is said to be 90 degrees. This is merely a method of expressing time in degrees of rotation instead of in seconds. The value of the voltage at any instant is known as the **instantaneous** value and of course varies from zero to a maximum value in one direction, through zero again, and then to a maximum in the opposite direction.

Sine waves. The curve plotted in Fig. 1–8 is known as a sine wave because of the following facts. The ratio of the vertical height of the end of the radius to the length of the radius is known as the **sine** of the angle between them. Thus, $\sin \theta = e/E$, where θ is the Greek letter "theta." This equation is read "sine theta equals e/E" and means simply that the sine of the angle θ is equal to e divided by E. Sine waves have very important properties, as we shall see later; they are representative of many natural phenomena such as the rise and fall of the tides and the motion of a pendulum. From this expression for the sine of an angle, we may obtain the value for e, the instantaneous value of the voltage in terms of E, the maximum value, at any time we may wish.

Another expression for time in circular measure is the **radian.** The radian is defined as follows. If the length of the circumference of a circle is stretched out straight, it will be found to be 6.28 times as long as the radius of the circle. That is, the circumference divided by the radius is 6.28, and the circumference divided by the diameter is 3.14. Now the ratio between the circumference and the diameter of any circle is known by the Greek letter π (pi) and the numerical value is 3.1416 for all circles. If an angle is marked out in a circle such that the length of its arc is equal to the radius (see arc PQ, Fig. 2–8), this angle is called 1 radian. There are, therefore, 2π radians in a circle, and any portion of the circle may be indicated in radians as well as in degrees.

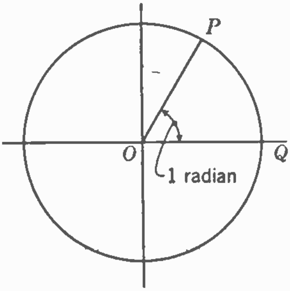


FIG. 2–8. The arc PQ, if stretched out straight, is as long as the radius OP. Then the angle at O is said to be 1 radian.

Triangle functions. Consider the right triangle in Fig. 3–8. Calling the angle between CB and AB theta (θ), the side AC is called the **opposite** side, CB the **adjacent** side, and AB the **hypotenuse.** The relations between the lengths of these sides and the angle theta are as follows:

$$\frac{AC}{CB} = \tan\theta \qquad \text{or} \qquad AC = CB\tan\theta \qquad [1]$$

$$\frac{AC}{AB} = \sin\theta \qquad \text{or} \qquad AC = AB\sin\theta \qquad [2]$$

$$\frac{CB}{AB} = \cos\theta \qquad \text{or} \qquad CB = AB\cos\theta \qquad [3]$$

The tangent (tan), sine (sin), and cosine (cos) are called the functions of the angle θ; if we know any two of the three functions of the right triangle and the angle involved, the other function may be found by means of the table at the end of this book. The cofunctions are, respectively, the cosine (cos), cotangent (cot), and cosecant (csc).

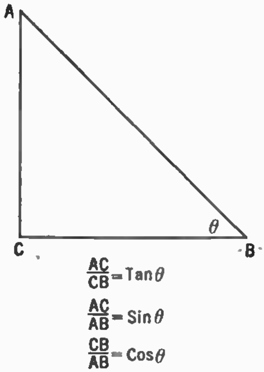


FIG. 3–8. Trigonometric "functions."

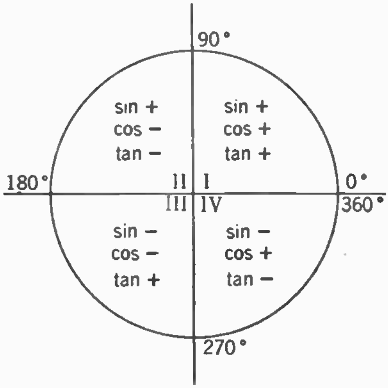


FIG. 4–8. A circle divided into four quadrants with signs of the functions of angles in each quadrant.

The function of an angle greater than 90° may be found from the equation:

Function $(N \times 90° + A)$ = Function A if N is even and cofunction A if N is odd.

Example 1–8. What is the sine of an angle of 195°? 120°?

Solution. This angle of 195° is equal to twice 90° plus 15°. Thus $195° = 2 \times 90° + 15°$, and $N = 2$; since N is even, the sine of 195° is equal to the sine of 15°.

The sine of an angle of 120° is equal to $1 \times 90° + 30°$ = cos 30°, since here N is an odd number.

The numerical sign (+ or −) of the functions varies from "quadrant" to quadrant as shown in Fig. 4–8.

Means of expressing instantaneous values. We may express the instantaneous value of an a-c voltage or current as follows:

$$e = E \sin \theta \qquad \text{or} \qquad i = I \sin \theta$$

where e or i = the instantaneous value.
E or I = the maximum value.
θ = the phase angle in degrees.

Lower-case letters denote instantaneous values; capital letters denote maximum values.

Let us look at Fig. 5–8, which represents the rotating arm E and the vertical height e in four typical cases, 45, 90, 135, and

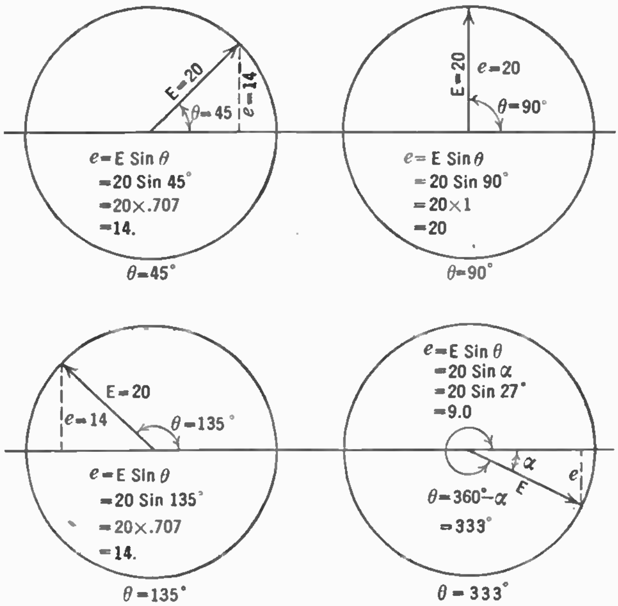


FIG. 5–8. At various times in the cycle the instantaneous value, e, of the vector E is as shown in these vector diagrams.

333 degrees. The actual height compared to the maximum length of the arm E may be calculated by means of the above table and formulas.

Since the sine of an angle of 0 degrees is zero, the instantaneous value of the voltage at 0 phase is zero; since the sine of an angle of 90 degrees is 1, the instantaneous value of the voltage at this point in the cycle is equal to the maximum value; and so on.

The three methods of representing an a-c voltage or current are:

1. By a graphical illustration such as Fig. 1–8, called a sine wave.

2. By an equation, such as

$$e = E \sin \theta \qquad \text{or} \qquad i = I \sin \theta$$

3. By the pictures shown in Fig. 5–8, known as **vector diagrams.** Such a line as E in Fig. 6–8 which moves about a circle is called a **vector;** the vertical distance of its end point from the horizontal axis is called its **vertical component.** The angle θ between the horizontal and the vector is called the **phase angle;** the value of the vertical component may be found by multiplying the maximum value E by the sine of the phase angle.

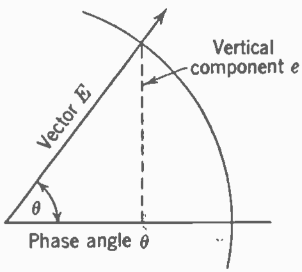


FIG. 6–8. The maximum value of the vector is E; the instantaneous value is $e = E \sin \theta$.

Example 2–8. Represent a voltage whose maximum value is 20 volts. On graph or cross-section paper, 12 divisions to the right equals one alternation and 20 divisions vertically from the horizontal line represents the maximum values of the voltage. The voltage starts at zero, increases to a maximum at 90° or 6 divisions, then decreases to zero at 12 divisions, etc. What is its value at other times? We can tell by using the table of sines. At the 30° phase the instantaneous value, or the height of the rotating arm above the time axis, is $e = E \sin 30° = 20 \times 0.5 = 10$ volts. Other instantaneous values can be found similarly and the entire sine wave plotted similar to Fig. 1–8.

Lay off on cross-section paper a length, say 20 divisions, equal to the maximum value of the voltage. Using this value as a radius, draw a circle. Then the instantaneous value at any time in the cycle can be found by measuring the vertical distance of this point on the circle to the horizontal axis. Thus at the 30° phase the vertical distance is 10 because $\sin 30° = 0.5$, and the instantaneous value is equal to the maximum value (20) multiplied by the sine of 30° (0.5) or $20 \times 0.5 = 10$

A voltage of maximum value 20 may be represented mathematically as

$$e = 20 \sin \theta$$

Problem 1–8. The maximum value of an alternating voltage is 110. What is its value at the following phases: 30°? 60°? 110°? 180°? 270°? 300°? 360°?

Problem 2–8. The instantaneous value of an alternating voltage is 250 volts at 35°. What is its maximum value? What is its value at 135°?

Problem 3–8. The instantaneous value of an alternating voltage is 400 volts at 75°. Plot to some convenient scale its sine wave.

NOTE. In all this discussion voltages or currents can be spoken of with the same laws in mind. Thus the form of a sine wave of current looks exactly like the sine wave of voltage with the same maximum value. The vector diagram looks the same because it is only necessary to label the rotating arm I instead of E and the mathematical formula reads $i = I \sin \theta$ instead of $e = E \sin \theta$. The answers to the above problems will be the same numerically whether we speak of voltage or current.

Effective value of alternating voltage or current. Since the voltage (or current) in an a-c system is rapidly changing direction, and since the needle and mechanism of an ordinary d-c measuring instrument require appreciable time for a deflection, they cannot follow the rapid changes of voltage or current and would only waver about the zero point of the meter even if they could follow the fluctuations. We can, however, compare direct and alternating currents by noting their respective heating effects. An alternating current is said to be equal in value to a direct current of so many amperes when it produces the same heating effect. The value is known as the **effective value** of the alternating current; it is equal to the maximum value multiplied by 0.707 or divided by $\sqrt{2}$. Thus

$$I_{\text{eff}} = 0.707I = I/\sqrt{2} \qquad \text{or} \qquad I_{\text{eff}} = \frac{I}{1.41}$$

where I_{eff} = effective value of an alternating current.
I = maximum value.

The effective value is also known as the **root mean square** or **rms** value for the following reasons. The heating effect of a direct current is proportional to the square of the current. If we take the instantaneous values of the current over a cycle of alternating current, square them, find the average of these squares, and extract its square root, the final value will be equal to the direct-current value that will produce an identical heating effect. The value of current secured in this manner is 0.707 times the maximum value. Since it is the square root of the

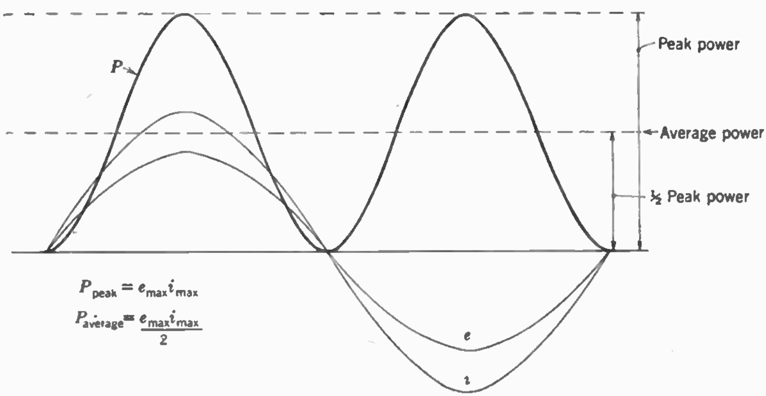


FIG. 7–8. Power is the product of e and i, and the curve marked P is actually secured by multiplying together these factors. Note that the power curve has twice the frequency of the e and i curves.

average or mean squares of several current values, it is abbreviated to the root mean square or rms. An rms voltage is one that will produce a current whose heating effect is the same as a given direct current as discussed above.

$$E_{rms} = E_{eff} = 0.707 E_{max} = \frac{E_{max}}{\sqrt{2}}$$

and

$$E_{max} = \frac{E_{eff}}{0.707} = E_{eff} \times \sqrt{2}$$

Voltage or current is considered as effective unless otherwise indicated or stated.

Example 3–8. What is the effective value of an alternating voltage whose maximum value is 100 volts?

$$E_{\text{eff}} = 0.707 \times 100 = 70.7 \text{ volts}$$

Problem 4–8. The effective value of an alternating current is 12 amp. What is the maximum value?

Problem 5–8. The maximum value of an alternating voltage is 110 volts. What is the effective value?

Problem 6–8. At the 45° phase the instantaneous value of an alternating current is 10 amp. What is its effective value?

Problem 7–8. The effective value of an alternating current is 100 ma. What is the instantaneous value at the 60° phase?

Adding alternating currents. Consider Fig. 8–8, in which there are two radius vectors which rotate at the same speed.

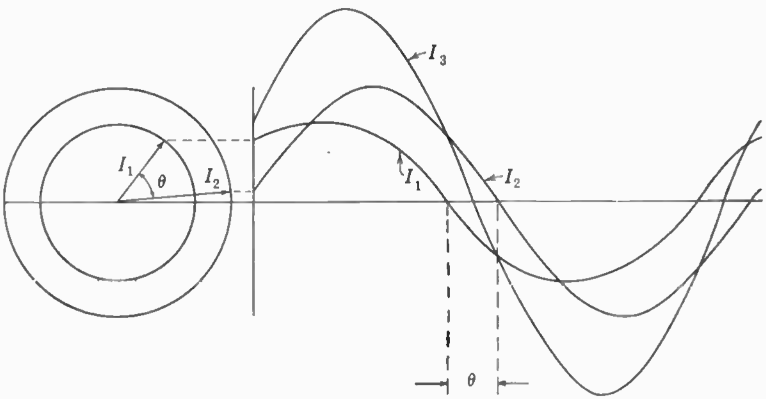


FIG. 8–8. In this curve I_3 is the sum of I_1 and I_2 at each instant.

They represent two alternating currents which have the same frequency but different maximum values and reach their maximum values at different times. The effect of these two currents may be obtained by adding at each instant the vertical heights of the two vectors and by plotting these values as a function of time in degrees or radians or seconds. It will be found that a new sine wave is produced as shown. This characteristic of sine waves is very useful. The phase difference between the two sine waves is constant, but, since the two maxima do not occur at the same instant, it is not possible merely to add the maximum values to get the maximum of the new sine wave. This phase angle must be taken into account.

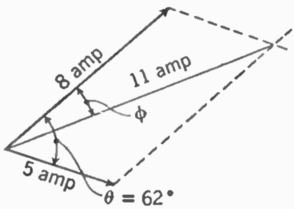


FIG. 9–8. A vector diagram of two sine waves combined to produce a new sine wave of 11 amp maximum value.

If the two vectors are used to form two sides of a parallelogram, the diagonal of the completed figure will represent the maximum value of the new sine wave, and the instantaneous values of the new wave can be found in exactly the same way as they can be found for either component.

The two sine waves need not have the same frequency to be added, but in this case the phase angle is changing continuously and the resulting wave is not a sine wave. To show the effect of two sine waves of different frequencies as well as different maximum values, it

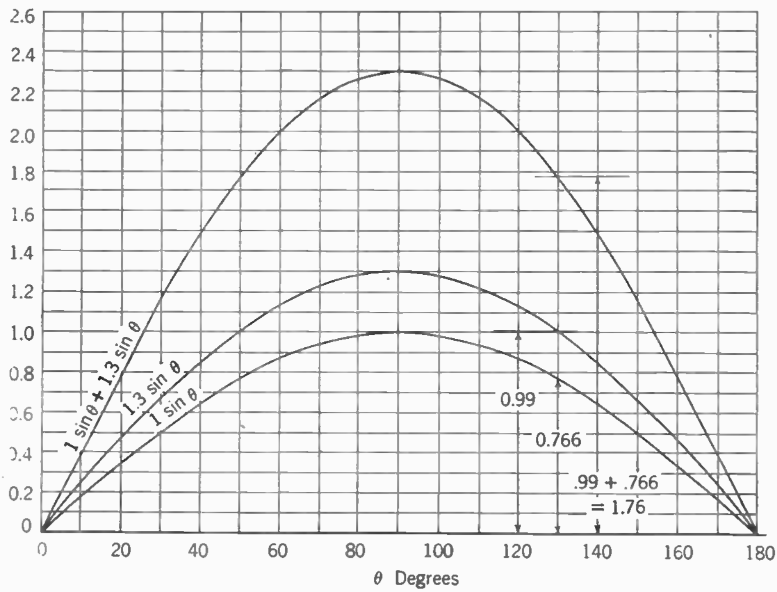


FIG. 10–8. How two sine waves of the same frequency but different amplitudes may be added by means of graph paper.

is only necessary to draw the waves to a convenient scale, to measure the vertical heights of the two waves at each instant, to add these values together, and to plot a new curve representing the values.

Phase relations between current and voltage. Whenever an a-c voltage forces a current through a resistance, the wave forms of the voltage and the current look alike; so do their mathematical formulas and their vector diagrams. This is explained by the fact that the current and voltage start at the same instant, rise to a maximum value at the same instant, and carry on throughout their respective cycles in perfect step, or **in phase.**

When an inductance or a capacitance or any combination of these quantities with each other or with resistance is in the circuit, other phenomena take place which differ entirely from those in a d-c circuit. For example, when an a-c voltage forces a current through an inductance, the current does not attain its maximum value at the same instant as the voltage, but at a later time; when the inductance is replaced by a capacitance, the opposite is true: the maximum value of the current takes place before the maximum voltage is reached.

CASE I. *Current and voltage in phase.* Figure 11–8 represents the current and voltage in phase, i.e., in a resistive circuit.

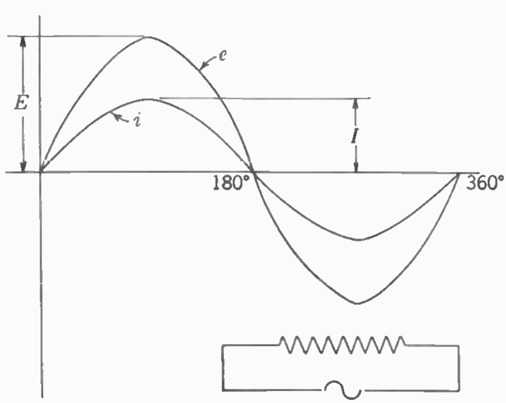


FIG. 11 8. Current and voltage in phase, true of a circuit containing resistance only.

Since the forms of the voltage and current waves are exactly similar they can be drawn on the same horizontal axis, or in the vector diagram as in Fig. 12–8. They can be thought of as two vectors, which may or may not have the same length, rotating at the same speed.

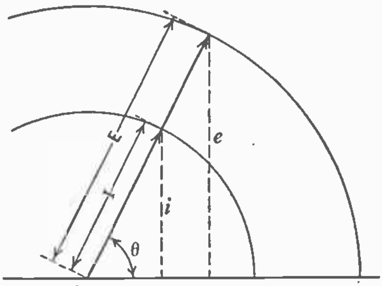


FIG. 12–8. Current and voltage in phase. The maximum value of voltage is greater (or is drawn to different scale) than the maximum value of the current.

In a resistive circuit Ohm's law $I = E \div R$ states the relations between current, voltage, and resistance, just as it does in a d-c circuit.

Example 4–8. Suppose that a lamp of 55 ohms is placed across a 110-volt 60-cycle a-c line. What current will flow through it at the 30° phase?

We must first find the maximum value of the voltage.

$$E = E_{eff} \times \sqrt{2}$$

$$= 110 \times 1.41 = 155 \text{ volts}$$

Since there is no phase effect in the circuit due to the resistance, the current is given by Ohm's law

$$I = E \div R$$

$$= 155 \div 55 = 2.82 \text{ amp}$$

This is the maximum current. The current at any phase may be found by

$$i = I \sin \theta$$

$$= 2.82 \sin 30°$$

$$= 2.82 \times 0.5 = 1.41 \text{ amp}$$

Problem 8–8. A resistance of 10 ohms is in a circuit with an alternating voltage of 20 volts maximum. At what phases will the current through the resistance be 1 amp?

Problem 9–8. A certain alternating current has the same heating effect as a direct current of 8 amp. It flows through a resistance of 25 ohms. What are the effective and the maximum voltage? At what phases will the instantaneous value of the voltage be equal to one-half the maximum value?

CASE II. *Current lagging behind the voltage.* Let us consider the case where the current does not attain its maximum value at the same instant that the maximum voltage is reached, as is illustrated in Fig. 13–8. It will be noted that the current maximum is 67.5 degrees behind the maximum value of the voltage curve. Therefore the maximum value of the current is said to lag behind the voltage maximum 67.5 degrees. The current

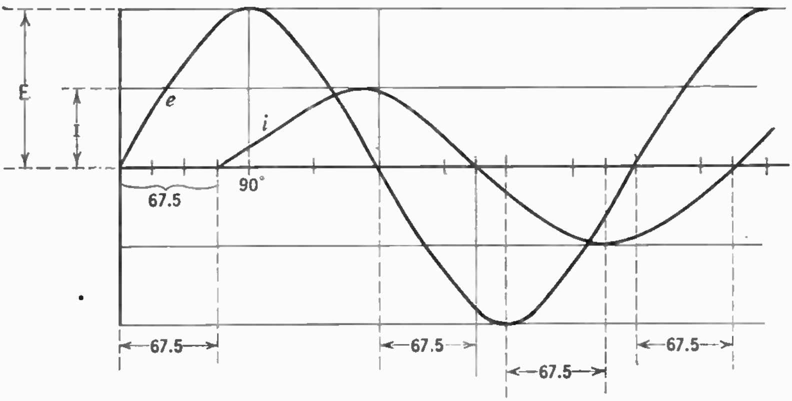


FIG. 13–8. Current and voltage in an inductive circuit where the current lags behind the voltage.

and voltage here may be thought of as two vectors, or arms, moving in two circles such that the current radius does not start until the voltage vector has completed 67.5 degrees of its total movement of 360 degrees.

The formulas for Case II where the current is lagging are

$$e = E \sin \theta$$

$$i = I \sin (\theta - \phi)$$

where θ = phase of the voltage in degrees.

ϕ = difference in phase between E and I, or the angle of lag. (The angle ϕ is called phi.)

Example 5–8. The current lags behind the voltage by 60°. The maximum value of the current is 40 amp. What is the instantaneous value of the current at the 75° phase?

Solution. Lay off on graph paper the voltage at the 75° phase and, 60° behind it, the current which has a maximum value of 40 amp. The vertical component then is equal to the instantaneous value of the current at this phase.

The problem may also be solved by the mathematical formula

$$i = I \sin (75° - 60°)$$

$$= 40 \sin (75° - 60°)$$

$$= 40 \sin 15° = 40 \times 0.26 = 10.4$$

Problem 10–8. If the maximum voltage is 110, what is the instantaneous voltage at the 110° phase? at the 90° phase? at 45°?

Problem 11–8. In an inductive circuit there is a phase difference of 25°. When the voltage is a maximum, the instantaneous value of the current is 10 amp. What is the maximum value of the current?

The cause of lagging current is inductance, which makes the maximum of current take place later than the maximum of voltage. If a circuit is purely inductive (the resistance is negligible) the difference between these maxima is 90 degrees. If there is appreciable resistance the difference is less than 90 degrees.

Why is the maximum value of the voltage ahead in time of the maximum value of the current? The maximum induced voltage across an inductance will occur when the current through the inductance is changing most rapidly, since an inductive voltage is numerically equal to the inductance multiplied by the rate of change of current. The current is changing most rapidly when it is going through zero. Consider a pendulum. When it is exactly vertical its velocity is the greatest; when it is at the peak of its swing to right or left its velocity is zero for a fraction of a second, since at this moment it comes up to its greatest swing and then reverses its direction. At the bottom of the swing, gravity has had its greatest effect and the bob is, at this moment, moving with greatest velocity. Similarly, in an alternating-current circuit, the current is changing most rapidly at the moment it is going through zero.

When alternating current flows through an inductor there will be an alternate increase and decrease in the magnetic field about it. When the field is increasing, a counter emf will be produced (according to Lenz's law) which opposes the increase in current. During this period energy is stored up in the magnetic field. On the next quarter cycle the field collapses, and during this period the rate at which the current is decreasing slows down. This process is repeated during the negative half cycle. It is for this reason that the maxima (or any other corresponding values) of voltage and current in an inductive circuit do not occur at the same instant, the maximum of current occurring after the maximum of voltage.

Inductive reactance. If a coil of negligible resistance is placed in an a-c circuit, less current will flow than if the coil were not present. The effect of the coil which reduces the current is called the **inductive reactance,** which is proportional to the inductance of the coil and to the frequency. Thus

$$\text{Current} = \frac{\text{Voltage}}{\text{Reactance}} \qquad \text{or} \qquad I = E \div X_L$$

where I is in amperes.
E is in volts.
X_L is $2\pi f L$.

Inductive reactance, therefore, is numerically equal to

$$X_L \text{ (ohms)} = 6.28 \times f \times L = \omega L$$

where f = frequency in cycles per second.
L = inductance in henries.
ω = Greek letter omega = $6.28 \times f$.

NOTE: The use of omega here as equal to $2\pi f$ should not be confused with the use of omega as an abbreviation for ohms. It is unfortunate that the same symbol has been associated with both resistance and $2\pi f$.

Example 6–8. In an a-c circuit the following data are given: $E_{\text{eff}} = 100$ volts; inductive reactance, $X_L = 20$ ohms. Find the maximum and effective current and the instantaneous current when the voltage is at the 150° phase.

Solution. The effective current is found from

$$I_{\text{eff}} = \frac{E_{\text{eff}}}{X_L}$$

$$= \frac{110}{20} = 5.5 \text{ amp}$$

$$I_{\text{max}} = I_{\text{eff}} \times 1.41 = 7.8 \text{ amp}$$

The vector diagram in Fig. 14-8 shows the instantaneous current to be

$$i = I \sin (\theta - 90°)$$
$$= 7.8 \sin (150° - 90°)$$
$$= 7.8 \sin 60°$$
$$= 7.8 \times 0.866$$
$$= 6.75 \text{ amp}$$

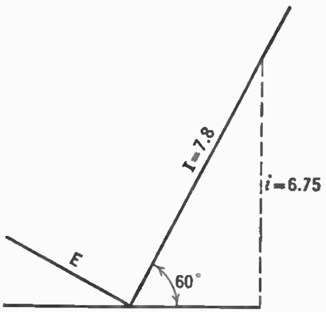


FIG. 14-8. Current lagging behind the voltage by 90 degrees.

Problem 12-8. In the above example find the instantaneous voltage when the instantaneous current is 6 amp.

Problem 13-8. What inductive reactance is needed to keep the maximum current down to 75 amp in a 110-volt effective circuit?

CASE III. *Current leads the voltage.* Here the maximum value of the current is reached before the corresponding maximum voltage is reached. The voltage lags behind the current, or, as it is usually stated, the current **leads** the voltage.

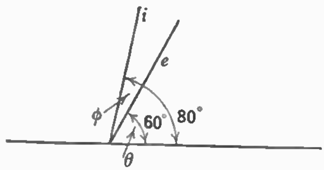


FIG. 15-8. Current leading the voltage by the angle ϕ at the 60° phase. The angle of lead is 20°.

A vector diagram for these conditions is shown in Fig. 15-8. The instantaneous values of voltage and current are

$$e = E \sin 60°$$

$$i = I \sin 80° \qquad \text{or} \qquad i = I \sin (60° + 20°)$$

The formulas applicable when the current leads the voltage are

$$e = E \sin \theta$$

$$i = I \sin (\theta + \phi)$$

where ϕ is phase difference between E and I or the angle of lead.

The current in such an equation will be the maximum current if the voltage is maximum, effective if the voltage is effective, etc.

Example 7–8. The effective current in an a-c circuit is 70 amp. The angle of lead is 30°. What is the instantaneous current when the voltage is at the 10° phase?

The maximum value of the current is found from

$$\begin{aligned} I_{max} &= I_{eff} \times 1.41 \\ &= 70 \times 1.41 \\ &= 98.7 \text{ amp} \end{aligned}$$

By the equation

$$\begin{aligned} i &= I \sin (\theta + \phi) \\ &= 98.7 \sin (10° + 30°) \\ &= 98.7 \times 0.643 \\ &= 64.5 \text{ amp} \end{aligned}$$

Problem 14–8. The instantaneous value of current in a certain capacitive circuit is 8 amp. The instantaneous value of the voltage is 25. The maximum values of the current and voltage are 15 amp and 80 volts respectively. What is the angle of lead between them?

The cause of leading current is capacitance, which makes the maximum of current occur before the maximum of voltage.

When current flows into a condenser it places a charge on the condenser plates and conveys to the condenser a given quantity of electricity. When the current starts to flow into the condenser there is no opposing voltage across the condenser because until the condenser receives a charge the voltage is zero. Therefore the rate of current flow is a maximum when the voltage across the condenser is zero. As current continues to flow into the condenser, the voltage across the condenser rises and opposes further flow of current into it. When the maximum of voltage is attained as the result of the storage of charges in the condenser it exactly balances the impressed emf and at this instant the current flow into the condenser is zero.

The maximum of current, then, corresponds in time to minimum voltage and is said to lead the voltage. The minimum of current occurs when the maximum of voltage is reached. In a purely capacitive circuit (no resistance and no inductance) these maxima are 90 degrees apart. If there is resistance in addition to capacitance the phase difference is less than 90 degrees.

Capacitive reactance. If a capacitance having no resistance is placed in an a-c circuit, less current will flow than if the capacitance were short-circuited. The opposition to the flow of current is called the **capacitive reactance,** and it is found to be inversely proportional to the capacitance and to the frequency. Thus

$$\text{Current} = \frac{\text{Voltage}}{\text{Capacitive reactance}} \quad \text{or} \quad I = \frac{E}{X_c}$$

Capacitive reactance is equal numerically to

$$X_c \text{ (ohms)} = \frac{1}{6.28 \times f \times C} = \frac{1}{\omega C}$$

where f = frequency in cycles per second.
C = capacity in farads.
ω is the Greek letter omega = $6.28 \times f$.

Current leads the voltage in a capacitive circuit because capacity tends to prevent any changes in voltage and so the maximum of current in a purely capacitive circuit takes place 90 degrees ahead of the maximum of voltage. If there is an appreciable resistance in the circuit the difference is less than 90 degrees; thus resistance tends to bring the current and voltage in phase.

Example 8–8. If a condenser which has a capacitive reactance of 5 ohms is in an a-c circuit the instantaneous value of the voltage at the 20° phase is 48 volts. What is the maximum current through the condenser? What is the instantaneous current through it at the 20° phase?

Solution.

$$e = E \sin 20°$$

$$48 = E \times 0.342$$

$$E = 140 \text{ volts}$$

$$I = \frac{E}{X_c}$$

$$= \frac{140}{5} = 28 \text{ amp}$$

$$i = I \sin (20° + 90°)$$

$$i = 28 \sin 110°$$

$$= 28 \cos 20°$$

$$= 28 \times 0.940$$

$$= 26.32 \text{ amp}$$

Problem 15–8. If a condenser in an a-c circuit whose voltage is 110 passes a current of 3 amp, what is the reactance of the condenser in ohms?

Problem 16–8. Condensers are usually rated at the maximum voltage at which they can be operated with safety. What should be the rating of a condenser to be used in a 220-volt a-c circuit?

Problem 17–8. An a-c circuit has a voltage of 115, and a current of 4.5 amp is flowing. It has a condenser in it. What is the reactance of this condenser? What is the instantaneous value of the current when the voltage is 80? At what phase is this?

Comparison of inductive and capacitive reactances. Coils and condensers have opposite effects upon an alternating current. The reactance of an inductor increases as the frequency increases; the reactance of a capacitor decreases as the frequency increases.

A coil which may pass considerable current at 60 cycles may pass practically nothing at 1000 kc. On the other hand, a condenser which will pass considerable current at 1000 kc will pass practically nothing at all at 60 cycles.

Where it is desired to pass a low frequency through a circuit but to exclude a high frequency from it, a shunt condenser and a series coil are placed in the circuit. The series coil prevents the flow of the high-frequency current because of its high reactance, and the capacitance shunted across the circuit provides an easy path for the high frequencies but its high reactance to low frequencies prevents any of the low frequencies from leaking away through this shunt circuit.

Where it is desired to permit the flow of a high-frequency current through a circuit but to prevent a low-frequency current from flowing, the capacitor and the inductor are interchanged, the capacitor being placed in series and the coil in shunt to the circuit.

In these ways advantage is taken of the different effect of coils and condensers upon the flow of alternating currents of high and low frequencies.

Example 9–8. Assume an a-c circuit composed of an inductance of 1 mh. What current will flow if E is 100 volts and the frequency is 600 cycles?

$$X_L = 6.28 \times f \times L$$

Since 1 mh = 10^{-3} henry

$$X_L = 6.28 \times 1 \times 10^{-3} \times 600$$

$$= 3.8 \text{ ohms}$$

$$I = \frac{E}{X_L} = \frac{100}{3.8} = 26.3 \text{ amp}$$

Example 10–8. What is the reactance of a 500-$\mu\mu$f condenser to radio waves of a frequency of 600 kc? Since 500 $\mu\mu$f = 500×10^{-12} farad, and 600 kc = 600,000 cycles

$$X_c = \frac{1}{6.28 \times 600 \times 10^3 \times 500 \times 10^{-12}}$$

$$= \frac{10^9}{6.28 \times 600 \times 500}$$

$$= \frac{10^5}{6.28 \times 30} = \frac{10^5}{188.4} = 530 \text{ ohms}$$

Problem 18–8. What would the current be in an a-c circuit if the frequency were 6000 cycles and if the inductance were 1 henry? If the inductance were 1 μh and the frequency 100 megacycles? Assume $E = 100$.

Problem 19–8. Calculate the reactance of 1 henry, 1 mh, 1 μh at the following frequencies: 100 cycles, 1000 cycles, 1,000,000 cycles.

Problem 20–8. What reactance is needed to keep the current into an electric iron down to 5 amp when it is placed across a 110-volt circuit (assuming that the iron has no resistance)?

Problem 21–8. The inductance in an a-c circuit is 0.04 henry. At what frequency will the current be 3 amp if the voltage is 110?

Problem 22–8. Calculate the reactance of a 1-μf, 0.001-μf, 50-$\mu\mu$f condenser at 60 cycles, 60,000 cycles, 600 kc, 15,000 kc.

Problem 23–8. The capacity in an a-c circuit is 1.0 μf. At what frequency will the current through it be 415 ma if the voltage is 110?

By-pass condensers and choke coils. Condensers that are shunted across a circuit to provide a low-reactance path for the higher frequencies to follow are called **by-pass** condensers. Coils may be used in series with a circuit to prevent the flow of the higher frequencies in that circuit; they are known as **choke coils.** When they are shunted across a circuit, they provide a path for lower frequencies to follow. The proper combination of inductors and condensers will enable certain bands of frequencies to pass or to be excluded.

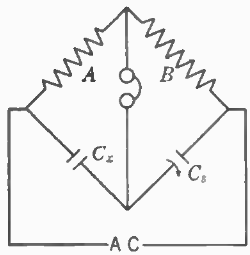


FIG. 16–8. Wheatstone bridge for measuring capacitance. When bridge is balanced, no sound is heard in head-phones, and $C_x = \frac{B}{A} C_s$.

Measurements of capacities. Condensers may be measured for capacitance by comparing them with a known capacitance by means of a Wheatstone bridge as shown in Fig. 16–8. The unknown capacity C_x may then be calculated from

$$C_x = \frac{B}{A} C_s$$

The capacitance of condensers from 0.01 to 10 or more microfarads may be measured by noting the current through them with a known a-c voltage across them. The condenser is first tested for open or short as indicated on page 122. Then, in series with a milliammeter, the condenser is plugged into a light socket (a-c, of course). Then the voltmeter may be put across the condenser, if the voltage of the line is not known. The capacitance is

$$C\ \mu\text{f} = \frac{I\text{ ma} \times 1000}{6.28 \times f \times E}$$

Combinations of resistance with capacitance or inductance. Coils and condensers always have some resistance in them, although in most radio apparatus the resistance is negligible compared to the reactance. Since a reactance as well as a resistance impedes the flow of current, we must combine them to determine what current flows through a piece of apparatus under a certain voltage and at a certain frequency.

Because inductances and capacitances have different effects upon an alternating current from a resistance, we cannot merely add the reactances in ohms to the resistance in ohms to determine the resultant effect upon the circuit. They must be added *vectorially,* not *algebraically.*

Impedance. A combination of resistance and reactance is called **impedance.** The value in ohms may be found as follows. Two factors whose effect is at right angles to each other may be combined and the resultant secured by the formula found in plane geometry which states that "the square of the hypotenuse of a right triangle is equal to the sum of the squares of the other two sides." Thus

$$Z^2 = R^2 + X^2$$

and, if $R = 3$ ohms and $X = 4$ ohms,

$$Z^2 = 9 + 16 = 25$$

whence

$$Z = \sqrt{25} = 5$$

This is called getting the **vectorial** sum. The **algebraic** sum, obtained by simple addition—as when two resistances are combined in series—would be, in this example, 7 ohms, whereas the vectorial sum is 5 ohms. It does not matter in computing whether X is positive or negative since the square of either a plus or minus quantity has a positive sign. Thus

$$(+X)^2 = (-X)^2 = X^2$$

The resultant of combining a resistance and a reactance can be found graphically. Lay off on a horizontal line a number of units corresponding to the number of ohms resistance. Then, if the reactance is inductive, erect a perpendicular and lay off on it a number of units equal to the number of inductive reactance ohms. The length of line connecting the extremities of these two lines is the resultant impedance in ohms.

The reasoning behind this procedure is as follows. Remember that in a resistive circuit the voltage and current are in phase; but that in an inductive circuit the current lags 90 degrees behind the voltage; and that in a capacitive circuit the voltage lags 90 degrees behind the current. Thus, comparing the circuits, we may see that in the inductive circuit the current lags 90 degrees behind the current in the resistive circuit—there is a 90-degree phase difference between these two currents—and similarly, the current in a capacitive circuit is 90 degrees ahead of the current in a resistive circuit.

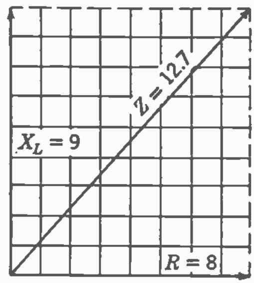


FIG. 17–8. How to compute impedance by vectors using cross-section paper.

Since the effects of inductance and resistance are at right angles (90 degrees) to each other and since the effects of capacity and resistance are at right angles to each other, circuits combining resistance with either inductance or capacitance may be represented in this manner, using two vectors, one pointing at a 90-degree angle to the other. The effect upon the circuit of the two properties may be determined by drawing a diagonal to the parallelogram, two of whose sides are the resistance and the reactance due to inductance (or capacitance).

Because capacitive reactance has an effect opposite to that of inductive reactance, the line representing it should be pointed downward. Graph paper is of great aid in solving a-c problems in this manner.

Example 11–8. An alternating current of 8 amp maximum flows through a coil whose inductance is 0.043 henry and whose resistance is 5 ohms. What voltage is required if the frequency is 60 cycles?

The current in such a circuit is

$$I = E \div Z$$

$$Z = \sqrt{R^2 + X^2}$$

and

$$X = 2\pi f \times L = 6.28 \times 60 \times 0.043 = 16.25 \text{ ohms}$$

$$Z = \sqrt{5^2 + 16.25^2} = \sqrt{289} = 17 \text{ ohms}$$

whence

$$E = IZ = 17 \times 8 = 136 \text{ volts}$$

Example 12–8. What is the impedance in a circuit in which there is a condenser of 1.66 μf and a resistance of 800 ohms? The frequency is 60 cycles.

$$Z = \sqrt{R^2 + X_c^2}$$

$$X_c = \frac{1}{2\pi f C}$$

$$= \frac{10^6}{6.28 \times 1.66 \times 60} = 1590 \text{ ohms}$$

$$Z = \sqrt{800^2 + 1590^2}$$

$$= \sqrt{(64 \times 10^4) + (256 \times 10^4)} \quad \text{(approx.)}$$

$$= 10^2\sqrt{320}$$

$$= 1790 \text{ ohms}$$

General expressions for impedance. If an a-c circuit is composed of resistance and both inductive and capacitive reactance the impedance is figured as follows. Since the inductive and capacitive reactances are opposite in effect, the negative sign is fixed to the capacitive reactance, the positive sign to the inductive reactance. That is, a capacitive reactance is a negative reactance of so many ohms; an inductive reactance is a positive reactance of so many ohms. Before they are combined with the resistance vectorially, their algebraic sum is obtained to get the net reactance. Thus the general expression for impedance is

$$Z = \sqrt{R^2 + (X_L - X_C)^2}$$

and after the capacitive reactance has been combined with the inductive reactance (it is actually subtracted because the signs of the two reactances are different) the form of the impedance becomes as before

$$Z = \sqrt{R^2 + X^2}$$

Again note that the sign before X^2 in the above equation is positive. This sign is always positive since two negative quantities multiplied together (or a negative quantity squared) result in a positive quantity. If, for example, the actual value of the capacity reactance in ohms was greater than the inductive reactance in ohms, the effective reactance in the circuit would be negative but X^2 would be positive.

Example 13–8. What is the impedance of a circuit consisting of a capacitive reactance of 6 ohms, an inductive reactance of 10 ohms, and a resistance of 4 ohms?

The vector diagram for this example is shown in Fig. 18–8. Here we have

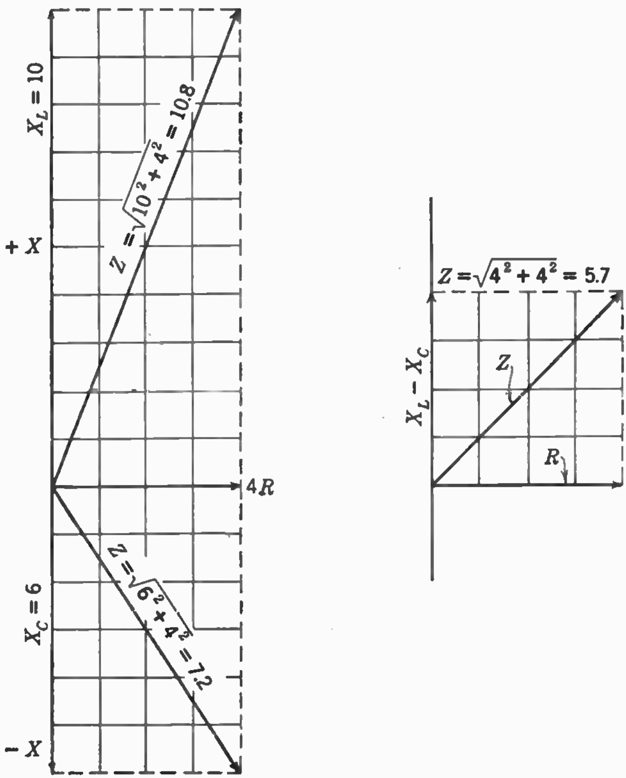


Fig. 18–8. At left is resultant of combining reactances and a resistance to form two Z vectors. At right, the two reactances have been added to form a net reactance of 4 ohms which is then combined with a resistance of 4 ohms to form an impedance of 5.7 ohms.

X_L pointing upward at an angle of 90° from the resistance and X_c pointing downward at an angle of 90° from the resistance. The total effect of the reactances is 10 − 6 or a positive 4 ohms which points upward.

If, however, the values of X_L and X_c are interchanged, so that the resultant of adding the reactances is a negative 4 ohms which point downward, then

$$\text{In Case 1:} \quad Z = \sqrt{R^2 + (X_L - X_c)^2} = \sqrt{4^2 + (10 - 6)^2} = \sqrt{4^2 + 4^2} = 5.7$$

$$\text{In Case 2:} \quad Z = \sqrt{R^2 + (X_L - X_c)^2} = \sqrt{4^2 + (6 - 10)^2} = \sqrt{4^2 + (-4)^2} = \sqrt{4^2 + 4^2} = 5.7$$

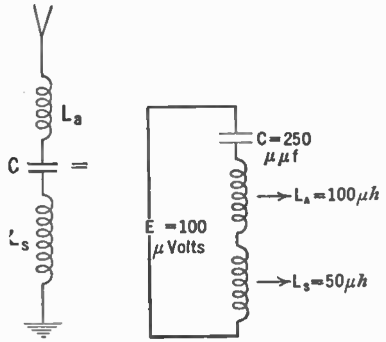


FIG. 19–8. An antenna and its equivalent.

Problem 24–8. An antenna (Fig. 19–8) may be considered as an inductance, L_a and a condenser in series. If the voltage in Fig. 19–8 is 100 μv, f = 1000 kc, what is the current through the coil L_s in series with L_a? (There is no mutual inductance between L_a and L_s.)

Problem 25–8. What is the total reactance in a circuit which has 45 ohms inductive reactance, 70 ohms capacitive reactance, and 20 ohms resistance? What is the impedance? What current would flow if the voltage were 110 effective?

Problem 26–8. What would be the values of capacitance and inductance in Problem 25–8 if the frequency were 500 cycles?

Problem 27–8. What voltage is required to force 1 ma through the following circuit: resistance 8 ohms, inductance 300 μh, capacity 500 μμf, frequency 750 kc?

Series a-c circuits. In Fig. 20–8 is an inductance in series with a resistance. The current flowing in the circuit may be found by dividing the voltage across the circuit by the impedance of the circuit, that is, $I = E \div Z$, in which

$$Z = \sqrt{R^2 + X^2}$$

which is quite different in numerical value from $R + X$. For example, if $R = 3$ and $X = 4$, the vector sum $Z = 5$ whereas the arithmetical sum = 7.

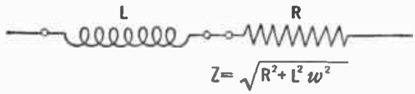


FIG. 20–8. A series circuit: $\omega = 6.28 \times f$.

As in a d-c circuit, the voltage across an impedance, a reactance, or a resistance is equal to that impedance, reactance, or resistance in ohms multiplied by the current in amperes.

Voltage across a resistance	$E_R = I \times R$
Voltage across an inductance	$E_L = I \times X_L$
Voltage across a capacitance	$E_C = I \times X_C$
Voltage across L and C	$E_{L+C} = I(X_L - X_C)$

The voltage across two resistances or reactances in series is the *algebraic* sum of the individual voltages, remembering that a capacitive reactance has a negative sign and that the voltage across it is negative with respect to that across an inductance. The voltage across two impedances in series, however, must be determined by adding the individual voltages *vectorially*. This is because the impedance is a vector sum of a resistance and reactance.

Let us take a typical example. The current in the circuit of Fig. 20–8 is

$$I = \frac{E}{\sqrt{R^2 + X^2}}$$

or

$$E = I\sqrt{R^2 + X^2}$$

$$E^2 = I^2(R^2 + X^2)$$

or

$$E^2 = I^2R^2 + I^2X^2$$

$$= E_R{}^2 + E_X{}^2$$

whence

$$E = \sqrt{E_R{}^2 + E_X{}^2}$$

Therefore the resultant voltage across a resistance and a reactance is the vector sum of the individual voltages.

Example 14–8. If $E = 15$, $R = 3$, $X = 4$

$$I = \frac{15}{\sqrt{R^2 + X^2}} = \frac{15}{\sqrt{9 + 16}} = \frac{15}{5} = 3 \text{ amp}$$

$$E_R = IR = 3 \times 3 = 9 \text{ volts}$$

$$E_X = IX = 3 \times 4 = 12 \text{ volts}$$

$$E = \sqrt{E_R{}^2 + E_X{}^2} = \sqrt{81 + 144} = \sqrt{225} = 15$$

Experiment 1–8. To measure the capacitance of a condenser. Place a condenser of about 10 μf in series with an a-c milliammeter and measure the current through it when placed across a 110-volt, 60-cycle line.

Then

$$E = IX_c = I \times \frac{1}{6.28 \times f \times C}$$

or

$$C = \frac{I}{E \times 6.28 \times f}$$

where E = line voltage or

$$C\ \mu\text{f} = \frac{I}{41.5} = I \times 0.024$$

when $E = 110$.
$f = 60$.
I = ma.

Vector diagrams of series circuit. In a resistive circuit, the voltage and the current are in phase, their maximum values occurring at the same instant. If the circuit is purely reactive (no resistance), there is a 90-degree phase difference between the current and the voltage. If there are both resistance and reactance in the circuit, the angle between the current and voltage is less than 90 degrees, the exact value depending upon the relative values of R and X.

Vector diagrams furnish a convenient means of studying the relations existing in such complex series circuits. If the vector diagram is laid off on cross-section paper, the work is greatly simplified.

Since the current is the same in all parts of a series circuit, the current can be taken as a "reference vector" and the voltages determined from it. As an example let us suppose that we have a circuit made up of 5 ohms resistance and 4 ohms inductive reactance. Through these two components 1 amp of current flows. We will draw the reference vector horizontal and label it I, current. Since, however, the voltage across the resistance is the product of R and I, we can also use this horizontal vector as representing the voltage across the resistance. This line, therefore, represents both the total current in the circuit and the voltage drop across the resistance. Let us lay off 5 units to the right and label it IR. Note that, if we were considering the current and voltage in a reactive circuit, we could not do this because we know that there is a 90-degree phase difference between current and voltage, and the same vector cannot be used to represent both of them. In this example the current and voltage are in phase because only the resistance part of the circuit is under consideration.

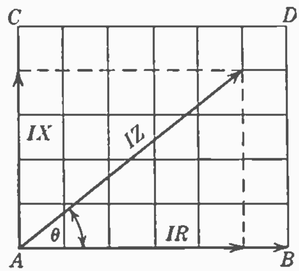


FIG. 21-8. Vector diagram of a series circuit.

The horizontal line, therefore, represents both the current in the circuit and the voltage drop across the resistance. How shall we represent the voltage drop across the inductive reactance?

We know that the voltage across the inductance is 90 degrees ahead of the current through it. We erect a vector perpendicular to the current vector and label it the reactive voltage vector. How long should it be? The actual voltage drop across the inductance will be equal to the product of I and X, and so we lay off 4 units vertically. This is now equal in direction and length to the voltage drop across the inductance, taking the current through the circuit as the reference vector.

What is the applied voltage which will produce 1 amp through 5 ohms resistance and 4 ohms reactance? If we complete the parallelogram as shown by the dotted lines, we will have a representation of the applied voltage both in direction and in magnitude. With cross-section paper and a compass, it is a simple matter to determine the value of the voltage, in this example approximately 6.3.

The 6.3 volts required to force 1 amp through this series circuit must be the voltage across the entire circuit, and the line representing this voltage must also be representative of the impedance of the circuit. Thus, the vector diagram has IR as the horizontal vector, IX as the vertical vector, and IZ as the vector representing the total applied voltage.

Now each of these vectors has been multiplied by a constant factor I, and if I is divided out of each of them we shall have left R, X, and Z. Therefore not only can we use a vector diagram to represent the voltages across component parts of a series a-c circuit, but also the vectors can be made to represent the parts themselves. Furthermore the angle between the vector representing Z and R is the angle by which the voltage across the circuit and the current through it differ in phase. Clearly in this example the voltage leads the current, and the angle between the directions of the R and the Z vectors is the angle of lead. Thus, in Fig. 21–8, $BD \div AB$ is the tangent of the angle θ, or

$$\frac{BD}{AB} = \tan \theta$$

and, since $BD = AC = IX$ (or the voltage across X) and $AB = IR$ (or the voltage across R),

$$\frac{E_X}{E_R} = \frac{IX}{IR} = \frac{X}{R} = \tan \theta$$

The reactance and the resistance in ohms being known, the tangent of the angle may be determined, and the angle itself looked up in a table. When the tangent of an angle is known but not the angle, the expression is written $\theta = \tan^{-1}\frac{X}{R}$ and is read "θ (theta) the angle whose tangent is X over R."

Similarly, $X \div Z = \sin\theta$, and $R \div Z = \cos\theta$.

The effect of a resistance in series with a reactance is to decrease the angle of phase difference between the current and the voltage, that is, to bring them more nearly in phase. In a pure reactance circuit, the angle is 90 degrees; when resistance is added, this angle decreases. In a pure resistance circuit, there is no angle, the current and voltage are in phase; they reach their maximum values at the same instant. If the reactance is capacitive, the procedure is exactly the same except that the X vector is directed downward vertically.

Example 15–8. In an a-c circuit there are a resistance of 10 ohms and an inductive reactance of 8 ohms. A current of 8 amp is flowing. What voltage exists across each part of the circuit and across the entire circuit? What is the phase difference between the current and the voltage?

$$\text{Voltage across } R = IR = 8 \times 10 = 80 \text{ volts}$$

$$\text{Voltage across } X = IX = 8 \times 8 = 64 \text{ volts}$$

Draw the vector diagram to scale as in Fig. 22–8; then the diagonal $IZ = 102.5$ volts. (Note that the algebraic sum of the voltages across R and X is 144 volts.) The tangent of the angle of phase difference is equal to $X \div R$ or

$$\tan\theta = X \div R = 8 \div 10 \text{ or } 0.8$$

$$\theta = \tan^{-1} 0.8$$

or

$$\theta = 38° 40'$$

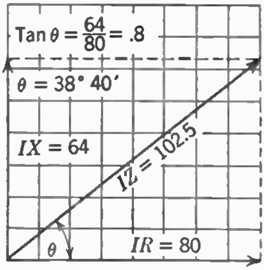


FIG. 22–8. Vector representation of a typical problem in a-c circuits.

Note in Fig. 21–8 that the right-angled triangle has three sides respectively representing the voltage across the resistance, IR, across the reactance, IX, and across the impedance, IZ. Now if each of these voltages is divided by the current, I, three sides of the triangle can be made to represent R, X, and Z. Then, if we know two components of an impedance, or the impedance and one component, the third component (R or X) can be found. Thus

$$\cos\theta = R \div Z \qquad \sin\theta = X \div Z \qquad \tan\theta = X \div R$$

or

$$R = Z\cos\theta \qquad X = Z\sin\theta \qquad Z = R \div \cos\theta \qquad \theta = X \div \sin\theta$$

Case I. X and R given, to find Z. Look up in a table of trigonometric functions the angle whose tangent is $X \div R$; find the sine or cosine of this angle and divide it into X or R.

Case II. X and Z given, to find R. Note that the sine of an angle is equal to the cosine of 90° minus that angle. Thus sin 70° = cos 20°. Since the sum of the angles of a triangle is equal to two right angles (180°) and since one right angle exists between X and R, the other two angles must total 90°.

Now $R = Z\cos\theta$, but θ is not known. However, the angle $(90 - \theta)$ can be found since that is the angle between X and Z, and both are given. The cosine of the angle between X and Z is equal to $X \div Z$. Therefore, if θ is not known, $R = Z\cos\theta$ cannot be solved, but $R = \sin(90 - \theta)$ is known.

Look up the angle whose cosine equals $X \div Z$; at the same time find the sine of this angle and multiply it by Z; or in mathematical language

$$R = Z\cos\theta = Z\sin(90 - \theta) = Z\sin(\cos^{-1} X \div Z)$$

Case III. R and Z given, to find X. This situation resembles Case II, and in mathematical language

$$X = Z\sin\theta = Z\cos(90 - \theta) = Z\cos(\sin^{-1} R \div Z)$$

Example 16–8. A series circuit is made up of resistance = 46 ohms, inductive reactance = 26.6 ohms. What is the impedance?

Solution: $X \div R = 0.577$. This is the tangent of an angle of 30°. The cosine of this angle is 0.866; and $Z = R \div 0.866$ or $46 \div 0.866 = 53$ ohms.

Suppose, however, that Z and X are given, to find R. $X \div Z = 26.6 \div 53 = 0.5$. Look up the angle whose cosine is 0.5. Since it is 60° and the sine of this angle is 0.866, R, therefore, is equal to $0.866Z$ or $53 \times 0.866 = 46$ ohms.

Problem 28–8. Two resistances, one of 8 ohms and the other of 24 ohms, are in a 60-cycle 110-volt circuit. What current flows? What is the current if the frequency is increased to 500 cycles?

Problem 29–8. What current exists in the above circuit if the second resistance is replaced by an inductive reactance of 24 ohms? If the frequency is 60 cycles, what inductance will be in the circuit?

Problem 30–8. In a circuit with 550 volts across the terminals are the following pieces of apparatus: a coil with 15 ohms reactance, a condenser with 7 ohms reactance, two resistances of 10 and 5 ohms. What current flows? What is the voltage across each part? What is the phase relation between voltage and current?

Problem 31–8. In an a-c circuit appear a voltage of 34 volts across a resistance and a voltage of 66 volts across a capacitance. What is the voltage across the combination?

Problem 32–8. Two condensers are in series with two inductances and a resistance. The condensers have reactances of 8 and 10 ohms, the inductances are 20 and 6 ohms, and the resistance is 4 ohms. What current flows, what voltage appears across each component, and what is the phase between current and voltage? Assume that $E = 110$.

Problem 33–8. What is the phase difference in each of the following: (*a*) pure resistance circuit; (*b*) pure inductive circuit; (*c*) pure capacitive circuit; (*d*) 100 ohms resistance and 100 ohms inductive reactance; (*e*) 100 ohms resistance and 50 ohms inductive reactance; (*f*) 100 ohms resistance and 100 ohms capacitive reactance; (*g*) 100 ohms resistance and 50 ohms capacitive reactance; (*h*) 100 ohms inductive reactance, 50 ohms resistance, and 25 ohms capacitive reactance; (*i*) 100 ohms each inductive and capacitive reactance and 100 ohms resistance?

Problem 34–8. In a series circuit there are 45 ohms inductive reactance and 20 ohms resistance. It is desired to increase the phase angle between the current and voltage to 85°. How can this be done? How can the phase angle be decreased to 30°? Solve by means of vector diagrams.

Characteristics of a series circuit. 1. The voltage across a series combination of resistance and reactance is the vector sum of the voltages across the separate units.

2. The combined resistance of several resistances in series is the algebraic sum of the individual resistances.

3. The combined reactance of several reactances, whether inductive or capacitive, is the algebraic sum of the individual reactances.

4. The impedance, or combined effect of a resistance and reactance, is the vector sum of their individual values.

5. The impedance of several impedances in series must be obtained by breaking down the individual impedances into their corresponding resistances and reactances, adding the resistances,

and adding the reactances with due regard to the fact that a capacitive reactance has a negative sign. The total impedance then is equal to

$$Z = \sqrt{(R_1 + R_2)^2 + (X_1 + X_2)^2}$$

Example 17–8. Suppose that we combine two resistances, reactances, etc., of 3 and 4 ohms respectively. The table below shows the resultant values.

Combination	Sum		Resultant
1. $R = 3; X = 4$	$\sqrt{9 + 16}$	$= Z$	5
2. $R = 3; R = 4$	$3 + 4$	$= R$	7
3. $X_L = 4; X_c = 3$	$4 - 3$	$= X$	1
4. $X_L = 3; X_c = 4$	$3 - 4$	$= X$	−1
5. $X_L = 3; X_L = 4$	$4 + 3$	$= X$	7
6. $X_c = 3; X_c = 4$	$-4 - 3$	$= X$	−7
7. $R = 3, X_L = 4, X_c = 3$	$\sqrt{9 + (4 - 3)^2}$	$= Z$	3.16
8. $R = 3, X_L = 3, X_c = 4$	$\sqrt{9 + (3 - 4)^2}$	$= Z$	3.16
9. $R = 3, X_L = 3, X_c = 3$	$\sqrt{9 + (3 - 3)^2}$	$= Z$	3

Note in 7 and 8 above that the resultant is the same although the conditions are different. This occurs because a negative number when squared, or multiplied by itself, becomes a positive number. In other words, a negative reactance and a resistance always produce a positive impedance.

In a series circuit, the reactance which is the greater determines the reactance of the resultant or equivalent series circuit. For example, if there are 4 ohms capacitive reactance and 11 ohms inductive reactance, the circuit as a whole will be inductive. The greatest share of the total voltage will appear across the inductance. Voltages are sometimes called the controlling factors in series circuits.

Resonance. Since reactances may be positive or negative in effect, a very important phenomenon can take place. When the capacitive reactance of a circuit is equal to the inductive reactance, their respective effects cancel out and the resultant impedance is equal to the resistance in ohms alone. To a circuit possessing inductive (or capacitive) reactance one may add capacitive (or inductive) reactance and thereby actually reduce the impedance to the value of the ohmic resistance existing in the circuit. This reduction in impedance occurs at a single frequency only since, at all other frequencies, the inductive and capacitive reactances will be unequal. This phenomenon, known as **resonance,** is so important that the entire following chapter is devoted to it.

Parallel circuits. In a circuit like that in Fig. 23–8, in which several reactances or combinations of resistance and reactance may be connected in parallel, the following rules hold:

The *voltage* across each branch equals the voltage across the combination.

The total *current* taken from the voltage source is the vector sum of the currents through each branch.

The *impedance* offered to the flow of current by the combination is the voltage divided by the total current.

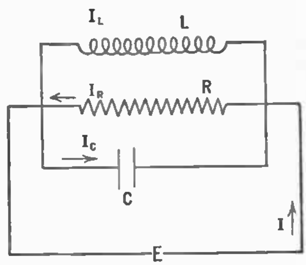


FIG. 23–8. In a parallel circuit the current I may be very small compared with I_L or I_C.

Thus in Fig. 23–8 the current through the entire combination may be found as follows, assuming that $E = 120$, $X_C = 8$, $X_L = 5$, $R = 3$;

$$I_C = \frac{E}{X_C} = \frac{120}{8} = 15 \text{ amp through the condenser}$$

$$I_R = \frac{E}{R} = \frac{120}{3} = 40 \text{ amp through the resistance}$$

$$I_L = \frac{E}{X_L} = \frac{120}{5} = 24 \text{ amp through the inductance}$$

$$I = \sqrt{I_R^2 + (I_L - I_C)^2} = \sqrt{1681} = 41 \text{ amp}$$

$$Z = \frac{120}{41} = 2.92 \text{ ohms}$$

Since the impedance is the ratio between the voltage across the circuit and the current through it, to find the impedance of several branches in parallel we must know the voltage and the current. Often we would like to know the impedance without knowing either the voltage or the current. The procedure then is to assume a voltage, find the currents that would flow, divide the voltage by the total current, and get therefrom the impedance, or

$$Z = \frac{E}{I}$$

Example 18–8. What is the impedance of 630 ohms capacitive reactance shunted by 100 ohms of resistance?

Assume a voltage of 100.

$$I_C = 100 \div 630 = 0.159 \text{ amp}$$

$$I_R = 100 \div 100 = 1.0 \text{ amp}$$

Total current

$$I = \sqrt{I_C^2 + I_R^2}$$

$$= \sqrt{0.159^2 + 1^2} = \sqrt{1.025}$$

$$= 1.015 \text{ amp}$$

$$Z = \frac{E}{I} = \frac{100}{1.015} = 98.5 \text{ ohms}$$

Vectors in parallel circuits. Since the voltage across all elements of a parallel circuit is the same, while the currents differ, we may take the impressed voltage as the reference vector in solving problems by vector diagrams.

As in a series circuit, if resistance only is in the circuit, the current and voltage are in phase. The total current taken from the source by the individual resistive branches is, then, merely the arithmetic sum of the currents taken by the individual branches. If, however, there is reactance as well as resistance, the resultant current must be found in exactly the same way that the resultant voltage is found in a series circuit, with the exception that one uses the voltage as the reference instead of the current.

Example 19–8. A resistance of 40 ohms is shunted by an inductance of 60.4 ohms. What is the impedance of the combination?

Solution. Lay off the horizontal vector as representing the voltage E across the circuit. Since the voltage across and the current through the resistance are in phase, this vector can also represent the current through the resistance, provided that due care is taken to indicate its correct magnitude. In this example the current through the resistance, I_R, is equal to $E \div R$, and as a start one must assume a voltage. It is always easier to assume a voltage value which is a power of 10 (that is, 10, 100, 1000, etc.) since, as in this example, reciprocals are often required, and they can be obtained easily from the slide rule or from a table of reciprocals.

Assuming a voltage of 100, the current through the resistance is $100 \div 40$ or 2.5 amp, and the current through the inductance is $100 \div 60.4$ or 1.66

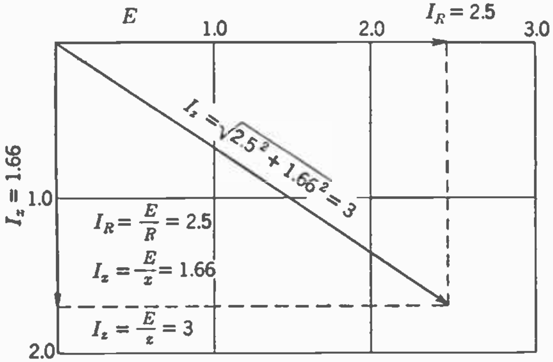


FIG. 24–8. Vector diagram of the parallel circuit of Example 19–8.

amp. The current through the resistance is laid out, 2.5 units along the voltage vector, and since the current through the inductance must be behind the voltage vector by 90 degrees, the inductive current is laid out on a vector vertically downward, 1.66 units long. When the parallelogram is completed, the length of the diagonal is the magnitude of the total current, and the angle between the current taken by the resistance and the combined current is the phase angle or angle by which the total current lags behind the impressed voltage. Note here that the length of the diagonal of the parallelogram does not represent the magnitude of the impedance of the parallel circuit but the impressed voltage divided by the impedance. Since the total current (represented by the length of the diagonal) is 3.0 units, the impedance is $100 \div 3$ or 33.3 ohms.

Power in a-c circuits. In a d-c circuit the power is the product of the voltage across the circuit and the current through it.

Thus, if 1 amp is fed into a device under a pressure of 100 volts, the power used is 100 watts.

In a-c circuits the voltage and current are not always in phase. In fact in many circuits there is a decided difference in phase between the current and voltage. What is the power?

The power at any instant is the product of the instantaneous current and the instantaneous voltage. Thus in Fig. 25–8, where the voltage and the current are in phase, as in a resistance circuit, the height of the voltage line e above the horizontal time axis multiplied by the height of the current line i above this axis gives the instantaneous power. This is plotted in the curve P.

When, however, the current and voltage are not in phase, as in an inductive (Fig. 26–8) or capacitive (Fig. 27–8) circuit, a different-looking curve results although the instantaneous power is still the product of the instantaneous values of current and voltage. The part of the power curve marked B is interesting. It is the result of multiplying a positive current (or voltage) by a negative instantaneous value of voltage (or current). The product is negative because the product of a negative and a positive number is a negative number. The power represented by the product of a positive current (or voltage) and a negative voltage (or current) is considered a negative power.

Power actually consumed in a circuit is considered positive power. Negative power is power that is returned to the generator from the line. Power is returned to the generator only when there is reactance in the circuit. A pure resistance consumes the entire amount of power fed to it by the generator; a pure reactance, that is, a coil or a condenser having no resistance, consumes no power from the source since it returns to the generator on one half cycle as much power as it received on the other half cycle. The instantaneous power, however, represents on one half cycle the energy going from the generator to establish the magnetic field about the coil or the electrostatic field in the condenser, and in the next half cycle the instantaneous power (again the product of the instantaneous volts and amperes) is returned to the generator. A reactance that has some resistance in it,

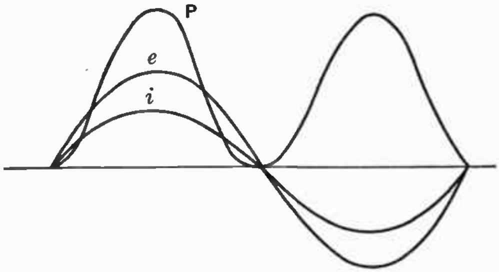


FIG. 25–8. In a resistive circuit, all power delivered by the generator is consumed by the external circuit. Voltage and current are in phase. The curve of instantaneous power has twice the frequency of the voltage and current curves.

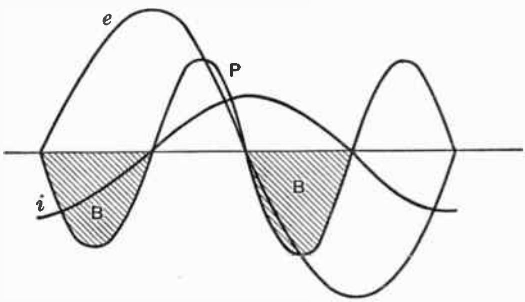


FIG. 26–8. In a purely reactive circuit, the average power is zero, as much being delivered back to the generator (shaded areas) as is delivered to the circuit. The curve of instantaneous power is a double-frequency curve.

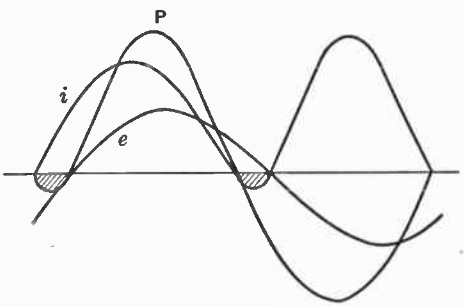


FIG. 27–8. In this case the circuit is slightly capacitive; some power is stored in the capacitor, but most is used up in the resistance of the circuit.

and all do, consumes only the amount of power represented by the I^2R loss in the resistance.

The **apparent power** taken by a load on an a-c circuit is

$$P_A = EI$$

The **actual** (or true) **power** is that taken by the resistive component

$$P = I^2R$$

The ratio between the true power to the apparent power is known as the **power factor** of the circuit.

$$\text{P.f.} = \frac{P}{P_A} = \frac{I^2R}{IE} = \frac{IR}{E}$$

Since $E = IZ$

$$\text{P.f.} = \frac{IR}{IZ} = \frac{R}{Z}$$

The power factor of a circuit, or of a piece of apparatus, is a measure of its resistance compared with its impedance, or vice versa. In radio circuits one is usually more interested in the reactance and there is every desire to keep the resistance low; but in power circuits the opposite is true, at least as far as the power companies are concerned. One uses a condenser or an inductor in a radio circuit because one wants a condenser or an inductor; whatever resistance the component has, in addition to its reactance, is a nuisance because it consumes power. A pure inductor or a pure capacitance will not take any power from the line; thus, a pure capacitance or inductance plugged into a light socket will not cause the electric meter to turn. The current the device takes from the line merely sets up a field which is given back to the line when the direction of flow of current changes. A flow of current through the line, however, is a positive thing; it flows through the resistance of the power line, resulting in a voltage drop and a power loss, and it is a direct waste as far as the power company is concerned.

The effort of the power-company engineers, therefore, is to secure as high a power factor as possible. On the other hand, radio engineers strive for the lowest possible power factor. One of the characteristics which makes mica such an excellent dielectric for condensers is its low power factor, 0.02 per cent at radio frequencies. A condenser made from it has very low losses.

Now on page 156 it was stated that R/Z is equal to $\cos \theta$, where θ is the angle between the resistive component of an impedance and the impedance. Therefore another expression for the power factor is $\cos \theta$, and the true power taken by a circuit is

$$P = EI \cos \theta$$

where E and I are the rms values of voltage and current.

Since the cosine of an angle has values between zero and 1, the power factor may have any value between zero and 1 and may be expressed as an angle or as a percentage. Thus unity power factor represents a pure resistance or a load with 100 per cent power factor. In a purely reactive circuit the power factor is zero.

Example 20–8. A 220-volt a-c motor takes 50 amp from the line; but a wattmeter in the line shows that 9350 watts is taken by the motor. The apparent power is the product EI or 11,000 watts; the power factor is true power divided by apparent power = 9350 ÷ 11,000 or 0.85 or 85 per cent. What does this mean?

Power W taken by motor from line = $EI \cos \theta$ = 9350
Current taken from line = $W \div E \cos \theta$ = 50 amp
But the *useful* current = $W \div E$ = 42.5 amp

Therefore, 50 − 42.5 = 7.5 amp flows through the line and from the generator supplying the system, which would not be taken if the machine had unity power factor (like a pure resistance). This extra current represents a waste because of the voltage loss in the resistance of the lines and a power loss in the resistance of these lines.

Example 21–8. A single-phase induction motor with 440 volts across it delivers 12 hp. The motor is 89.5 per cent efficient and has a power factor of 84 per cent. What current is taken from the line, and what power is taken from the line?

$$\text{Power taken by motor} = \frac{\text{Power delivered}}{\text{Motor efficiency}} = \frac{12 \times 746}{0.895} = 10{,}000 \text{ watts}$$

This is the real power taken from the line,

$$\text{Current taken from line} = \frac{\text{Real power}}{E \cos \theta} = \frac{10{,}000}{220 \times 0.84} = 54.25 \text{ amp}$$

If the motor had unity power factor (no reactance) it would require $W \div E$ or 45.5 amp from the line. The difference, 8.75 amp, flows through the generator at the power plant, through the line and feeder transformers between the plant and the motor, through the wires in the factory using the motor, and finally through the wires of the motor.

Knowing the power factor, one can determine the ratio of the resistance to the impedance of the circuit; but one cannot find the absolute value of the ohmic resistance directly. For example, a motor may have a power factor which is more or less independent of the current drawn or the power delivered. That is, its ratio of resistance to impedance is a constant. The current taken from the line and the power taken from the line will depend upon the power the motor is called upon to deliver. If one can call the effective resistance of the motor the value of a pure resistance taking an equivalent amount of power from the line, i.e., R equals power divided by I^2, then one can determine this value of R, and since the ratio between R and Z is known (power factor) all three elements of the circuit, R, Z, and X, can be found. However, not all this R will be made up of the ohmic resistance of the windings of the motor—in fact if much of it is, the motor is inefficient.

Problem 35–8. The reactance of a condenser is 300 ohms at 680 kc. What is its reactance at 1200 kc? *Hint:* Use simple proportion.

Problem 36–8. What current will flow through a 4-μf condenser when placed across a 400-volt source of 60-cycle voltage?

Problem 37–8. A resistance of 40 ohms is in series with a capacity of 60 $\mu\mu$f. What is the impedance at 28 megacycles? If 100 volts is placed across this series circuit, what voltage will be across each element?

Problem 38–8. In a series circuit containing resistance and inductance, the values of these two components in ohms is equal at a certain frequency. What is the impedance at this frequency in terms of the resistance? What will be the impedance of the circuit at twice this frequency?

Problem 39–8. A series circuit composed of $R_p = 10{,}000$ ohms, $R_o = 20{,}000$ ohms, and $L = 100$ henries has 100 volts at 100 cycles across it. What is the voltage across L and R_o in series?

Problem 40–8. What is the impedance of 240 ohms resistance shunted by an equivalent value in capacity reactance?

Problem 41–8. Electric lamps with the transformers supplying them commonly operate at a power factor of about 95 per cent. What power is taken from a generator by eight lamps each taking 5 amp at 115 volts?

Problem 42–8. A by-pass condenser across a 500-kc circuit has a capacitance of 0.01 μf. If the voltage across the condenser is 100, and if the ratio of the reactance to the resistance is 40, what is the current taken by the condenser, what is its power factor, and how much power does it consume?

Problem 43–8. A certain loud speaker has an inductance of 1.5 henries and a resistance of 3000 ohms. What power will it take from a 100-volt, 100-cycle source? What is its power factor?

CHAPTER 9

RESONANCE

The most important circuits in radio are those in which either series or parallel resonance occurs. In transmitting and receiving systems resonance is used to build up large voltages and currents at certain desired frequencies and to discriminate against undesired signal frequencies by keeping their voltages and currents low. When one tunes a radio receiver, he actually adjusts the a-c circuits within the receiver so that a condition of resonance exists. Everyone who has operated a receiver has, in tuning it, performed one of the most interesting experiments in all a-c theory and practice. It is necessary that we look into the phenomenon of resonance very closely.

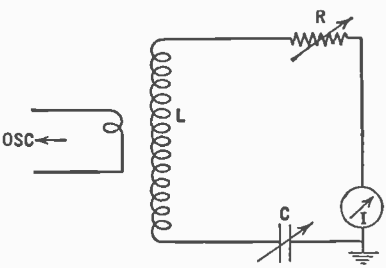


FIG. 1–9. When L is coupled loosely to an oscillator and C is varied, the current indicated at I will go through a maximum like that in Fig. 3–9.

Series-resonant circuit. Although a general idea may be obtained of what takes place in a resonant circuit when a radio receiver is tuned, a much more exact idea may be had as a result of a laboratory experiment.

Experiment 1–9. The purpose of this experiment is to determine the effect of changing the capacitance and the frequency in a resonant circuit. Connect in series, as in Fig. 1–9, a coil of about 200 μh, a variable capacitance of about 1000 $\mu\mu$f, a resistance of about 10 ohms, and a radio-frequency ammeter. Couple the inductance loosely to a r-f generator so that at resonance the maximum current will produce a reading at the top of the r-f ammeter scale. Then vary the capacitance of the condenser, noting down the current that flows at each value of capacitance.

Plot these data as a curve. Now set the condenser so that it has the proper capacitance to produce maximum current at the frequency of the generator. Then vary the frequency of the generator above, through and below resonance, plotting the data secured. Change the value of resistance and repeat both experiments.

Calculate the voltage across the condenser and across the coil and the phase angle between current and voltage, and plot the data for both experiments performed above.

Experiment 2–9. Connect in series with a lamp an inductance of several henries. Add sufficient resistance so that the lamp does not light when placed across a 110-volt, 60-cycle line. Then put a condenser (of the filter type) in series with the resistance, the line, and the lamp. Add more capacitance until the lamp lights up. The added capacitance has brought the circuit to resonance so that the only hindrance to the flow of current was the lamp and the resistance. Adding something to the circuit actually made more current flow.

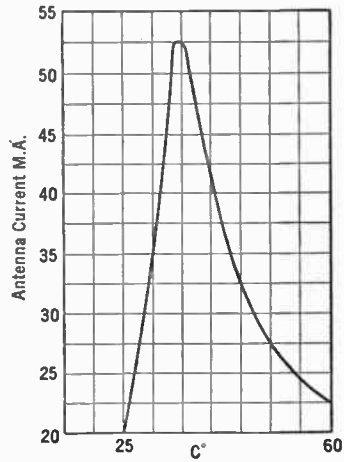


Fig. 2–9. How the antenna current of a radio station varies as the series antenna condenser is varied.

The curve in Fig. 3–9 shows what happens as the voltage across a series circuit is kept constant while the frequency is increased. At first the current increases slowly, then as the resonant frequency, 356 kc, is approached the current increases very abruptly and after passing through a sharp maximum at 356 kc falls very rapidly at first and then more slowly. The voltages across coil and condenser go through similar changes. The phase between current and voltage changes also, being a negative angle (current leading voltage) below resonance, being zero at resonance (current and voltage in phase), and becoming a positive angle above resonance (current lagging behind voltage).

At zero frequency, that is at direct current, the current in such a circuit will be zero because the condenser will not permit direct current to pass. At very low frequencies, the reactance of the condenser is very high, so that little current will flow. At very high frequencies the reactance of the coil becomes very great and therefore little current will flow. At intermediate frequencies more current flows.

When a series circuit is resonant, the current and voltage are in phase, the current is a maximum, the impedance is a minimum, and the voltages across the condenser and the inductance are equal and opposite in sign and greater in value than the voltage impressed across the combination. This voltage may become so high that the condenser will be punctured. The voltage across the coil or the condenser at resonance is equal numerically to the voltage across the entire circuit multiplied by the factor X_L/R or X_C/R, ($\omega L/R$ or $1/\omega CR$), often called the Q of the circuit. This factor, Q, is an important and useful "operator," which will be employed many times in circuit design. If R_s is the resistance in series with the tuned circuit,

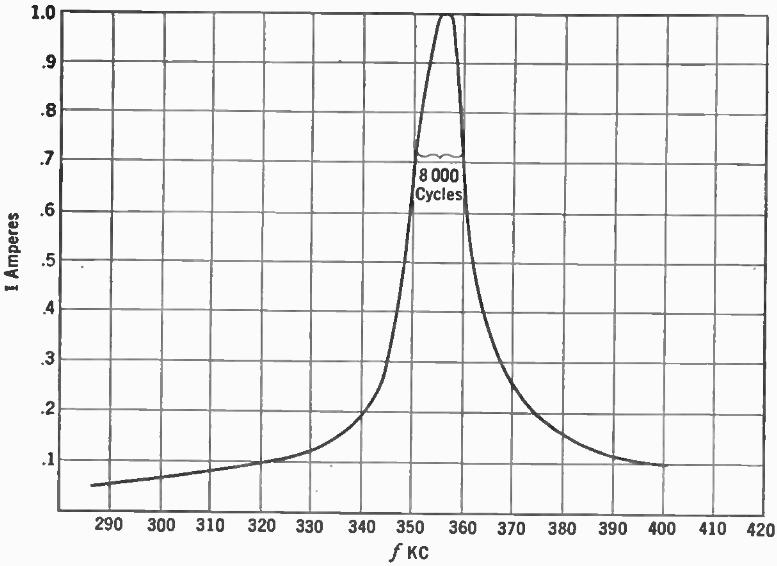


FIG. 3-9. The resonance curve of a circuit like that of Fig. 1-9.

$Q = \frac{1}{R_s}\sqrt{L \div C}$, and if R_p is the resistance shunted across the tuned circuit, $Q = R_p\sqrt{C \div L}$.

A curve showing how the current in the circuit changes as the variable factors are changed, that is, a graph of I against the capacitance, the inductance, or the frequency, is called a **resonance curve** and is symmetrical about the resonant frequency if the circuit is adjusted by changing the inductance, and is dissymmetrical when the capacitance or the frequency is the variable factor.

Characteristics of series-resonant circuit. Below the resonant frequency the reactance is mainly capacitive; above this frequency the circuit is mainly inductive, that is, the capacitive reactance is the main deterrent to the flow of current below resonance; above resonance the inductance offers the greatest opposition to the flow of current. For a narrow band of frequencies in the neighborhood of 356,000 cycles in Fig. 3–9 the total impedance of the circuit is less than 100 ohms. Far from the resonant frequency the reactance and impedances are much greater, and very little current will flow.

Below resonance, where the capacitive reactance predominates, the current leads the voltage; at resonance the current is in phase with the voltage; above resonance the current lags behind the voltage. At resonance, therefore, there is an abrupt change in phase between the voltage across the circuit and the current taken by it from the generator.

At all frequencies the voltage across the inductance is 90 degrees ahead of the current and the voltage across the condenser is 90 degrees behind the current. Between the two reactive voltages, then, is a 180-degree phase difference; that is, they are exactly out of phase. Their resultant may be found by subtracting the lesser from the greater. The resultant of combining them with the voltage across the resistance must be the voltage across the entire series circuit, which is the vector sum. Thus,

$$E = \sqrt{E_R^2 + (E_L - E_C)^2}$$

At any frequency but that of resonance, one of the two reactive voltages is greater than the other. At resonance, however, the two voltages are equal in magnitude and opposite in phase so that the resultant of combining the reactive voltages is zero. When added vectorially to the IR drop in the circuit, the resultant is the voltage which is impressed by the external source.

The vector sum of the reactive and resistive voltages is equal to the impressed voltage.

At resonance the reactances are equal to each other and equal to $\sqrt{L \div C}$, i.e., $X_L = X_C = \sqrt{L \div C}$.

Example 1–9. What are the voltages and phase relations in a series circuit like that of Fig. 1–9 at a frequency of 370 kc if I, X_C, X_L and R are as follows?

$$I = 0.274 \text{ amp}$$

$$X_C = 430 \text{ ohms}$$

$$X_L = 466 \text{ ohms}$$

$$R = 10 \text{ ohms}$$

$$E_R = I \times R = 0.274 \times 10 = 2.74 \text{ volts}$$

$$E_C = I \times X_C = 0.274 \times 430 = 118 \text{ volts}$$

$$E_L = I \times X_L = 0.274 \times 466 = 128 \text{ volts}$$

$$E_{R+L} = \sqrt{2.74^2 + 128^2} = 128 \text{ volts (approx.)}$$

$$\phi_{R+L} = \tan^{-1}\frac{128}{2.74} = \tan^{-1} 46.6 = 88.46°$$

$$E_{R+C} = \sqrt{2.74^2 + 118^2} = 118 \text{ volts (approx.)}$$

$$\phi_{R+C} = \tan^{-1}\frac{118}{2.74} = \tan^{-1} 43 = -88.38°$$

$$E_{L+C} = E_L - E_C = 9.6 \text{ volts} = E_X$$

$$E = I\sqrt{R^2 + X^2} = \sqrt{E_R^2 + E_X^2} = \sqrt{2.74^2 + 9.6^2}$$

$$= 10 \text{ volts}$$

$$\phi_{R+L+C} = \tan^{-1}\frac{X}{R} = \tan^{-1}\frac{X_L - X_C}{R} = \frac{466 - 430}{10} = 3.6$$

$$= 74° \, 30'$$

For example, the reactances of a capacitance of 1000 $\mu\mu$f and an inductance of 200 μh at resonance may be found by

$$X_L = X_C = \sqrt{\frac{L}{C}}$$

$$= \sqrt{\frac{200 \times 10^{-6}}{1000 \times 10^{-12}}}$$

$$= \sqrt{0.2 \times 10^6}$$

$$= \sqrt{0.2} \times 10^3$$

$$= 0.447 \times 10^3$$

$$= 447 \text{ ohms}$$

At resonance the inductive reactance and the capacitive reactance in the equation for the impedance $Z = \sqrt{R^2 + (\omega L - 1/\omega C)^2}$ cancel out, that is, $\omega L - 1/\omega C = 0$ so that the resultant impedance is the resistance alone,

$$Z = R \text{ (at resonance)}$$

Reactance diagrams. One of the best ways to see what is occurring in a resonant circuit is to plot the reactances of the inductance and the capacitance as the frequency is increased, remembering that the capacitive reactance is negative. As the frequency increases, the negative reactance of the condenser decreases, approaching, but never reaching, zero reactance. At the same time the positive reactance of the inductance increases. The sum of the two reactances changes from a high negative value to a high positive value, and at one frequency is zero.

The effect of quite complex circuits can be determined in this manner, the reactance of each element being plotted versus frequency and these reactances being added properly to determine the combined effect.

Effect of resistance on series-resonant circuit. At resonance the magnitude of the current in the circuit is controlled solely by the resistance. Its effect is most important in any radio

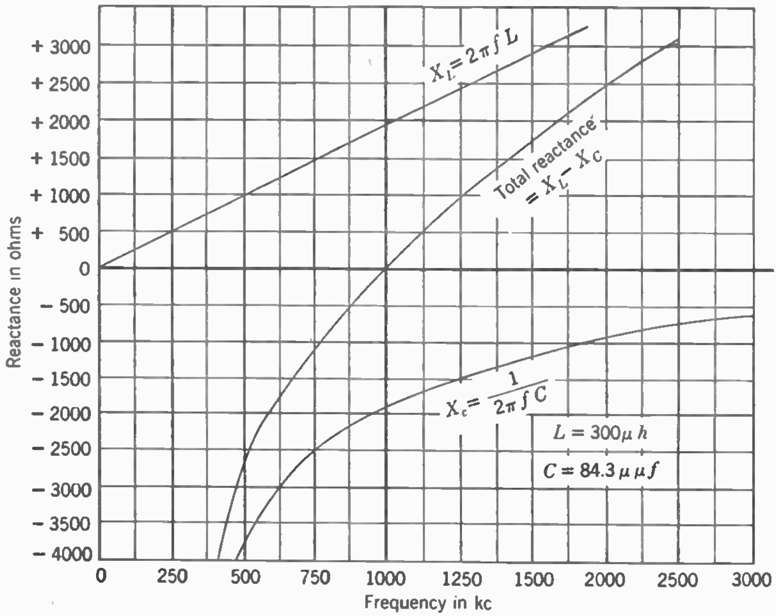


Fig. 4–9. A plot of the reactance of a coil and a condenser in series. X_L rises in direct proportion to f; X_C rises inversely as f. The sum, $X_L - X_C$, changes from a high negative value where X_C is greater than X_L, through zero where $X_L = X_C$, to a high positive value where X_L is greater than X_C.

circuits where resonance plays a prominent part. The curves in Fig. 5–9 show the effect of adding resistance to the circuit of Fig. 1–9. The voltages across the condenser and the inductance depend upon the resistance of the circuit. They are greater the smaller the resistance, because the voltage across these reactances is equal to the product of the reactance and the current. The current, controlled entirely by the resistance at resonance, in turn produces greater voltages across the reactance when less resistance is in the circuit. If E is the voltage impressed on the whole circuit, the voltage across the condenser is $E \div \omega CR$ and that across the inductance at resonance is $E \times (\omega L/R)$. The voltage across either L or C is equal to EQ. Note how often the factor Q enters into radio circuits. Here it

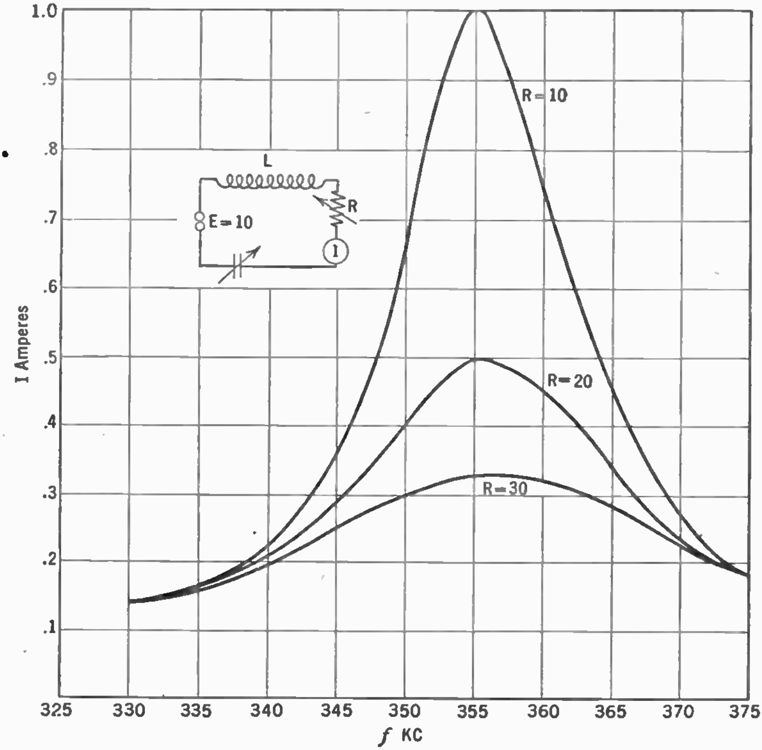


FIG. 5–9. Effect of resistance on a resonance curve. Note that the current far from resonance is not changed so much as the resonant current.

shows that the voltage rise in a low-resistance circuit may be very high at resonance.

Power into resonant circuit. No power is dissipated in heat in a pure inductance or capacitance, but energy stored at one instant in a magnetic or electrostatic field is turned back into the circuit at another instant. Power is expended only in the resistance of a circuit. This power is equal, as usual, to

$$P = I^2 \times R$$

where R is the resistance of the circuit. In Fig. 1–9, where the resistance is 10 ohms and the current at resonance 1 amp, the power is 10 watts. Thus the power fed into the circuit by the generator is the product of the current and the voltage, or 10×1 or 10 watts.

In other words, all the energy taken from the generator is used up in heating the resistance. None is necessary to maintain the magnetic and electrostatic fields of the coil and the condenser. The energy in these fields is transferred from one to the other, the sum at any one instant being equal to the sum at any other instant so long as the energy dissipated in the resistance is supplied from the outside. The only factor limiting the flow of current at resonance is the resistance of the circuit.

In actual circuits the resistance is not isolated as in our demonstration problems. All coils have resistance; so do all condensers, although the resistance of modern variable condensers is quite small. These resistances take power from the generator and reduce the maximum height of the resonance curve.

The resonant frequency of the circuit. The condition for series resonance—that the reactances of the circuit add up to zero—is fulfilled when

$$X_L - X_C = 0$$

or

$$X_L = X_C$$

or

$$\omega L - \frac{1}{\omega C} = 0$$

or

$$\omega L = \frac{1}{\omega C} \qquad \text{or} \qquad \omega^2 = \frac{1}{LC}$$

and, since $\omega = 2\pi f$,

$$(2\pi f)^2 = 4\pi^2 f^2 = \frac{1}{LC}$$

$$f^2 = \frac{1}{4\pi^2 LC}$$

and so we arrive at the familiar expression for the **resonant frequency** of a circuit as

$$f = \frac{1}{2\pi\sqrt{LC}}$$

in which f = the frequency in cycles.
L = the inductance in henries.
C = the capacitance in farads.
π is the Greek letter pi and is equal to $3.1416\cdots$.

Example 2–9. To what frequency will a circuit tune which has an inductance of 0.25 henry and a capacitance of 0.001 μf?

Let us write the above formula as

$$f^2 = \frac{1}{4\pi^2 LC} = \frac{1}{39.5LC}$$

$$= \frac{1}{39.5 \times 0.25 \times 0.001 \times 10^{-6}}$$

$$= \frac{10^9}{39.5 \times 0.25}$$

$$= \frac{10^9}{9.87} = \frac{10^3}{9.87} \times 10^6$$

$$f = \sqrt{101} \times \sqrt{10^6}$$

$$= 10.1 \times 10^3 \text{ cycles}$$

$$= 10.1 \text{ kc}$$

The expression for the resonant frequency of a circuit shows that the frequency depends upon the product of L and C, and not upon either of them alone. If L is doubled, C can be halved and the resonant frequency of the circuit will not be changed.

Wavelength. When higher and higher frequencies are used, as is the trend today, the numbers designating those frequencies become cumbersome and it is more convenient to express the voltages or currents in terms of wavelength. Since the frequency is the number of complete cycles per second and since radio waves travel at a fixed speed, it follows that a complete cycle occupies a given amount of space. Figure 6–9 pictures one of the cycles. The distance apart, in space, of two corresponding parts of two waves, say the two positive or negative crests, or the points where the two waves cross the zero axis in a given direction, constitutes the wavelength. The wavelength can be computed by remembering that it is equal to the speed at

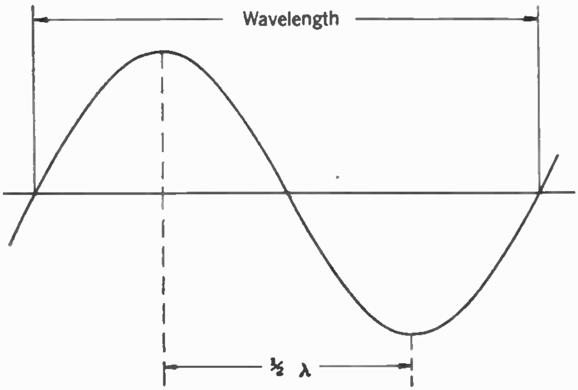


FIG. 6-9. A wavelength is the distance between any two corresponding portions of two waves. Thus, the wavelength may be the distance between two positive (or two negative) peaks, or between the points shown above.

which electric waves travel (300×10^6 meters per second) divided by the number of waves per second (frequency). Thus

$$\text{Wavelength} = \frac{\text{Velocity of radio waves}}{\text{Frequency}}$$

$$\text{Wavelength in meters} = \frac{300 \times 10^6}{f \text{ in cycles}} = \frac{300 \times 10^3}{f \text{ in kilocycles}}$$

$$= 300 \times 10^6 \times 2\pi\sqrt{LC}$$

The customary symbol for wavelength is the Greek letter lambda, λ, and so the above expression may be written

$$\lambda = \frac{300 \times 10^3}{\text{kilocycles}} = 1.884\sqrt{LC}$$

where L is in microhenries and C is in micro-microfarads.

Example 3–9. What wavelength corresponds to 1000 kc?

Solution.
$$\lambda_{\text{meters}} = \frac{300 \times 10^3}{1000} = 300$$

Example 4–9. What frequency corresponds to a wavelength of 100 cm (1 meter)?

Solution.
$$\lambda_{\text{meters}} = \frac{300 \times 10^6}{f_{\text{cycles}}}$$
$$f_{\text{cycles}} = \frac{300 \times 10^6}{\lambda_{\text{meters}}}$$
$$= \frac{300 \times 10^6}{1} = 300 \times 10^6$$

Since 10^6 cycles = 1 megacycle, 1 meter wavelength = 300 megacycles.

Problem 1–9. What inductance must be placed in series with a 2-μf condenser to resonate at 60 cycles? If the voltage across the combination is 110 (effective) and the resistances in the coil and condenser add up to 200 ohms, what power is consumed in the circuit at resonance, what is the resonant current, and what voltage then appears across condenser and inductance?

Problem 2–9. A coil of 0.15 henry is in series with a condenser of 28.5 μf and a resistance of 5.8 ohms. The voltage across the circuit is 22 volts; the frequency is the resonant frequency of 77 cycles. What would the voltage across the condenser be if the resistance were doubled? What power would then be wasted in heat at resonance?

Problem 3–9. A variable condenser has a range from maximum to minimum capacitance of 9 to 1, that is, from 0.0005 to 0.0000555 μf. What frequency range will it cover with a given coil; that is, what is the ratio between the highest and lowest frequency to which it will tune the coil?

Problem 4–9. What happens to the voltage across a condenser of a series-resonant circuit if the capacitance is reduced to half, resonance being maintained by other means which also keep the original current?

Problem 5–9. An antenna may be represented by an inductance of 50 μh in series with 0.00025-μf capacitance and 30 ohms resistance. What is its resonant frequency? If a distant station transmitting on this frequency produces a voltage of 1000 μv across the ends of this antenna system, what current will flow?

Problem 6–9. What can be done to increase the current at 2000 kc in the antenna of Problem 5–9? At resonance, what voltage will appear across the 50-μh inductance if R = 30 ohms? What power is being lost in this antenna at resonance?

Problem 7–9. The primary of an audio-frequency transformer has 100 henries inductance. In many circuits a condenser is placed across the primary so that high radio frequencies will not have to pass through the transformer. If this condenser has a capacitance of 0.001 μf, what is the decrease in effective impedance of the circuit to a frequency of 10,000 cycles?

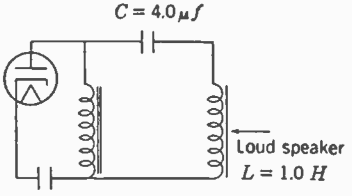


FIG. 7–9. An example of a resonant circuit.

Problem 8–9. A loud speaker may be coupled to a power tube through a condenser as in Fig. 7–9. If the speaker has an inductance of 1 henry and the condenser is a 4-μf unit, to what frequency will the combination become resonant?

Problem 9–9. Plot a curve of the reactance of the loud speaker in Problem 8–9 from 100 to 10,000 cycles.

Problem 10–9. In a 7100-kc transmitting station a 100-$\mu\mu$f condenser is in series with the antenna. What voltage has this condenser across it, if the antenna current is 1 amp?

Problem 11–9. In another transmitter an amplifier is used to raise the voltage coming from a 7500-kc oscillator. In the input circuit of the amplifier is an inductance of 4.5 μh; 80 volts is to appear across it. How much current must flow through this inductance? What capacitance must be across it if it is to tune to 7500 kc?

Speed of radio transmission. Radio signals, like light, travels at the rate of 186,000 miles per second. Although we often think of this as instantaneous transmission, it is a perfectly definite velocity and is not infinite. When signals are transmitted through wires, the speed is less than that in free space.

Important applications have been made of the fact that radio waves travel at known speeds. For example, a radio signal sent out from a transmitter may be reflected back to the region of the transmitter by some electrical conductor, such as a metallic airplane. The time required for the signal to go out to the plane and to be returned back to the transmitter vicinity is a measure of the distance of the airplane from the transmitter.

The signals travel 186,000 miles per second, or 0.186 mile per microsecond. This is equal to $0.186 \times 1760 = 327$ yards per μsec. Since it takes as long to go out as to return, if the time interval between sending out the signal and receiving it back is 10 μsec, the reflecting object is 1635 yards distant.

Parallel resonance. Many of the circuits used in radio involve resonance in a branched or parallel circuit. Figure 8–9 shows a typical parallel circuit composed of an inductance shunted by a condenser, the combination forming what is sometimes called an **anti-resonant** circuit. The effects of varying the frequency of the voltage across the circuit are widely different from the effects in a series circuit. In a series circuit, the currents become very large at resonance and the resultant series impedance of the circuit becomes small. In the parallel connection the circuit at resonance offers a large impedance and the current from the generator becomes very small. In the series connection the same current flows through the condenser and the coil. The voltages across these units differ. In the parallel connection the same voltage is across each branch, but the currents through them differ.

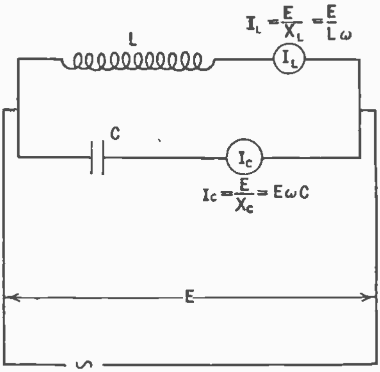


Fig. 8–9. A circuit in which parallel resonance occurs. A parallel circuit like this is sometimes called an anti-resonant circuit.

Experiment 3–9. Connect as in Fig. 8–9 the coil and condenser used in Experiment 1–9. If sufficient meters are available read the alternating current in the two branches as well as the current from the generator as the frequency of the generator is changed. Then fix the generator frequency and adjust the condenser capacity until maximum resonance occurs. Plot the currents against frequency and against condenser capacity. The generator in this experiment may be a small oscillating tube. A 5-watt output is sufficient to produce currents in the branches of the circuit of 100 ma which can be read with a Weston Model 425 thermogalvanometer.

If laboratory apparatus is not available, the current may be calculated after L, C, and E values have been chosen.

The voltage across the branches is equal to the voltage across the circuit as a whole. The current taken by each branch is the ratio between the voltage and the reactance (or impedance if a combination of X and R exists) of that branch. Thus,

$$I_L = \frac{E}{X_L} = \frac{E}{\omega L}$$

$$I_C = E/X_C = E\omega C$$

$$I = I_L - I_C = E\left(\frac{1}{\omega L} - \omega C\right) = E\left(\frac{1 - C\omega^2 L}{\omega L}\right)$$

As the frequency is increased, more and more current is taken by the capacitive branch, less and less by the inductive branch. In a series circuit the *voltages* across coil and condenser are out of phase; their algebraic sum at any frequency combined vectorially with the IR drop is the voltage across the combination. In a parallel circuit the *currents* are out of phase; at any frequency their algebraic sum combined vectorially with the shunt resistance current (if any) gives the current taken from the generator. In the simple example discussed here, where the resistance is neglected, the algebraic sum of the currents gives the generator current. Since these two currents are out of phase (the capacity current has a negative sign), adding them algebraically actually means subtracting I_C from I_L.

At resonance the currents taken by the two branches are equal and if there is no resistance in the circuit the current taken from the generator is zero, because it is the difference of the two branch currents which is read in the generator circuit ammeter.

The impedance of the circuit as a whole, that is, the impedance into which the generator must feed current, is the ratio between the voltage and current:

$$Z = E \div I$$

Therefore, if no current flows, the circuit has infinite impedance. Actually there is always some resistance in the circuit. This may be in an additional shunt path, or resistance may exist in one or both of the other branches. Actually, then, the generator current does not fall to zero but passes through a minimum value. In most radio circuits by far the greater part of the resistance which is in the circuit resides in the coil since the resistance of the average well-constructed condenser used at radio frequencies is very small.

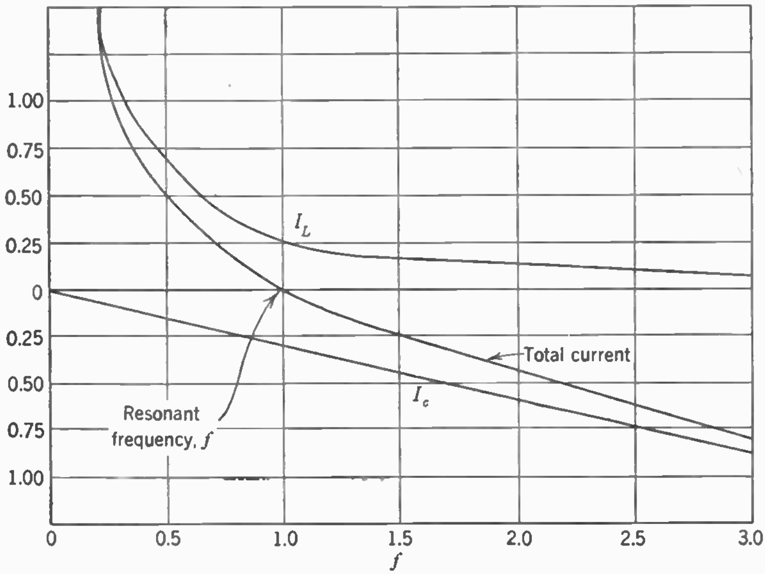


FIG. 9–9. In studying parallel circuits, it is simpler to plot the current taken by the reactive branches than to plot the individual reactances as is done in series circuits, for example in Fig. 4–9.

Resonance conditions. The three possible resonance conditions in a parallel-resonant circuit are:

1. Inductive and capacitive reactances are made equal.

2. The circuit is tuned for minimum current taken from the line.

3. The parallel circuit acts like a pure resistance.

The three frequencies at which these distinct conditions may exist are but slightly different provided that the resistance in the circuit is small, as it almost invariably is in practical radio systems.

At any rate, at resonance the anti-resonant circuit acts approximately like a pure resistance. The current through the inductive branch equals

$$I_L = \frac{ER}{R^2 + \omega^2 L^2}$$

and the impedance represented by the circuit is

$$Z = \frac{E}{I} = \frac{R^2 + \omega^2 L^2}{R}$$

where E is the impressed voltage.
R is the total resistance of the circuit.

Since R^2 is small compared to $\omega^2 L^2$

$$Z = \frac{\omega^2 L^2}{R} = QX_L = Q^2 R_L = \frac{L}{RC}$$

where R_L is the coil resistance.

The current taken from the external source by the tuned circuit is, at resonance

$$I = \frac{E}{Z} = \frac{E}{\omega^2 L^2 / R} = \frac{ER}{\omega^2 L^2} = \frac{E}{\omega L Q}$$

Thus it can be seen that the line current may be made very small by tuning the circuit to resonance and by keeping the circuit resistance very small. The current is actually the impressed voltage divided by Q times the inductive (or capacitive) reactance.

Values of Q which are easily attained are of the order of 100 or more, and so the line current may be $1/100$ of the impressed voltage. This means, simply, that the parallel tuned circuit represents a high impedance when looked at from the outside. The loop around the circuit, however, has low impedance (acting like a series-tuned circuit), and heavy currents may flow within the circuit. Off resonance the impedance may be small, and so such a circuit may be used to by-pass frequencies higher or lower than the resonant frequency without much loss of the resonant frequency.

At resonance the current circulating in the circuit (often called a "tank" circuit) is Q times the line current.

The resonance frequency is the same as for a series circuit, provided that the total circuit resistance is small.

$$f_r = \frac{1}{2\pi\sqrt{LC}}$$

For example, in the circuit in Fig. 8–9, let

$$L = 200\ \mu\text{h}$$

$$\omega = 2\pi \times 356{,}000$$

$$R = 10 \text{ ohms (in the coil)}$$

$$X_L = \omega L = 200 \times 10^{-6} \times 2\pi \times 356{,}000$$

$$= 447 \text{ ohms}$$

Here we may neglect the effect of resistance and use the simple relation for resonant current and for impedance.

If

$$E = 10$$

$$I = \frac{ER}{\omega^2 L^2} = \frac{ER}{(X_L)^2}$$

$$= \frac{10 \times 10}{447^2}$$

$$= 0.5 \times 10^{-3} \text{ amp} = \tfrac{1}{2} \text{ ma}$$

This is the line current. The circulating current will be Q times this value or $0.5 \times \omega L/R = 0.5 \times (447 \div 10) = 0.5 \times 44.7 = 22.35$ ma.

Effective resistance. The thing to remember about the impedance of an anti-resonant tank circuit at its resonant frequency is that it is practically a pure resistance as far as the generator which feeds it power is concerned. The power taken from the generator, therefore, is equal to the square of the current from the generator times the impedance of the circuit or

$$P = I^2\left(\frac{L}{CR}\right) = I^2Z$$

where P = power from generator.
I = current taken from generator.
L = inductance of the tank.
C = capacitance of the tank.
R = resistance of tank circuit.

Now if the inductance of the tank is coupled to a load, for example to an antenna, then the load will draw some power from the tank. This power must be supplied by the generator feeding power to the tuned circuit since the tank by itself supplies no power—it merely acts as a transformer to couple the generator to the load. When the antenna takes power, the tank acts as though its resistance had been increased. For example, suppose that 5 watts is taken by the tank from the generator and that 1 amp flows to the tank circuit. Now couple the coil to an antenna which may take 5 watts. A total of 10 watts now flows from the generator, and so far as it is concerned it seems as though the resistance of the tank circuit had been doubled, since twice the power is now taken from it and the line current has been maintained constant.

The value of R, tank circuit resistance, therefore must represent not only the series ohmic resistance of the coil, the leads, and the condenser, but also the resistance "reflected" into the tank by the load which it feeds.

Tuned circuit applications. Series- and parallel-resonance circuits are very widely used in radio apparatus. Whenever it is desired to secure a large current and a low-impedance circuit, series resonance is utilized. When it is desired to build up a high impedance or a high-voltage circuit an anti-resonant circuit is used. Let us consider the antenna-ground system in Fig. 10–9. The antenna has in series with it an inductance across which a voltage is to be developed at a desired frequency. In series with this inductance are a capacitance for tuning purposes and an anti-resonant circuit. Voltages of various frequencies, among them the desired frequency, are impressed on the antenna by distant transmitting stations. The maximum voltage is desired across the inductance, L, at the desired frequency and the minimum at other frequencies. If there is an especially strong signal which is setting up a voltage across the antenna, then the anti-resonant circuit can be tuned to this interfering frequency when it will act as a rejector circuit.

The condenser, C, is adjusted until the antenna system as a whole is resonant to the desired signal. A large current flows through the series system, building up a large voltage across L. Voltages of other frequencies cause small currents to flow in the antenna system, and consequently small voltages at these frequencies are built up across the coupling coil, L. The anti-resonant circuit being tuned to the unwanted signal makes the antenna system as a whole have a very high impedance at this frequency and so very small currents will flow through it, building up small voltages of this frequency across the coupling coil.

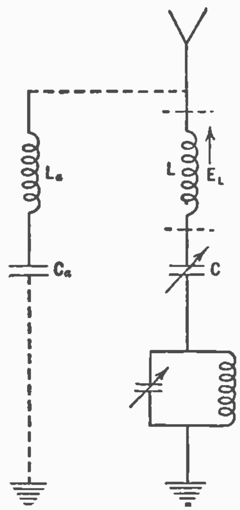


Fig. 10–9. The anti-resonant circuit in series with the antenna rejects undesired signals by making the series impedance to them very high.

Such an anti-resonant circuit is often called a **rejector** circuit because it rejects signals of the frequency to which it is tuned. The series-resonant circuit is called an **acceptor** because it accepts signals of the resonant frequency. The rejector used in this circuit is commonly known as a wave trap because it traps out unwanted signals.

Let us suppose that signals are fed into the input of an amplifier which has an internal resistance, R, in series with an output circuit, as shown in Fig. 11–9. The voltage across this output, Z, is to be made as high as possible. The amplifier has available a certain voltage, E, which must be divided between the internal resistance of the amplifier and the output load. The proportion of the voltage that appears across this load increases as its ohmic impedance increases with respect to the amplifier's resistance. Thus, if the output impedance is equal to the internal resistance of the amplifier, one-half of the total voltage available will appear across it. If it is higher than this value, a greater proportion of voltage will be usefully applied across the load and less used up in the resistance.

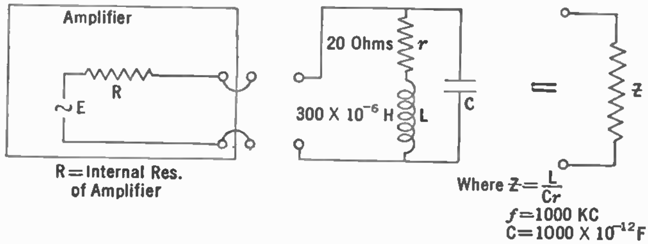


FIG. 11–9. A high impedance is desired for the amplifier to work into. A tuned circuit does the trick. Numerically it is equal to Z.

Here the anti-resonant circuit is used. At resonance its impedance becomes equal to $\dfrac{L}{Cr}$ or

$$\frac{\omega^2 L^2}{r} = \frac{(6.28 \times 1000 \times 1000)^2 \times (300 \times 10^{-6})^2}{20}$$
$$= 180{,}000 \text{ ohms}$$

If the amplifier's internal resistance is equal to 20,000 ohms, the voltage across the tuned circuit is 180/200 or 9/10 of the total available voltage. This is because the total voltage is divided between two resistances which add up to 200,000 ohms. The voltage across the 180,000-ohm tuned circuit is the useful voltage, which is equal to 180,000/200,000 of the total voltage.

Problem 12–9. A screen-grid tube gives the greatest voltage amplification when worked into a very high impedance. A condenser of 1500 $\mu\mu$f is available. Calculating the size of the inductance required to tune to 30 kc and assuming it to have a resistance of 30 ohms, what is the impedance ($\omega^2 L^2/R$) that can be presented to the tube by shunting the coil and condenser and using the anti-resonant circuit as a load?

Problem 13–9. A wave trap is to be put into an antenna and tuned to a station whose frequency is 750 kc. Convenient sizes of coil and condenser to use are 100 μh and 450 $\mu\mu$f. They are to be shunted across each other and the combination put in series with the antenna. If the coil has a resistance of 10 ohms and the condenser a resistance of 1.0 ohm at this frequency, what impedance will the trap offer to the offending signal?

Problem 14–9. In Problem 13–9, neglecting phase differences between the trap and the rest of the antenna, if the total impedance of the antenna to the offending signal is double that of the trap alone so that one-half of the total antenna voltage is across the trap, what current will flow through the condenser if the total 750-kc voltage across the system is 10 μv?

Sharpness of resonance. The effect of resistance is to reduce the maximum current flowing in a series-resonant circuit, to make less pronounced the minimum of current flowing into a parallel-resonant circuit from an external source, and to decrease the impedance (L/CR) of the parallel circuit.

Since the maximum current is desired in a series circuit, and the maximum impedance in a parallel one, the inclusion of resistance in either is deleterious.

Let us consider the antenna illustrated in Fig. 10–9. Suppose that its inductance, L, is 200 μh, and C at resonance (356 kc) is 1000 $\mu\mu$f. For the moment we shall neglect the presence of the wave trap. Assume a voltage of 10 volts. What is the effect on the resonance curve of this antenna system if it has a resistance of 10 ohms or of 40 ohms? The current at resonance in the 10-ohm case is 1 amp whereas at 370 kc the current is 0.215 amp, a ratio of 46.5. In the 40-ohm case, the resonant current would be only 0.25 amp—one-fourth of its value with the lower resistance—and the current at 370 kc, i.e., 14 kc off resonance, would be 0.0209 amp. This is a current ratio between the resonant and the off-resonant current of only 12.

In other words, if the antenna had impressed on it from equally distant and equally powerful radio stations two voltages —one of 356 kc, the desired frequency, and one of 370 kc, the unwanted frequency—46.5 times as much current flows at the desired frequency as the unwanted. In the 40-ohm antenna, however, not only is the desired current cut to one-quarter of its other value but the ratio of wanted to unwanted current has been decreased to 12. The low-resistance antenna is said to be more selective, and its selectivity is decreased when resistance is added to it. The sharpness of resonance, or the selectivity, is proportional to the Q of the circuit. The circuit increases in

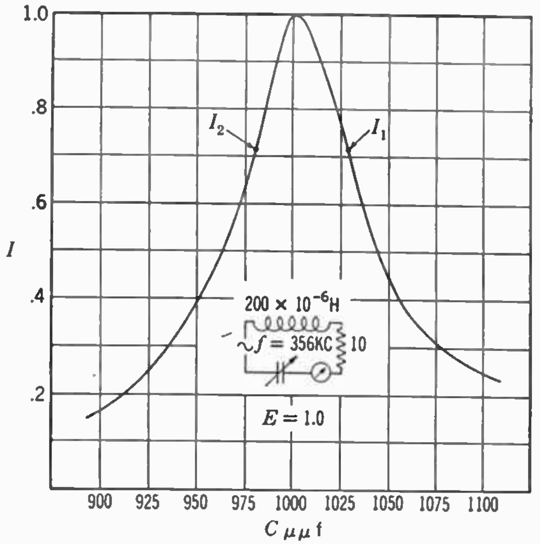


FIG. 12–9. If I_1 and I_2 are equal to 70 per cent of the maximum value attained at resonance, the resistance of the circuit may be calculated.

selectivity as Q increases; i.e., the ratio of the response at resonance to the response to some frequency off resonance is greater. Low-resistance, high-Q circuits are highly selective.

To measure two circuits comparatively it is convenient to have a fixed frequency at which the selectivity is determined compared to the resonant frequency. This is taken usually as the frequency at which the reactance in the circuit is equal to the resistance of the circuit.

Suppose that, as in Fig. 12–9, we plot a resonance curve of current against capacity. Suppose the capacity to be adjusted until the total reactance in the circuit $(X_L - X_C)$ is equal to the resistance in the circuit, that is,

$$(X_L - X_C) = R$$

when

$$I = \frac{E}{\sqrt{R^2 + X^2}}$$

$$= \frac{E}{\sqrt{2R^2}} = 0.707 I_r$$

where

$$I_r = \text{current at resonance}$$

and if

$$C_2 C_1 = C_r^2$$

then

$$\frac{\omega L}{R} = \frac{2C_r}{C_2 - C_1} = Q$$

in which C_r = the capacitance at resonance.

C_1 and C_2 = the two values of capacitance which make $I = 0.707 I_r$.

Width of resonance curve. If, however, the frequency of the impressed voltage is so adjusted that two currents are reached, above and below the resonant frequency f_r, which are equal to $0.707 I_r$,

$$\frac{\omega L}{R} = \frac{f_r}{f_2 - f_1} = Q$$

whence the width of the frequency band

$$f_2 - f_1 = \frac{R f_r}{\omega L} = \frac{R}{2\pi L}$$

Example 5–9. What will be the width in cycles of the resonance curve at a point where $I = 0.707 I_r$ when $L = 200$ μh, $R = 10$ ohms, $f = 356{,}000$ cycles?

$$f_2 - f_1 = \frac{10 \times 356{,}000}{447} = \frac{R \times f_r}{\omega L}$$

$$= 8000 \text{ cycles}$$

and if

$$\frac{\omega L}{R} = \frac{2C_r}{C_2 - C_1}$$

$$C_2 - C_1 = \frac{2C_r \times R}{\omega L}$$

$$= \frac{2000 \times 10}{447}$$

$$= 44.7\ \mu\mu\text{f}$$

This equals the change in capacitance required to change the current from $I = 0.707I_r$ below resonance to $I = 0.707I_r$ above resonance, or from I_1 to I_2 in Fig. 12–9.

Problem 15–9. In Fig. 12–9, suppose that the resistance is 20 ohms instead of 10. Calculate the width of the band at the point where the current is 0.7 of its maximum value, and the change in capacitance required to produce this change in current.

Problem 16–9. A certain coil-condenser combination has a resistance of 16 ohms at 400 meters. The inductance is 170 μh. What is the width of band passed at the point where the current is equal to 0.7 of its maximum value? What is the Q of this circuit?

Problem 17–9. A circuit is to pass only 0.707 of its maximum current at a point 2.0 kc off resonance, which occurs at 500 kc. The condenser to be used has a capacity of 0.0006 μf. Calculate the maximum resistance the circuit can have.

Problem 18–9. Suppose that increasing the size of an inductance by a factor of 2.0 increases the resistance in a circuit by a factor of 1.5. The circuit is to tune to the same wavelength. What has happened to the selectivity of the circuit?

Problem 19–9. If the expression $\omega L/R$ of a coil remains constant over a fairly wide band of frequencies, does the selectivity of a tuned circuit differ at different frequencies? Does the width of band passed differ at 1500 kc from what it is at 500 kc?

Ratio of L/C. Since the selectivity is proportional to Q which is equal to the inductive or capacitive reactance divided by the circuit resistance, it is apparent that increasing L and decreasing C so that the circuit is still tuned properly will increase the selectivity.

$$Q = \frac{X_L}{R} = \frac{X_C}{R} = \frac{\omega L}{R} = \frac{1}{\omega CR} = \frac{1}{R}\sqrt{\frac{L}{C}}$$

If, however, the coil resistance increases as fast as the inductance increases, or faster, then the benefit of using high L and low C will not be secured.

Coil resistance. An inductor for radio frequencies may have a d-c resistance of only a fraction of an ohm. Yet that coil placed in a tuned circuit in series with a condenser having still lower resistance will exhibit an effect as though the resistance were many times the d-c resistance of the wire.

A wire stretched out straight will have one value of resistance to direct currents and another value of resistance to high-frequency currents. Therefore the increase in apparent resistance is not due to the fact that the copper wire is wound up in the shape of a coil. The fact is that the resistance is higher to higher-frequency currents than it is to direct currents. Because of the rapid change in direction of flow and because the current within the cross section of the conductor changes rapidly, small countervoltages are induced within the wire itself, according to Lenz's law. Because magnetic lines originate at the center of the wire and expand outward, being more concentrated at the center than near the surface of the conductor, the induced voltage would cause a decided opposition to the flow of current at the center. The result is a decrease in effective conductor area, and a rise in resistance. This effect is known as **skin effect,** and it explains why the resistance of an inductance increases as the frequency increases. Over the normal tuning range in which the coil is to be used, the resistance does not vary greatly (perhaps 2 to 1), but the effect is important, nevertheless. As the frequency is increased, the current tends to flow more on the surface (or "skin") of the conductor and less in the interior.

Distributed capacitance of coils. Whenever two objects which conduct current are insulated from each other and are at different voltages, electrostatic charges can be stored on them; they constitute a condenser. The capacitance of this condenser depends upon the proximity of the objects, the insulation between them (the dielectric constant), and their shape. In a coil of wire, each turn is insulated from the adjacent turns, and a difference of potential exists between two turns. Therefore a coil is not a pure inductance but may be thought of as a coil shunted by a capacitance made up of the sum of many small turn-to-turn capacitances. At some frequency the coil shunted by the resultant capacitance becomes as shown in Fig. 13–9. Here the tuning condenser is no longer in series with a coil but is in series with an anti-resonant circuit made up of parallel inductance and capacitance. At the resonant frequency of this parallel combination little current will flow from the generator through the series circuit because, at this frequency, the anti-resonant circuit has very high impedance.

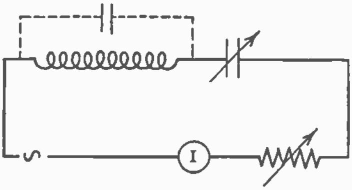


Fig. 13–9. When the dotted capacitance across the coil tunes it to the frequency of the generator, the series impedance of the circuit becomes very high because it then acts as an anti-resonant circuit.

Near the resonant frequency of the coil, the total series impedance of the entire circuit becomes very high. This inherent coil capacitance is known as a **distributed capacitance** since it is not concentrated in any one place or form but is more or less evenly distributed along the whole length of the inductor.

If one measures the resistance of a coil at higher and higher frequencies, he will find that, near the resonant frequency, the apparent resistance rises very rapidly. What is happening is the phenomenon described above. Actually it is the parallel impedance of the inductor shunted by its own distributed capacitance that is being measured.

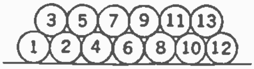


Fig. 14–9.– Bank winding type of coil.

Various attempts have been made to calculate the capacitance of coils. It has been found by experiment that the radius and shape of the coil control the distributed capacitance to a large extent. Thus a coil of average proportions, i.e., the length about equal to the diameter, has a capacitance of approximately $0.6R$ $\mu\mu$f, where R is the radius of the coil in centimeters. As a rough rule it has been stated that the capacitance in micro-microfarads is always less than the radius in centimeters in a solenoid of a single layer. Another experiment showed that the natural wavelength of solenoids was about 2.54 times the total length of wire on the coil.

To obtain large inductance in small space and less internal capacitance, it has become customary to make multilayer coils of peculiar types of winding called bank winding and universal winding.

Air cores have a permeability of unity; if coils could be wound on a core of higher permeability, the inductance per length of wire and per unit of space would increase. Permalloy is a kind of iron alloy dust on which coils may be wound with increased inductance at the medium frequencies. At higher frequencies one may use cores made up of finely powdered iron bound together with some sort of very fine binder. The permeability of this material may be as high as 12. With a given length of wire, therefore, of a given high-frequency resistance a greater inductance may be wound. The Q of such a coil will be higher because the ratio of inductance to resistance will be higher.

Q's of the following values are typical: 456 kc litz wound universal coil, 80; same coil with powdered iron core, 145; transmitter coil for 5000 kc, 650; gang condenser, ceramic insulation, 1000 kc, 3000.

Problem 20–9. A coil with a Q of 100, an inductance of 18 μh, and a condenser of 30 $\mu\mu$f are in series. To what frequency does the circuit resonate, what is the resonant impedance, and what is the resistance of the coil? How wide is the resonance curve at the points where the current is 0.707 of the maximum current?

Problem 21–9. At 1 megacycle, by how much must the frequency be changed to lower the resonant current to 0.707 of its value if the circuit Q is 200?

Comparison of series and parallel tuned circuits. In the series circuit the line current is high, the circuit impedance is low, the voltages across coil and condenser are high, current through L and C is high, and the actual impedance at resonance is equal to R, the resistance of the circuit.

In the parallel circuit the line current is low, the circuit impedance is high, the currents through coil and condenser are high, voltage across coil and condenser is equal to the impressed line voltage, and the actual impedance of the circuit is equal to ω^2L^2/R or L/CR.

Far from resonance there is a large phase difference between the voltage across the coil, or condenser, and the line voltage, but as resonance is approached these voltages come more and more into phase with the impressed voltage.

	Current				
	Line	Circuit	Impedance	E_X	Z
Series	High	High	Low	High	R
Parallel	Low	High	High	Line	$\frac{\omega^2L^2}{R}$

Transients. In many radio and radar circuits, advantage is taken of the fact that the final value of current (or voltage) in the circuit is not attained instantaneously after voltage (or current) is applied. Proper combinations of R and C or R and L can delay the attainment of the final value of current (or voltage) for an appreciable time. In R-C circuits, for example, current flows for a short time only after d-c voltage is applied. In R-L circuits, current flows in a coil coupled to the inductance for only a brief time after d-c voltage is applied.

Let us consider the simple case of R and C in series with a battery and a switch. One cathode-ray tube voltmeter which acts very quickly can be placed across R and another across C. What happens when the switch in Fig. 15–9 is closed?

Since there is no charge in C there is no voltage across it. At the instant of closing the switch, the full battery voltage is impressed across R. Current flows through it. This current must also flow into and charge the condenser. The charge builds up a voltage across the condenser which opposes the battery voltage. As soon as the condenser voltage is equal to the battery voltage, current flow ceases, and all the voltage drop in the circuit appears across the condenser; none across the resistance. The voltage across the resistance falls off at the same rate as the condenser voltage rises.

The conditions during the time in which this momentary current flows are called **transient.** After current flow ceases, the conditions are called the **steady-state** conditions.

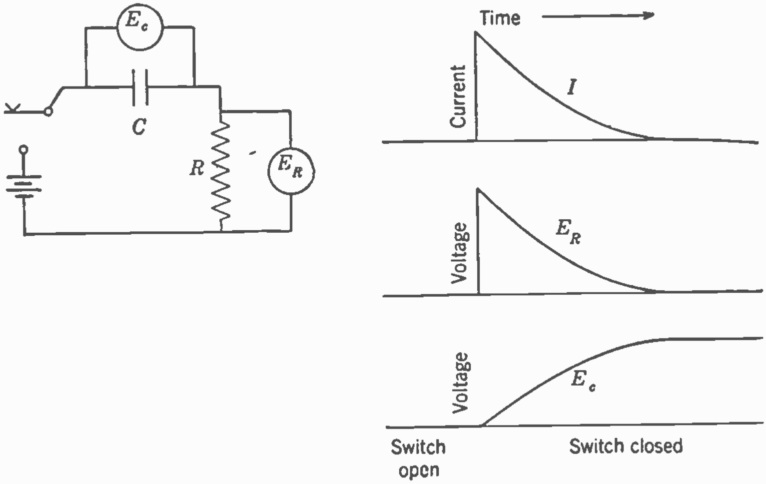


FIG. 15-9. When the key is closed, current rushes into the capacitor, building up a voltage there which opposes the battery voltage and ultimately equals the battery voltage. Then current around the circuit ceases to flow, the condenser is charged, and, since there is no current, there is no voltage across R. This chain of events may occur very quickly (in a microsecond or less) or may require a much longer time if C and R are large.

When the switch in Fig. 16-9 is thrown so that the condenser is discharged through the resistance, what are the transient conditions? At the instant the switch is thrown, there is but one voltage in the series circuit, that existing across the condenser. Current begins to flow and to build up a voltage drop across the resistance. This voltage drop must be opposed to the condenser voltage according to Kirchhoff's law. The current flowing through the resistance produces a power loss in this resistance, due to heat, of I^2R. Here the energy that existed in the condenser in the form of potential energy is dissipated and soon current flow ceases. Maximum current flows at the instant the switch is closed; maximum voltages exist across condenser and resistance at this instant. The sum of the two voltages must equal zero at all times; that is, $E_C - E_R = 0$.

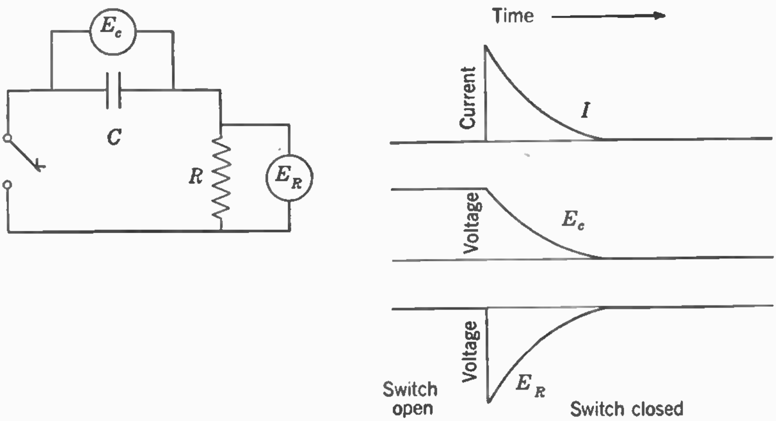


FIG. 16–9. In this figure the condenser is charged. When the key is closed, current flows through the circuit, building up a voltage across R which opposes E_c. When the charge in C has been dissipated, current ceases and E_c and E_R both become zero.

Conditions in *R-L* circuits. When voltage is impressed on a circuit made up of resistance and inductance, the voltage across the inductance rises to the full battery voltage instantly, but the current through the circuit builds up slowly. Therefore the voltage drop across the resistance builds up slowly (at the same rate as that at which the current increases). The voltage across the inductance decreases at the same rate as that at which the current increases since there is induced voltage across an inductance only when the current through the inductance changes and it is proportional to the rate at which the current changes.

If the *R-L* circuit is shorted, the impressed voltage is reduced instantly to zero; the current decreases at a rate depending upon R and L; the voltage across the resistance decreases at the same rate; and the instantaneous voltage across the inductor, which has a polarity opposite that of the original impressed voltage, decreases at the same rate as the current decreases.

Square waves. Now that we see that currents and voltages in *R*, *L*, and *C* circuits do not build up instantly, we can understand why the voltages across the individual components of the circuits differ from one another when a square wave is impressed upon the circuits. The exact wave forms of the voltages will depend upon the length of time that elapses between charge and discharge in an *R-C* circuit or between the instants of impressing the battery voltage on and shorting the *R-L* circuit. If the flat-topped portion of the square wave is long compared to the time

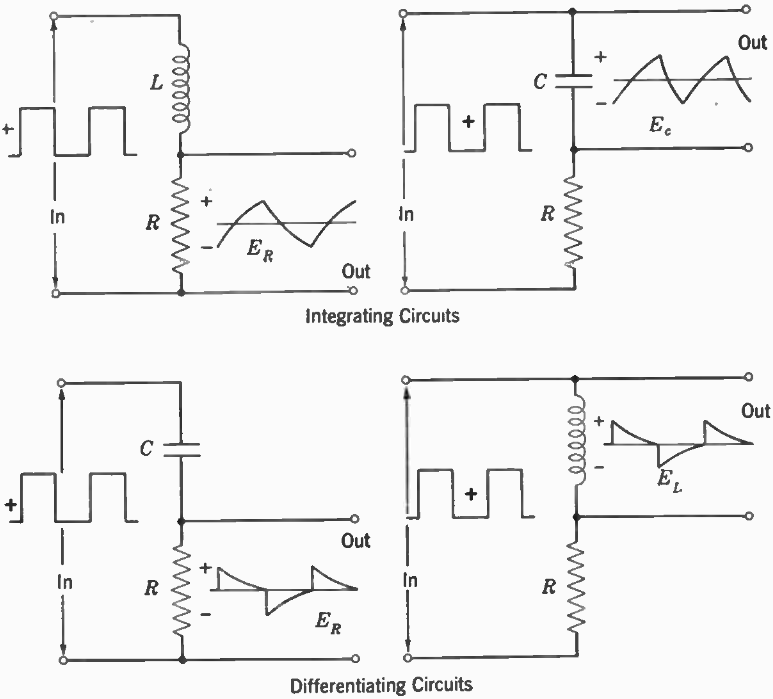


FIG. 17–9. Circuits in which the transient response is important. The forms of E_R, E_c, and E_L vary according to the time constant of the circuit compared to the dimensions of the square wave input.

constant of the R-C or R-L circuit, the wave forms shown in Figs. 15–9 and 16–9 will exist; but if the square wave has a narrow flat top, then only a portion of the full transient wave form will be realized.

For example, Fig. 17–9 shows the effect of impressing a pulse on R-L and R-C circuits. These circuits and the corresponding wave forms are employed in television, in welding control, and in radar circuits. They are known as **differentiating** and **integrating** circuits as indicated. Note how the square wave is changed into a peaked wave (in integrating circuits) or a saw-tooth wave (in differentiating circuits). This is one way of producing wave forms of these two kinds. If sine waves are impressed on these circuits, sine waves of voltage will appear across the individual components, the only change occurring being a shift in phase between the voltage across the resistance compared to that across the capacitor or the inductor.

CHAPTER 10

PROPERTIES OF COILS AND CONDENSERS

Coils and condensers form the nucleus of every radio circuit. To understand their role in the reception of radio messages, let us look at a simple receiving system.

Tuning a receiver. A simple receiving circuit consists of an antenna-ground system connected to a coil and a detector, such as a crystal of galena or silicon or other sensitive mineral which has the property of separating the audio tones from a radio wave. A pair of headphones may be put in series with the detector so that the audio tones which are filtered out of the radio wave by the detector may be made audible. A small condenser across the phones will pass the radio frequencies but not the audio frequencies, which must go through the phones.

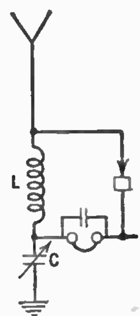


FIG. 1–10. A simple radio receiver.

One way to get louder signals is to tune the antenna-ground system to the frequency of the desired wave. This is done by varying C in Fig. 1–10. When the circuit is series resonant, a large current flows through the inductance. The voltage across it ($X_L \times I$) will be large, and the response from the crystal will be greater.

The voltage across the inductance can be amplified and then impressed across the detector. This amplification may take place in several stages so that very weak signals may finally be heard with the strength of nearby strong signals which are detected directly from the antenna inductance. If desired, the signals may be amplified again after detection by means of audio-frequency amplifiers.

As we have already seen, there is another advantage of tuning the antenna, the advantage of selectivity. Signals of low frequency find considerable impedance in the condenser of such a series-tuned antenna; signals of high frequency find impedance in the coil; signals of the desired or resonant frequency find a minimum of impedance, and so the filtering action of the tuned system is advantageous. If, in addition to the series-tuned circuit, we use an anti-resonant or parallel-tuned circuit as in Fig. 2–10, we impose more hardships upon unwanted signals. In this case when maximum current flows through L_1 maximum current is induced in L_2. If, then, C_2 is tuned so that L_2C_2 forms an anti-resonant circuit, the impedance to the resonant frequency will be very high and any current through it will build up a large voltage across it so that the detector gets a high voltage at the desired frequency and a low one at all other frequencies—and the selectivity of the system as a whole is improved.

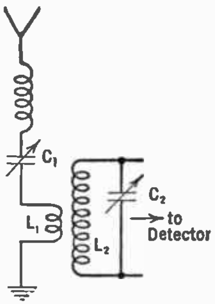


FIG. 2–10. Varying C_1 until the antenna system as a whole is series resonant increases voltage across L_1.

If, in addition, each radio-frequency amplifier stage is tuned to the desired signal, the selectivity of the entire receiver may become very great. In the present congestion of radio stations, the necessity for selectivity of a high degree is evident.

The frequency meter. An instrument for measuring the frequency of signals is called **a frequency meter.** It consists of a coil and a condenser and some means of indicating when this simple circuit is tuned to resonance with a radio wave. The indicator may be a current meter, a lamp which lights up at maximum current through it, or a crystal detector and a d-c milliammeter. It may be connected directly into the circuit or, preferably, coupled loosely to it.

The circuit of a simple and effective frequency meter is shown in Fig. 3–10. The indicating device is a crystal detector and a meter which indicates the rectified direct current. If the indicator is coupled loosely to the tuned circuit its resistance will not broaden the response curve of the frequency meter. The inductance is usually fixed and the capacitance varied to obtain resonance, but to cover a wide band of frequencies it is frequently necessary to have several coils which fit into the instrument by means of plugs and jacks. If the coils are arranged so that the larger coils have exactly four times the inductance of the next smaller the wavelengths to which the larger coil will tune will be twice those of the next smaller coil or the frequencies will be one-half those of the next smaller coil.

A series of coils in which the same winding space is used but in which the number of turns in this space is doubled for each next larger coil will approximate these conditions very closely.

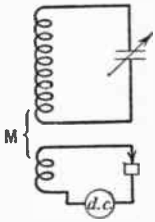


FIG. 3-10. A simple frequency meter.

Sometimes the instrument is equipped with a buzzer so that it will send out a modulated wave. A receiver can be tuned to a desired frequency by starting the buzzer, tuning the meter to the desired frequency, and adjusting the receiver until the buzzer tones are heard at maximum loudness.

Oscillating frequency meter. A most useful type of frequency meter has an oscillating vacuum tube and a meter, usually connected in the grid circuit of the tube. The circuit diagram for such a meter is shown in Fig. 4–10. Tube 1 generates radio-frequency currents, which are modulated when desired by the low- or audio-frequency generator tube 2. The grid meter gives very sharp indications of resonance, and because the device is a small modulated source of radio-frequency energy it can be used to tune receivers to any desired frequency. It is a much more accurate instrument than the buzzer frequency meter. The data in Table I are those of an oscillating frequency meter. The coils have dimensions as indicated.

Calibrating a frequency meter. A frequency meter, to be most useful, must be properly calibrated. This may be done in several ways. If the meter uses an oscillating tube the process is simple; all one needs is a source of known frequency and a

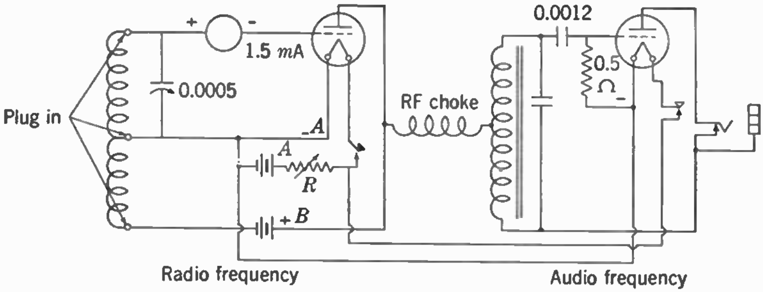


FIG. 4–10. Circuit diagram of a modulated oscillator useful as a wave-meter. The r-f oscillator has a d-c meter in its grid circuit and employs a series of plug-in coils. The modulating oscillator may use an audio output transformer or other center-tapped inductance of the correct value. When the r-f oscillator is coupled to an external tuned circuit, a sharp dip in the grid current will occur when the two circuits are tuned to the same frequency.

TABLE I

Coil	f, kc	λ, meters	Kilocycles per Dial Degree
A	2500–6660	45–120	31.6
B	1430–3750	80–210	23.3
C	750–1820	165–400	10.7
D	485–1130	265–620	6.5

Coil	Turns	Size Wire	Diameter, in.	Length of Winding	L, mh
A	15	21	$2\frac{11}{64}$	$1\frac{5}{8}$	0.014
B	30	21	$2\frac{11}{64}$	$1\frac{1}{2}$	0.055
C	60	21	$2\frac{11}{64}$	$1\frac{5}{8}$	0.217
E	90	27	$2\frac{11}{64}$	$1\frac{3}{8}$	0.495

receiver. Tune the receiver to a station whose frequency is known. Then turn on the oscillating tube, and, when a whistle is heard from the receiver, the known station, the receiver, and the frequency meter are all tuned to the same frequency. Next tune the receiver to another frequency and repeat the performance. A curve can then be plotted showing the calibration of the frequency meter.

The following description of how to calibrate a frequency meter over a wide range of frequencies by means of but a single accurately known frequency is an interesting experiment. It follows from the fact that an oscillating vacuum tube generates not only the frequency governed by the *LC* product of its circuit but also multiples (harmonics) of this frequency.

Experiment 1–10. To calibrate a frequency meter by harmonics. The necessary apparatus consists of:

1. An oscillating frequency meter connected as in Fig. 4–10.
2. An oscillating detector tube preferably followed by a stage of audio amplification.

Tune the oscillating detector to the frequency of some known station by listening in the headphones and bringing an antenna wire near the detector inductance. The tuning condenser of the detector should be equipped with a vernier or worm gear so that very accurate settings are possible. Tune as nearly as possible to zero beat with the known station. As the tuning dial is adjusted near resonance with the known station, now acting as the frequency standard, a note will be heard in the phones which represents the difference in frequency between the known station and that of the detector tube. When this difference tone (or beat note) disappears, the two oscillations are at the same frequency. Since frequencies lower than about 100 cycles cannot be heard in the phones, it will not be possible to tune closer than this to the desired frequency. By estimating the two points at which the audible beat disappears and by finally setting the oscillating receiver detector at the mid-point between these two dial settings, a sufficiently accurate setting will be made.

We have now equipped ourselves with a local generator whose frequency is accurately known. For example, suppose that it is 610 kc and that we are set to within 100 cycles of this frequency. We are within 100 parts in 610,000 of being exactly correct, or 1 part in 6100, which is sufficiently accurate. It is much more accurate than the dial reading of the frequency meter we are to calibrate.

Now move away the antenna coupling and see if the beat note changes. If it does, again adjust for true zero beat. Then start up the oscillating-tube frequency meter and, after giving it a few minutes to warm up, tune its dial slowly until a whistle or beat note is heard in the headphones which are still plugged into the detector-amplifier. This means that the frequency meter is being tuned to the frequency of the oscillating tube.

If we use the broadcast band coil of the frequency meter we ought to get a very loud beat note when the two circuits are in near resonance and another loud note when the dial is tuned to the half wavelength, in this case 1220 kc. Between these points several other weaker beat notes may be heard.

Now turn the dial slowly and note the dial reading each time a beat note is heard. For example, the table of such points may look like Table II, in which the loudest beat notes are marked with an asterisk. Then use another frequency-meter coil and repeat, always marking down the loud notes.

TABLE II

Dial Degrees	Difference	Units Difference	f (approximate)	f (exact)
10.2 *	....	..	1220	1220
34.0	23.8	2	1020	1016
47.0	13.0	1	920	915
60.0	13.0	1	820	813
85.0 *	25.0	2	610	610

* Loudest beat notes.

Now prepare data like those in Table III, in which the numbers along the top are obtained by multiplying the detector frequency by whole numbers from 1 to 10, and the vertical numbers are obtained by dividing this frequency by whole numbers. Thus our fundamental frequency is 610 kc. Twice this is 1220 kc, one-half is 305, etc.

TABLE III

	1	2	3	4	5	6
1	610	1220	1830	2440	3050	3660
2	305	610	915	1220	1525	1830
3	202.5	406	610	813	1016	1220
4	152.5	305	457	610	763	
5						
6						

Then make a list from this table of the frequencies that may be looked for from our calibration, namely, 610, 763, 813, 915 kc, etc.

What actually happens as we tune the frequency-meter dial and hear beat notes? The oscillating detector and the frequency-meter tubes are generating additional or harmonic frequencies as well as the fundamental to which they are set. These additional frequencies are much weaker than the fundamental. When we tune the meter to 1220 kc it beats with the second harmonic of the detector and gives an audible note. But how are we to recognize the 1220 point? How do we know that it is not the third or the fourth harmonic instead of the second?

Consider the data in Table II. We got loud notes at 10.2 and 85 degrees. We guess that these are the second harmonic and fundamental. We subtract the dial settings as in column 2. Then, assuming that 13 degrees is a unit, we note that there are two units between the 10.2- and 34-degree beat notes. We see then that there are six units between 1220 and 610 kc. We guess again and say that each beat note represents about one-sixth of the difference between 1220 and 610 kc, or about 100 kc per unit. Looking in our list of expected frequencies we can pick out these frequencies exactly.

We might guess at these frequencies from the original assumption that the two loud notes were from the 1220 and the 610 kc frequencies and noting that between them, a difference of 610 kc, were 85 − 10.2 dial degrees or about 8 kc per degree.

When the smaller coils are to be used, care must be taken to see that no harmonics are missed. Fortunately, if the coils have the dimensions given in Table I, the harmonics will fall at almost the same points on the dial. Thus on the largest coil 610 kc is found at 85 degrees. On the next smaller coil the 1220-kc frequency will be found within a degree or two of 85 degrees. And so on until the entire set of coils is calibrated.

Standard frequencies. In this country standard frequency signals are sent out from a Bureau of Standards' station, WWV. These signals are on 5000 kc, modulated at 400 cycles, and should be heard all over the United States. In addition there are many long-wave and intermediate-wave stations whose frequencies are kept within very close limits and which are on the air 24 hours of the day. Broadcasting stations form good standards of frequency, covering the band from 550 to 1600 kc, and above this are many short-wave stations whose signals may be heard the world over.

Calibrating by clicks. A common method of calibration is the click method. When a tuned circuit is brought near the inductance of the oscillating frequency meter, a sharp dip of the grid-current needle will be noted as the two circuits are resonated to each other. If the same tuned circuit is brought near the inductance of an oscillating detector tube, a sharp click will be heard in headphones when the circuits are tuned to the same frequency, provided that one listens in the plate circuit of this tube or behind a stage of audio amplification. This click is produced by a sharp change in grid current and a corresponding change in plate current.

Experiment 2–10. Calibration by clicks. This method requires an oscillating detector, a calibrated meter, and the unknown meter to be calibrated.

Couple the calibrated meter to the inductance of the detector, and turn the dial until a sharp click is heard in the phones, indicating that the circuits are tuned alike. If the two inductances are closely coupled two clicks will be heard, one when the tube stops oscillating and one when it starts again. These two points may be several degrees apart. Loosen the coupling, and note that the two clicks approach each other. Keep on loosening it until a degree of coupling is reached when only a single resonance click is noticed. Note the dial setting of the standard meter. Now remove it from the tuned circuit and bring near the latter the frequency meter to be calibrated. Turn its condenser dial until a click is heard as before. Now the meter has the same frequency as the standard. Other points for a calibration curve may be noted in the same manner.

This method really constitutes setting a generator or miniature transmitter (the oscillating detector) to a given frequency by means of the calibrated meter and then tuning the uncalibrated meter to resonance with this generator.

The properties of coils and condensers. We may investigate the properties of coils and condensers by performing the various parts of the following experiment.

Experiment 3–10. Wind up on a form, about 3 in. in diameter, a coil of about 60 turns of rather large wire, preferably with silk or enamel insulation so that the distributed capacitance of the inductance will be rather large. Connect it across a condenser whose maximum capacitance is about 500 $\mu\mu$f. Start at the maximum capacitance of the condenser and measure the resonant frequency of the coil-condenser combination by clicking it into an oscillating receiver, or by coupling it to an oscillator. Then decrease the capacitance and repeat until several readings are taken, say at 500, 400, 300, etc., $\mu\mu$f. Plot the result against C as shown in Fig. 5–10, that is (wavelength)^2 against capacitance.

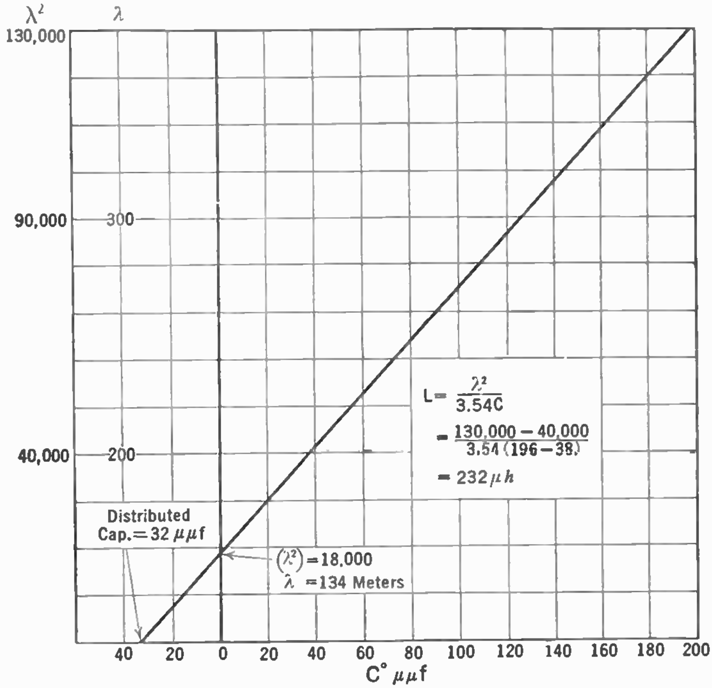


FIG. 5–10. A method of determining the distributed capacity of a coil.

A straight line results because the formula

$$(\text{Wavelength})^2 = 3.54L \times C$$

where L is in microhenries and C in micro-microfarads, is the equation of a straight line and states that the wavelength (in meters) squared is proportional to the capacitance in the circuit. The slope of the line divided by 3.54 is the inductance of the coil; that is,

$$L = \frac{1}{3.54} \times \frac{\lambda^2}{C}$$

It will be noticed that the straight line crosses the wavelength squared axis at some distance above the zero point. This gives us the natural wavelength squared of the coil itself and therefore the resonant wavelength to which the coil with no additional capacity will tune. The point where the line crosses the capacitance axis gives us the distributed capacitance of the coil. This value multiplied by the inductance as obtained above gives the LC product which when fitted into the proper formula gives the natural wavelength of the coil.

Thus, in one experiment we can determine not only the frequency or wavelength to which a coil-condenser combination will tune but also the coil's inductance, its distributed capacitance, and its natural wavelength.

As a check on these data: (*a*) Calculate the inductance from the formulas given in Fig. 4–6. (*b*) Disconnect the condenser from the coil and determine the natural frequency of the coil alone by means of an oscillating frequency meter.

Measurement of coil resistance. The effect of resistance upon the sharpness of resonance and the selectivity of the circuit has been mentioned. The **resonance curve** furnishes one method of measuring the resistance in a given circuit, provided that we know the inductance of the coil, which can be calculated from the formula in Fig. 4–6.

Experiment 4–10. To determine the resistance of a coil. Couple a series circuit composed of a coil, condenser, and indicating meter to a generator of about 5-watt output. Adjust the frequency of the generator to resonance with the series circuit. If the generator has a constant output over this frequency range the accuracy with which the coil resistance is determined will be greater. Pick out the two frequencies (f_2 and f_1) above and below resonance where the current in the circuit is 0.707 of its value at resonance (f_r), and calculate the width of frequency band at this point and the resistance of the circuit, from the equation

$$R = \frac{\omega L(f_2 - f_1)}{f_r}$$

Subtract from this value the resistance of the current meter. For example, a model 425 Weston thermogalvanometer will read currents of 115 ma and has a radio-frequency resistance of 4.5 ohms. The value of resistance remaining is the resistance of coil, leads, and condenser. Most of this resistance resides in the coil.

Experiment 5–10. To determine resistance of a circuit. Another method of determining the resistance of a coil is as follows: It necessitates the use of a decade resistance box or series of accurately known resistances of negligible inductance and capacitance and a variable condenser.

Small lengths of high-resistance wire (manganin) are to be preferred for frequencies higher than 1000 kc. Their d-c and high-frequency resistance is practically the same.

Connect the apparatus in series and couple to an oscillator.

With the resistance box short-circuited ($R = 0$), tune the circuit to resonance. Then add enough resistance to the circuit to reduce the current to one-half its resonant value, retuning to resonance, if necessary. Then, since we have halved the current, Ohm's law tells us that we have doubled the resistance. In other words, the added resistance is equal to the resistance already existing in the circuit. Again subtract the resistance of the current-indicating meter. What remains is the resistance of coils, condensers, and leads.

Repeat at several different frequencies, and calculate the Q and plot against frequency and wavelength.

If only one or two resistance units are available, say 5 or 10 ohms, and not a continuously variable standard of resistance like a decade box, the resistance of the circuit above may be determined by noting the current at resonance, and the current when some resistance has been added, retuning to resonance after adding the resistance if necessary. Then the currents, according to Ohm's law, are

$$I_1 = \frac{E}{R_1}$$

$$I_2 = \frac{E}{R_1 + R_2}$$

where I_1 = current at resonance and no added resistance.
I_2 = current at resonance and R_2 added.
R_1 = resistance of circuit.
R_2 = added resistance.

Then

$$R_1 = \frac{R_2 I_2}{I_1 - I_2}$$

If a current-indicating meter is used whose deflections are proportional to the current squared, such as a thermogalvanometer, it is only necessary in this experiment to add sufficient resistance to quarter the deflection of the instrument. This is equivalent to halving the current, and the added resistance is equal to the resistance already in the circuit.

The lower the resistance of the current-indicating device, the greater will be the accuracy with which such measurements may be carried out. For example, if the indicator has a resistance of 4.5 ohms and the circuit a resistance of 5 ohms, great accuracy cannot be attained, but if the circuit resistance is double or treble that of the indicator, much greater accuracy results. In any event the meter resistance must be subtracted from the measured resistance to get the resistance due to the circuit alone.

Condenser capacitance. We will now investigate, by means of an experiment, the capacitance of a condenser.

Experiment 6–10. To determine the capacitance of a condenser. Connect a variable condenser whose calibration is known across a coil and click into an oscillating receiver or into an oscillating frequency meter; attach the unknown condenser across the variable condenser and retune the variable condenser to resonance with the frequency meter. The difference in readings of the calibrated condenser is the capacitance of the unknown condenser. For example, suppose that resonance is obtained by the variable condenser alone when set at 400 $\mu\mu$f. Connecting the second condenser across the variable forces us to reduce the capacitance of the variable to 340 $\mu\mu$f. The difference $400 - 340 = 60$ $\mu\mu$f is the capacitance of the unknown. Such a method enables the experimenter to disregard the capacitance of the coil itself or of the leads since they are connected across the variable at all times and do not change when the unknown is attached to the circuit.

Antenna wavelength. We will proceed to determine the wavelength of an antenna by means of the following experiment.

Experiment 7–10. To measure the natural wavelength of an antenna. Connect in series with the antenna an inductance which can be adjusted in even steps, say a coil of 20 turns with taps at each turn. Measure the frequency to which the antenna tunes with the entire coil in the circuit by coupling the coil to an oscillating wavemeter. Then reduce the inductance by one turn, and repeat. Repeat until accurate readings are no longer possible. Plot wavelength, or frequency, against added turns of wire. Where the line crosses the wavelength or frequency axis is the natural wavelength or frequency of the antenna.

Antenna capacitance. The capacitance of an antenna may also be determined by experiment.

Experiment 8–10. To measure the capacitance of an antenna. Measure the wavelength of an antenna attached to an inductance, as in Fig. 6–10. Then replace the antenna-ground connections by a variable condenser (Fig. 6–10*b*), and tune the condenser until resonance with the frequency meter is indicated. The capacitance of the condenser at this point is the capacitance of the antenna.

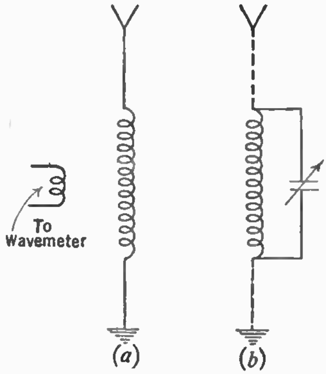


FIG. 6–10. To measure the capacitance of an antenna.

Antenna inductance. By the following experiment we can determine the inductance of an antenna.

Experiment 9–10. To determine the inductance of an antenna. Connect a known inductance, L_1, in series with the antenna, and measure the frequency f_1. Repeat, using a different inductance L_2, and get frequency f_2. The two frequencies are related as follows:

$$f_1 \propto \frac{1}{\sqrt{(L_1 + L_a)C_a}}$$

$$f_2 \propto \frac{1}{\sqrt{(L_2 + L_a)C_a}}$$

where L_a = antenna inductance.
C_a = antenna capacitance.
$\propto$ indicates "is proportional to."

Squaring both equations and solving for L_a we get

$$L_a = \frac{f_1^2 L_1 - f_2^2 L_2}{f_2^2 - f_1^2}$$

Problem 1–10. A frequency meter is being calibrated from a standard. At resonance the capacitance of the standard is 400 $\mu\mu$f, the capacitance of the other meter is 500 $\mu\mu$f. What is the ratio of their inductances? If the inductance of the standard is 300 μh, what is the inductance of the other frequency meter? At what frequency are they now set? What is the wavelength?

Problem 2–10. In calibrating a frequency meter from a source of 1000 kc which has many harmonics, the fundamental is received when the condenser capacitance is 253 $\mu\mu$f. Another indication is received when the capacitance is approximately 65 $\mu\mu$f. What harmonic is this?

Problem 3–10. In an experiment to determine the resistance of a coil, the current without added resistance is 100 ma and with 12 ohms added it is 60 ma. The resistance of the meter is 5 ohms. What is the resistance of the coil and condenser in series?

Problem 4–10. A certain coil and condenser (LC_1) tunes to a frequency, f_1. The condenser is changed by adding another to it so that the value is C_2. The circuit now tunes to a frequency, f_2. Prove that $f_2 \div f_1 = \sqrt{C_1} \div \sqrt{C_2}$.

Problem 5–10. Using the formula in Problem 4–10, what must be done to the capacitance to make a circuit tune to twice the frequency? half the frequency? double the wavelength? one-half the wavelength?

Problem 6–10. A coil-condenser combination tunes to 450 kc when the condenser is 600 $\mu\mu$f. When an unknown condenser is placed in series with the 600-$\mu\mu$f capacitance the circuit tunes to 600 kc. What is the unknown capacitance?

Problem 7–10. An antenna tunes to 1000 kc when 100 μh is in series with it, and 750 kc when 300 μh is in series. What is the inductance of the antenna? Remembering that the two inductances, L_1 or L_2, and the inductance of the antenna L_a are in series, and can be added to get the total inductance, what is the capacitance of the antenna? What is its natural frequency?

Problem 8–10. A radio receiver tunes to 600 meters (500 kc) with a 500-$\mu\mu$f condenser. What maximum capacitance must be put in parallel with this condenser in order to receive 800-meter (375-kc) radio compass signals and ship-to-shore traffic on 2200 meters?

Coil-condenser applications. Combinations of coils and condensers in circuits perform useful services throughout the electric communication art. Not only is the selectivity of a radio receiver produced by the properties of a resonant circuit, but such circuits may also be used to eliminate undesired frequencies or frequency bands and to change the response of a system to frequencies of certain values.

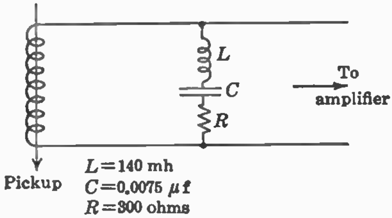


FIG. 7–10. Use of a series-resonant circuit as a low-impedance trap to reduce noise from needle scratch.

For example, a "scratch" filter may be used with a phonograph amplifier to reduce the noise caused by needle scratch. This filter is a series-resonant circuit placed across the circuit between the phonograph pickup and the amplifier it feeds into, as shown in Fig. 7–10. This circuit tunes to about 4500 cycles, and, since it is a series-resonant circuit, its impedance to frequencies near 4500 cycles will be very low. It will provide a low-impedance path for these frequencies, and since needle scratch is of this order of frequency the scratch noise will be reduced. Of course, any music frequencies in this region will be reduced also.

Filters. Since coils and condensers have varying effects upon a circuit, one having a reactance which increases as the frequency increases and the other decreasing as the frequency in-

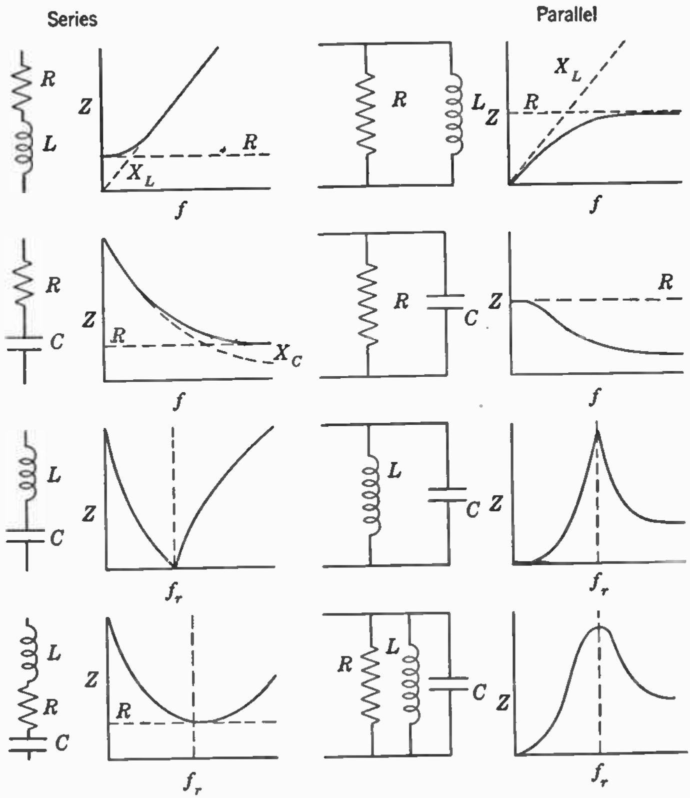


FIG. 8–10. Impedance diagrams of series and parallel combinations of L, C, and R.

creases, combinations of L and C may perform many useful functions. **Filters** are combinations of L and C that may be used to transmit only a required portion of a wide frequency spectrum or, conversely, to attenuate certain frequencies.

A filter that passes low frequencies but attenuates frequencies above the **cut-off frequency** is known as a **low-pass** filter; a combination of L and C that passes high frequencies but attenuates lower frequencies than the cut-off frequency is called a **high-pass** filter. If a band is attenuated while all other frequencies higher and lower than this band are transmitted, the filter is a **band-elimination** filter; if certain frequencies are passed but all others are attenuated, the unit is called a **band-pass** filter.

If sharper cut-off is desired, several sections of these elementary filters may be connected in series, each attenuating still further the energy transmitted by the one before it.

The values of L, C, and the terminal resistance, R, between which the filter works, for the simple filters shown in Fig. 9–10 are given in Table IV.

TABLE IV

VALUES OF L AND C FOR FILTERS

I. $L = R/\pi f_r;\quad C = 1/\pi f_r R$

II. $C = 1/4\pi f_r R;\quad L = R/4\pi f_r$

III. $\begin{cases} L_1 = (f_1 - f_0)R/\pi f_0 f_1; & C_1 = 1/4\pi(f_1 - f_0)R \\ L_2 = R/4\pi(f_0 - f_1); & C_2 = (f_1 - f_0)/\pi R f_0 f_1 \end{cases}$

IV. $\begin{cases} L_1 = R/\pi(f_2 - f_1); & C_1 = (f_2 - f_1)/4\pi f_2 f_1 R \\ L_2 = (f_2 - f_1)R/4\pi f_2 f_1; & C_2 = 1/\pi(f_2 - f_1)R \end{cases}$

Frequency regulation. A simple example of the use of the diverse effects of L and C is shown in Fig. 10–10. As long as the frequency is equal to the series-resonant frequency of L and C, the two tubes get equal voltages and produce equal currents through the output transformer. Since these two currents are opposite in direction, no current will pass through the secondary transformer. If, however, the input frequency rises, more voltage will be applied to the L tube than to the C tube, and more

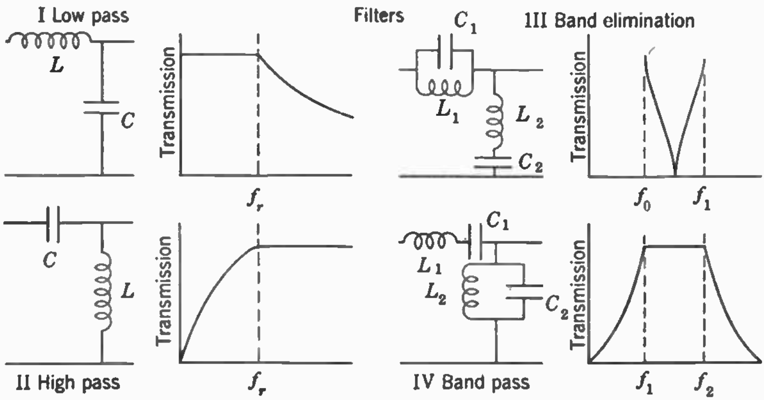


FIG. 9-10. Examples of the use of coils and condensers as filters.

current will flow through it. Conversely, if the frequency drops, a greater voltage will be impressed upon the C tube and more current will flow in it.

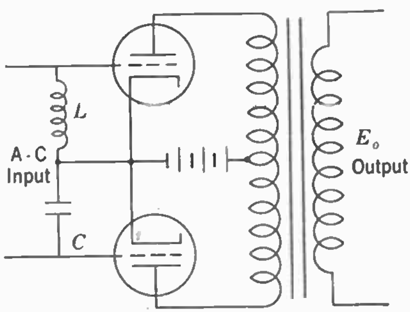


FIG. 10-10. Use of L and C as a frequency-regulating network.

These variations in current, since they are in opposite directions, can be applied to a circuit to maintain the input voltage to the two tubes constant in frequency. Thus the circuit may be used as a frequency-control system.

CHAPTER 11

VACUUM TUBES

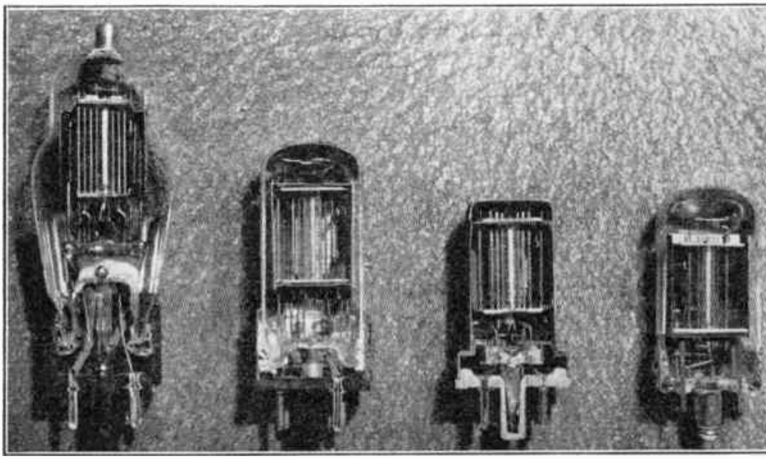
"The radio tube is a marvelous device. It makes possible the performing of operations, amazing in conception, with a precision and a certainty that are astounding. It is an exceedingly sensitive and accurate instrument. Its future possibilities even in the light of present-day acomplishments are but dimly foreseen; for each development opens up new fields of design and application. The importance of the radio tube lies in its ability to control almost instantly the flight of millions of electrons, with a minimum of control energy." *

The principle of the radio or electron tube is very simple. A source of electrons forms the basic part of the tube. These electrons are attracted to a positively charged electrode called the **anode** or **plate.** Electrons will not flow in the reverse direction. A simple two-element tube, therefore, will act as a rectifier, passing current only when the anode is positive with respect to the source of electrons known as the **cathode.** If a third element in the form of an open mesh of wires is put between the cathode and the plate, positive or negative voltages placed on this grid will control the rate at which electrons arrive at the plate. This control can be exercised almost instantaneously, and, since the grid catches very few electrons so that almost no current flows in the grid circuit, engineers are provided with a quick and efficient device for controlling the flow of an electric current which may be as small as a fraction of a microampere or as great as many amperes.

The tube will produce direct current from alternating current, or vice versa. It will act as a voltage or current amplifier or

* From the *RCA Receiving Tube Manual RC-14.*

will perform as a frequency changer. It will separate audio and radio frequencies, or it will mix them so that the radio frequencies act as a carrier for the audio frequencies. It will convert voltage changes into changes of light intensity, and vice versa. It will act as a voltmeter, current meter, wattmeter, or frequency meter. It will control industrial machinery with a



Constructional details of 6J7G, 6SJ7GT, 6SJ7, and 6C7. Courtesy of Sylvania Electric Products Co.

precision possible in no other way. The radio tube is indeed a modern miracle.

Edison effect. In his search for more efficient filaments for incandescent electric lamps, Thomas A. Edison discovered that the glass bulbs of the lamps were blackened in the form of a shadow of the filament. An investigation discovered that particles were leaving one side of the filament and crossing the evacuated space to the other leg of the filament. When they missed the filament and hit the glass they blackened it. Subsequent research disclosed the fact that the particles carried an electric current which could be collected by a positive metallic plate placed within the bulb. The particles were electrons, although Mr. Edison did not know it, and the phenomenon by which the bulbs were darkened is now known as the Edison effect.

This discovery lay dormant for a long time. Then Fleming, in England, made use of the ability of the filament to pass current in only one direction (rectification) to construct a more sensitive radio detector than had existed up to that time. Another long period elapsed during which the Fleming "valve" was widely used for this purpose but during which little or no improvement was effected. Then de Forest, in this country, made a most important invention. He introduced a third element, the control grid, between the filament (cathode) and the electron collector (anode).

Much work remained to be done before the radio tube (de Forest called it the "audion") reached its present stage of development, out of which have come broadcasting, the many applications of tubes to industrial processes, and, of course, radar and the numerous other important military uses.

Electron emission. The heart of the tube is the source of electrons. There are several ways in which free electrons are obtainable.

1. Thermionic emission. When a metallic filament such as tungsten is heated to a sufficiently high temperature (2700° C for pure tungsten) certain electrons escape from the tungsten atom, break through the surface of the filament, and are free to be moved about by electric fields. This process resembles the escape of molecules of water from the surface of water heated to the boiling point. At the surface of water, or of tungsten, exists a barrier through which the moving molecules, or electrons, cannot pass under ordinary circumstances; but when they have sufficient energy imparted to them, as by heating, they break through the "surface tension" and escape.

2. Secondary emission. When an electron is accelerated sufficiently it may have enough energy imparted to it to knock one or more electrons out of any material with which it comes in contact, either metal or insulator. A positively charged electrode situated near the source of these "secondary" electrons will collect them. In practical tubes the secondary electrons may be attracted back to the electrode from which they came, as from the plate, or they may be driven to another electrode which is maintained positively charged.

3. Photoelectric emission. When light of proper wavelength is allowed to fall upon certain metals, electrons are released from the surface of the metal. In the photoelectric cell, or phototube, the released electrons are attracted to a positively charged electrode.

4. Field emission. Electrons may be pulled out of the surface of metals if a sufficiently high positive potential is placed near it. The voltages must be quite high and concentrated in small space so that the voltage drop at a particular point is very great. This method of obtaining electrons is known as field emission since it is the high electric field which enables the electrons to escape from the metal.

In radio work, tubes employing thermionic emission are used almost to the exclusion of the other possible sources.

Types of electron emitters. Cathodes are of two general types. In one type the filament itself acts as the source of electrons. It is heated by battery or other d-c source or from an a-c source. In the second type the filament merely acts as a heater for raising the actual electron emitter to the proper temperature; it is placed within a metallic sleeve or cylinder which has a highly efficient emitting substance as an outer coating. This type of emitter is said to have an indirectly heated cathode.

If the first type of tube is operated from alternating current, care must be taken to see that the electrons from the anode or plate circuit are returned to the center of the filament rather than to one end. Otherwise, the cyclic change in voltage of one end of the filament with respect to the other, as the current changes its direction, produces hum in the output circuit. Usually the filament is heated by a step-down transformer, and, if the electrons from the plate circuit are returned to a center tap on this transformer, hum will be minimized.

Since the indirectly heated cathode need have no electrical connection to the filament, the filament can be operated from alternating current without having the alternating voltages impressed upon the cathode. The cathode, therefore, acts as though all portions of it were at the same potential, and it is called a **unipotential cathode.** No hum results from its use on alternating current if proper care is taken in constructing the tube and in its operation.

Emitter materials. Electron emitters are usually constructed of one of three substances, namely, tungsten, thoriated tungsten, and certain metallic oxides. Pure tungsten filaments operate at high temperature, consume appreciable filament power, but are mechanically sturdy. If the tungsten filament is impregnated in manufacture with thorium, the operating temperature required for copious emission is materially reduced. Less heating power is required; but the filament is not so strong as a pure tungsten filament and can be overloaded easily. If platinum or nickel or a nickel alloy is coated with certain alkaline earths (barium or strontium oxide) the temperature required for emission is still further reduced.

In general, receiving tubes and small transmitting tubes, say up to 30 watts plate dissipation, have oxide-coated cathodes; transmitting tubes up to about 1000 watts plate dissipation have thoriated-tungsten filaments; and larger tubes have pure tungsten emitters. All mercury-vapor tubes have oxide-coated cathodes.

Visually, tungsten filaments are white when heated to operating temperatures, thoriated tungsten filaments are yellow, and oxide-coated cathodes are dull red.

Directly heated cathodes (filament-type tubes) are more efficient than the heater-cathode (unipotential) types.

Diode tubes. The simplest radio tube consists of a **cathode** which acts as a source of electrons and an **anode** which acts as a collector of electrons. The characteristics of such a tube, known as a **diode** or two-element tube, can be studied by connecting it as shown in Fig. 1–11. When the cathode is heated, by means of a battery or other source of electrical power, and when the plate circuit is completed, current will flow as indicated by the milliammeter. The amount of current flowing depends upon two factors only: (1) the temperature of the cathode and (2) the voltage between plate and cathode. If the current is noted for each value of plate voltage we shall obtain a characteristic curve which shows the relationship between plate voltage and plate current for a given cathode temperature. It will be found that the current increases in proportion to the 3/2 power of the voltage.

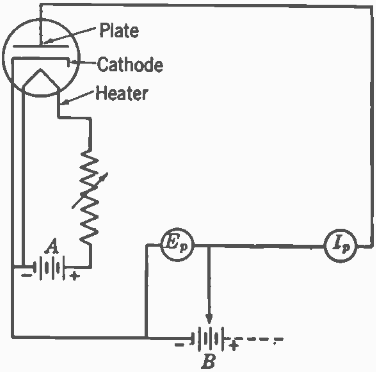


Fig. 1-11. Connections for studying action in a two-element or diode tube. Here the two elements are the cathode and the plate, the heater serving only to heat the cathode to emitting temperature.

If the plate-cathode voltage is increased continuously, it will be found that, ultimately, the plate current no longer increases. At this point all the electrons emitted by the cathode are attracted to the plate. Further increase in plate voltage merely speeds up the electrons in their flight through the tube but cannot increase the number per second that arrive.

Under these conditions the only way to increase the current flow is to increase the cathode temperature. Therefore if we fix the plate voltage and increase the cathode temperature (by increasing the filament voltage) we shall get another curve which expresses the relation between cathode temperature and plate current. If the cathode temperature is increased continuously, the plate current will rise sharply but the curve will ultimately level off. Under this condition the cathode may be emitting more electrons than actually reach the plate. An increase in plate voltage will again increase plate current, however.

Space charge. Electrons not collected by the plate tend to congregate in the space between cathode and plate and form a negative cloud. Electrons leaving the cathode now find this negative barrier between them and the plate and of course are repelled by it. Many of the electrons return to the cathode, some pass on to the plate, and others remain to form the **space charge.** The negative cloud has the same effect as decreasing the positive plate voltage; and the only way to increase the flow of electrons to the plate is to increase the plate voltage.

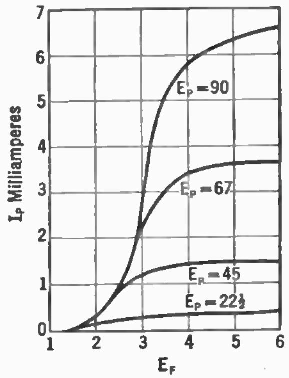


FIG. 2–11. Curves of typical diode, showing that plate current is controlled by plate voltage, E_P, and filament voltage, E_F.

Tubes are ordinarily so operated that more electrons are emitted than are taken by the plate. Space charge, therefore, is a factor in controlling plate current.

Tube voltages. The filament battery is called the A battery; the plate battery is called the B battery; and if there is a grid battery it is known as the C battery.

Diode operation. In practice the filament current is many times greater than the plate current. For example, in an ordinary rectifier tube such as the 5Y3, 2 amp (2000 ma) is required to heat the filament. From this filament acting as cathode a maximum of 375 ma plate current is permissible without overheating the plate.

If a plate battery is connected in series with a resistance and the diode, current flowing through the circuit represents power taken from the battery. The power consumed in the resistance is equal to the product of the current through and the voltage across that resistance and may easily be calculated. The power taken from the battery is equal to IE, and this is greater than the power consumed in the resistance. Where does the rest of the power go?

It is common knowledge that a rectifier tube in a radio receiver or transmitter gets hot in operation. Neglecting the power supplied by the filament battery, power is being consumed in the plate circuit of the tube, and it is this power that makes up the difference between the power taken from the B battery and that used up in heating the resistance. The tube, therefore, acts as though it were a resistance and the power consumed in it is equal to I^2R_p, where I is the plate current and R_p represents the d-c plate resistance of the tube. Thus

Battery power = IE
Power in resistance = I^2R
Power in tube = Battery power − Power in resistance

$$= IE - I^2R = I^2R_p$$

R_p can be found for any value of plate current by dividing the power consumed in the tube by the current squared.

In Fig. 3–11 is shown the $E_b - I_b$ characteristic of a typical transmitter rectifier, the 217-A. The value of the plate resistance of the tube at any value of plate current may be found by dividing the voltage by the current. The reader should calculate the d-c plate resistance for several values of plate current and should plot these values against plate current. He should note that the plate current does not follow Ohm's law; that is, doubling the voltage does not double the current. As a matter of fact, the current varies as the 3/2 power of the voltage. Thus, doubling the voltage increases the current $2^{3/2}$ times; multiplying the voltage by 3 increases the current $3^{3/2}$ times.*

* The reader need not be confused by these fractional powers. For example, 2^3 equals 2 multiplied by itself 3 times; $2^{1/2}$ equals the square root of 2. Therefore $2^{3/2}$ equals 2 cubed and then the square root of the product taken. Thus $2^{3/2} = (2 \times 2 \times 2)^{1/2} = \sqrt{2 \times 2 \times 2} = \sqrt{8} = 2.8$. Let the reader take values from Fig. 3–11 and see how closely the 217-A follows the 3/2 power law.

It will be noted that the plate resistance varies with the current through the tube. If the tube is acting as a rectifier, that is, if alternating current is impressed between cathode and plate, no current flows on the portion of the cycle which makes the plate negative with respect to the cathode. Under this condition the plate resistance is infinite. Throughout the half cycle in

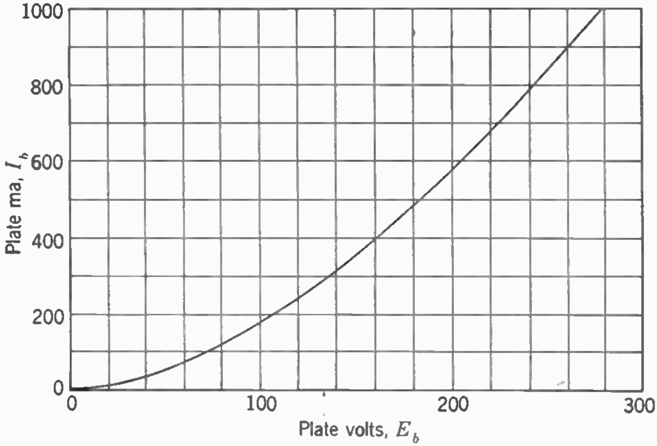


FIG. 3-11. Typical diode plate current curve for a fixed value of filament voltage sufficient to supply all the electrons that will be collected by the plate.

which the plate is positive, the current through the tube is varying and the tube resistance is varying also.

The d-c plate resistance of vacuum-type rectifiers is fairly high—hundreds or thousands of ohms. It may be made low in two ways: (1) by placing the plate and cathode very close together as in the 83-v and the 25Z5, where the spacing is 0.002 in.; or (2) by placing some gas in the tube.

The d-c plate resistance actually determines the power required to drive electrons through the tube. When the electrons strike the plate the power is dissipated in the form of heat.

Gaseous tubes. If a gas such as mercury vapor is introduced, the tube characteristics change markedly. The d-c plate resistance becomes very much lower, and the voltage drop across it becomes decidedly lower and is essentially independent of the current. The tube glows with a characteristic color when plate current flows.

Mercury vapor consists of atoms of mercury. Electrons leaving the cathode collide with the mercury atoms and liberate electrons from them. Thus a new supply of electrons is available to carry current to the plate. But another most important effect has been produced. The mercury atom, having lost one or more electrons, is positively charged. It is large and heavy and does not get very far from the vicinity of the cathode. But this is the region where the space charge, the cloud of negatively charged electrons, is located. A sufficient supply of positively charged mercury atoms (called **ions**) will neutralize the space charge and remove this inhibiting influence. Electrons are now free to move toward the plate in quantities limited only by the supply and the number taken by the plate. There is some voltage drop within the tube, but it is of the order of 15 volts, which is the emf required to ionize the mercury.

Since gaseous tubes have constant voltage drop, putting two in parallel will not increase the load current unless the resistance of the load is decreased. More current will not flow in a load if two equal-voltage batteries are connected in parallel. More current is *available*, but the load must be changed to utilize it.

Inverse peak voltage. How much voltage can be placed across a tube? The amount is limited by the construction of the tube. Excessive voltage will cause arcing and consequent damage. Spacing between cathode and plate, the kind of insulation, and the effect of heat on the electrodes and the glass all govern the maximum voltage permissible. On the half cycle that the tube conducts current, if it is operating as a rectifier, the voltage drop across the tube is fairly low. It is actually the impressed voltage minus the voltage drop across the load. But during the half cycle in which the tube is not conducting, the total impressed voltage is across the tube since there is no voltage drop in the load for the reason that no current is flowing.

The voltage across the tube during the half cycle when the tube is not conducting is known as the inverse voltage, and the maximum or peak value that may safely be placed across a tube is called the **maximum peak inverse voltage.** It is an important tube "rating."

Triodes. If a third element, like an open-mesh grid or similar structure, is placed between cathode and plate, the versatility of the tube increases enormously. Such a tube is called a **triode** or three-element tube. Now the tube will not only rectify, i.e., change alternating to direct current, but it will also perform the many other functions for which the electron tube is so well known.

The purpose of the grid. The grid has several important functions. It may be used to neutralize the space charge so that greater plate current may flow with a given filament temperature and given plate voltage.

Suppose that the grid is made positive with respect to the cathode. Since it is physically nearer the cathode than the plate is, a small positive potential will have the same effect as a large positive potential on the plate. A positive potential near the filament prevents the building up of a high negative space charge, making it easier for the electrons to go to the plate.

Suppose, however, that we make the grid negative. Its negative potential, near the cathode, will have more effect on electrons than the positive potential on the plate, which is farther from the cathode. A small negative voltage on the grid may prevent the plate from getting electrons even if the plate is at a positive potential. Since there is practically no time lag in the flow of electrons, the grid takes instantaneous control. Therefore the grid is a **control** electrode.

Characteristic curves. In practice the cathode is heated so that more electrons are emitted than are needed. The rate of electron flow between cathode and plate is determined by the plate and grid voltages. Raising these voltages in a positive direction will increase the plate current. The cathode temperature is not varied in operation.

Under these conditions the plate current depends upon two factors: the grid and plate voltages. If we choose a value for either and vary the other, we will get data for a characteristic curve showing the relation between plate current and either grid or plate voltage. A circuit for studying the characteristics of a tube is shown in Fig. 4–11. Meters measure the several voltages and the plate current.

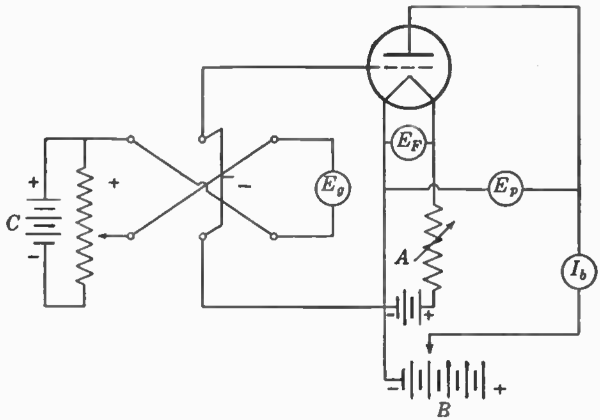


FIG. 4–11. Apparatus for studying tube characteristics. The DPDT switch reverses the polarity of the grid voltage.

Experiment 1–11. Effect of grid bias upon plate current. Set up the apparatus as shown in Fig. 4–11, using in succession several of the common types of tubes. Set the filament or cathode heater at the proper voltage. Fix a small voltage on the plate of the tube, and take down data showing how the plate current changes as the grid voltage (known as a **bias**) is varied from a value negative enough to make the plate current zero to a value a few volts positive. The DPDT switch in the grid circuit permits changing the polarity of the grid without changing the meter (E_g) connection. Then increase the plate voltage and repeat. Plot these data like those in Fig. 5–11.

Grid voltage-plate current curves. Several interesting and important facts may be discovered by looking at grid voltage-plate current curves, which we shall call the E_c-I_b curves. At large negative grid voltages there is little or no current in the plate circuit. As this negative voltage is decreased, some electrons get past the grid and through the space charge and to the plate. The current begins to flow, increases at a rather slow rate, then more rapidly, then in a steep and straight line, and finally, if the experiment is carried far enough, the curve flattens

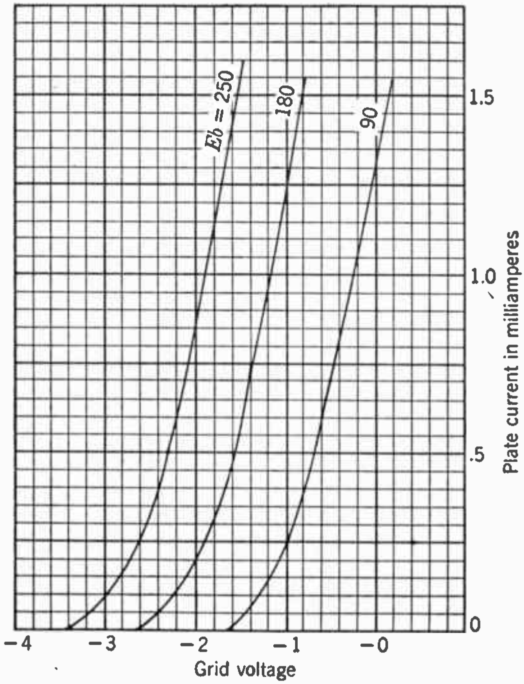


FIG. 5–11. Typical E_c-I_b curves. Note that the several curves corresponding to different plate voltages are parallel and have long straight portions.

out. Increasing the plate potential and again varying the grid potential produces a new curve which is essentially parallel to the first, but moved to the left. Increasing the plate voltage a like amount again produces a new curve displaced an equal distance to the left of the second line. Such a graphic collection of data, known as a family of curves, tells all we need to know of the effect of grid voltage upon plate current. The steady grid voltage is called a **C** or **grid bias**.

If we place a meter in the grid circuit we shall see that a very minute grid current, seldom more than one-tenth the plate current, flows.

Effect of plate voltage upon plate current. To determine this effect we will resort to experiment.

Experiment 2–11. Set up the apparatus as in Experiment 1–11. Set the grid voltage at some value, say −5 for an ordinary receiving tube, and note the plate current as the plate voltage is changed from 0 to perhaps 100 volts in 10-volt steps. Then change the grid voltage to −10 and repeat; then at −15, −20, 0, +5, +10, etc.

These data may be taken from the E_c-I_b curves plotted in Experiment 1–11 by picking from the curves the proper values of current, and plate and grid voltages.

Plate voltage-plate current curves. Here again (Fig. 6–11) the plate-current curves are essentially parallel over the straight

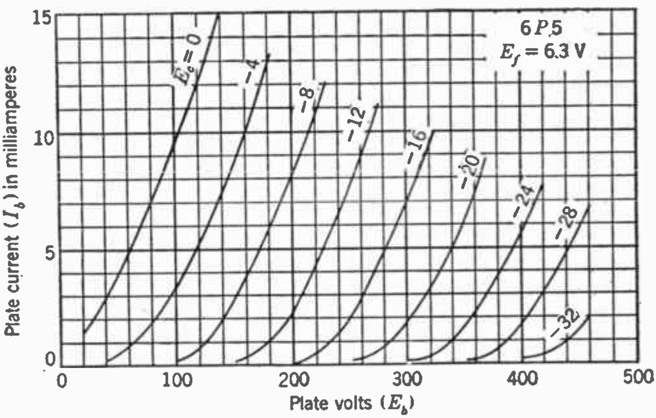


FIG. 6–11. Family of E_b-I_b curves for a triode tube.

parts. If the grid voltages chosen are in equal steps, the plate-current curves will be equal distances from each other.

Note that any one curve, for example, the E_b = 180 volt curve of Fig. 5–11, tells us what the plate current will be for values of grid voltage between −2.6 and −0.8 volts. Similarly in Fig. 6–11 the E_c = −4 curve shows what the plate current will be for values of plate voltage between 40 and 180 volts.

From characteristic curves of this type, we may calculate all the tube constants and foretell nearly all the properties of the tube when connected into a circuit with other apparatus whose electrical constants we know.

Amplification factor. For example, we know that the grid potential is relatively more important than the plate voltage in controlling plate current. Why? Because it is nearer the source of electrons. How much? We can tell from the E_b-I_b curves in Fig. 6–11. Looking at the line marked $E_c = -20$, we see that at a plate voltage of 360 the plate current is 7.5 ma but that if E_b is decreased to 300 volts the plate current decreases to 2 ma, a change of 5.5 ma for 60 volts or a net change of 0.092 ma per volt. The slope of this particular line, 0.092 ma per volt, gives us another important tube constant without further calculation, as we shall see later.

Now, looking at the two curves marked $E_c = -20$ and $E_c = -16$ at the points where they cross the 300-volt E_b line, we see that at a plate voltage of 300 the plate currents are respectively 2 and 7 ma. We see that changing the grid voltage by 4 volts causes a change of 5 ma in the plate current, a net change of 1.25 ma per volt. Dividing the change per volt caused by E_c variations by the change per volt produced by E_b variations gives the relative ability of the grid and plate potentials to influence the plate current. Thus,

$$\frac{\text{Ability of grid voltage to control plate current}}{\text{Ability of plate voltage to control plate current}} = \frac{1.25 \text{ ma/volt}}{0.092 \text{ ma/volt}} = 13.6$$

This ratio is defined as the **amplification factor** of the tube. It is the ratio between the plate-voltage change required to produce a certain plate-current change and the grid-voltage change required to produce the same change in plate current. The Greek letter mu, μ, is the symbol commonly used for the amplification factor of a tube.

The amplification factor for a given tube does not vary much under the conditions under which the tube is ordinarily used.

It is controlled largely by the mechanical construction of the tube and by the nearness of the grid to the cathode. A grid composed of many wires close to the cathode produces a high amplification factor; a grid with a wide mesh and not so close to the cathode produces a tube with a low amplification factor.

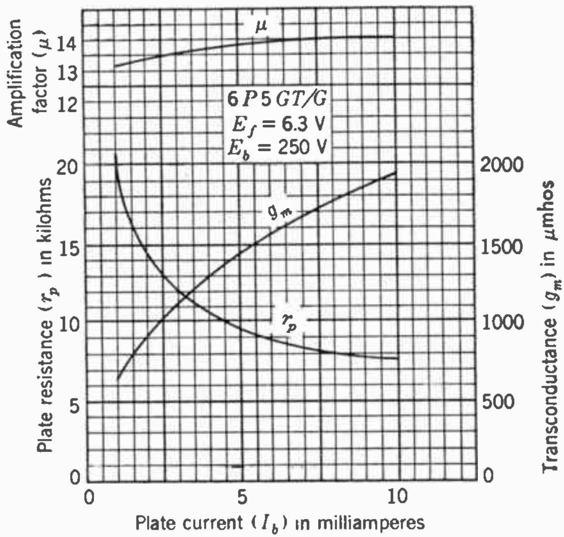


Fig. 7–11. Dependence of r_p, g_m, and μ on the plate current. Note that μ is almost constant in value and that g_m increases as r_p decreases.

The student should note that the amplification factor is *not* the ratio between plate and grid voltages but the ratio between *changes* in these voltages.

$$\mu = \frac{\Delta E_b}{\Delta E_c} \text{ to produce a given } \Delta I_b$$

where the symbol Δ signifies "a change in."

Internal plate resistance. Like a rectifier tube, the triode imposes resistance to the flow of electrons or current. The actual resistance to the flow of direct current is the ratio between the voltage and the current, but since the tube is more often used with both direct and alternating current flowing through the plate circuit, and since the direct current is merely incidental, it is important to have a knowledge of the internal resistance to a flow of alternating current.

Ordinarily the tube is operated with certain fixed grid (E_c) and plate (E_b) voltages. Then in addition an alternating voltage is applied to the grid (sometimes to the plate) and a certain amount of alternating current flows in the plate circuit. What is the resistance to the flow of this alternating current?

If the plate voltage is varied slightly, say a volt or so, and the corresponding change in plate current is noted, then the ratio between these **changes** is called the **dynamic plate resistance** r_p. Thus

$$r_p = \frac{\text{Change in plate voltage}}{\text{Change in plate current}} = \frac{\Delta E_b}{\Delta I_b}$$

For example, reducing the plate voltage from 180 to 140 volts (Fig. 8–11) produced a plate current change of 7.8 ma (0.0078 amp).

$$r_p = \frac{\Delta E_b}{\Delta I_b} = \frac{180 - 140}{0.016 - 0.0082} = \frac{40}{0.0078} = 5200 \text{ ohms (approx.)}$$

Problem 1–11. The plate current of a tube is 4.5 ma when $E_b = 90$ volts, and 0.9 ma when $E_b = 40$ volts. What is the plate resistance?

Problem 2–11. The plate resistance of a tube is 12,000 ohms. At $E_b = 140$, $I_b = 14.5$ ma. What is I_b when $E_b = 100$ volts?

Since the dynamic plate resistance, r_p, changes with plate- and grid-voltage changes, the conditions of both must be considered when the resistance is mentioned. Thus the 2A3 tube has a dynamic plate resistance of 800 ohms when the plate voltage is 250 volts and the grid voltage is 45 volts negative. Its internal resistance changes if either E_b or E_c is varied.

Transconductance of a tube.* One more important tube constant remains. The **transconductance** is the factor which tells

* In the older literature this tube constant is called "mutual conductance." Transconductance is the preferred present-day term.

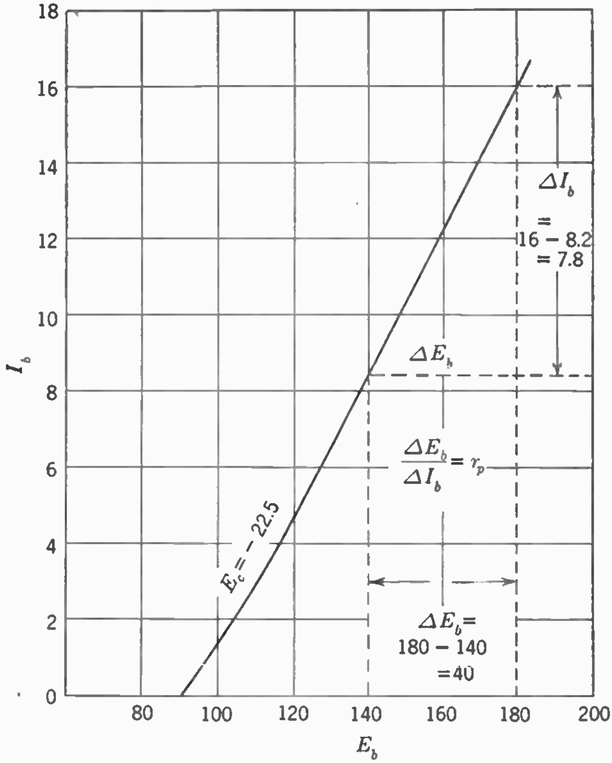


FIG. 8–11. Method of calculating plate resistance from E_b-I_b curve.

how much plate-current change is caused by a given grid-voltage change. Thus

$$g_m = \frac{\text{Change in plate current}}{\text{Change in grid voltage}} = \frac{\Delta I_b}{\Delta E_c}$$

Thus, if a change of 1 volt on the grid produces a change of plate current of 1 ma, the transconductance

$$g_m = \frac{1 \times 10^{-3}}{1} = 1 \times 10^{-3} \text{ mho or 1000 micromhos}$$

The **mho** is the unit of conductance. Note that it is ohm spelled backwards.

The transconductance is numerically equal to the ratio of the amplification factor to the plate resistance. Thus

$$g_m = \frac{\Delta I_b}{\Delta E_c} = \frac{\mu}{r_p}$$

Problem 3–11. A tube has a plate current of 7.75 ma at zero grid bias and 3.8 ma at $E_c = -4$. What is the transconductance?

Problem 4–11. The transconductance of a tube is 800 micromhos. What change in plate current is produced by a 1-volt change on the grid?

Problem 5–11. The transconductance of a tube is 775 micromhos. The plate current at $E_c = -6$ is 2 ma. What is the plate current at $E_c = -2$?

Importance of transconductance. The tube is ordinarily so worked that variations in the input a-c voltage (grid voltage) produce variations in output (plate circuit) a-c current, and it is important that the transconductance of a tube be high. Amplification factors as low as 3 and as high as 100 are used for simple three-element tubes with plate resistances varying from 800 to 100,000 ohms and with transconductances of the order of 1000 to 4000 micromhos.

Slopes of characteristic curves as tube constants. Since the plate resistance of the tube is defined as $r_p = \frac{\Delta E_b}{\Delta I_b}$, and the transconductance as $g_m = \frac{\Delta I_b}{\Delta E_c}$, these values may be taken directly from the E_b-I_b or E_c-I_b curves.

The transconductance is the slope of the E_c-I_b line. When changing the grid voltage produces no change in plate current the transconductance is zero, that is, the E_c-I_b curve no longer has any slope or steepness and it flattens out, as for high positive values of grid bias. On the other hand, when the E_b-I_b curve flattens out the plate resistance becomes infinite, and so the steeper the E_b-I_b curve the lower the plate resistance.

Experiment 3–11. From the curves plotted in Experiments 1–11 and 2–11 measure the slopes at various values of E_b and E_c and calculate the plate resistance, transconductance, and amplification factor. Plot these values against grid voltage for one curve and against plate voltage for another curve.

Note carefully that the "slope" of the characteristic curves relates only to the straight part—not the lower bend where the curvature, and hence the slope, is continually changing. In the

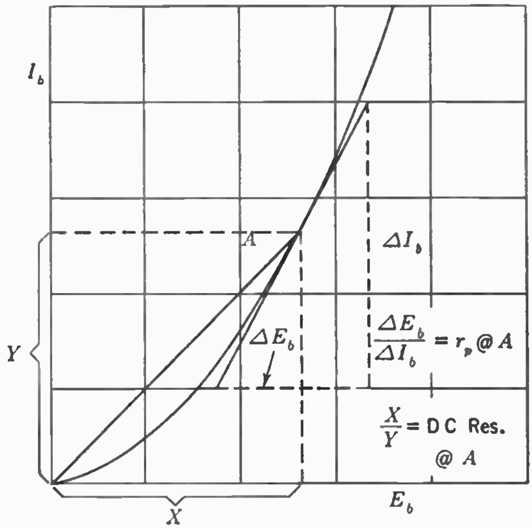


FIG. 9–11. Difference between d-c and a-c resistance of a tube. One (a-c) is the slope of the curve; the other (d-c) is merely the ratio between I_b and E_b.

problems above, it is assumed that the characteristic is straight (linear) over the region in which the values given are to be found.

The three important tube "constants" or "parameters" are determined by the methods indicated here or by means of a-c bridges.

CHAPTER 12

THE TUBE AS AN AMPLIFIER

The essential property of the triode (more complex tubes are really triodes with some modern improvements) is the ability to amplify voltages or currents or power. Small a-c voltages placed upon the grid produce higher a-c voltages in the plate circuit. How is this possible?

Basic triode circuit. In Fig. 1-12 is the basic amplifier circuit. The direct voltages on cathode, grid, and plate are fixed.

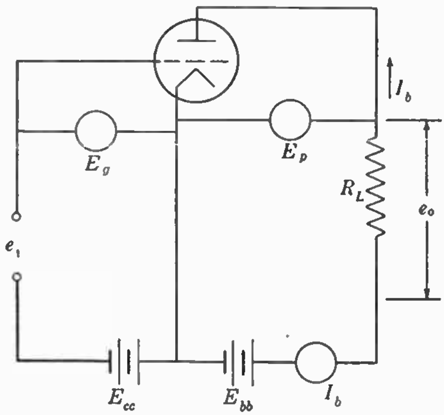


FIG. 1-12. Basic amplifier circuit with resistance load R_L.

In the plate circuit is a resistor; voltmeters and ammeters are placed at the points where voltages and currents are to be measured. Provision is made to apply some alternating voltage between the grid and cathode. An a-c voltmeter is placed across the plate resistor. Up to the moment when alternating voltages are applied to the grid, no alternating voltage will appear across the plate resistor (known as the **load**) because there is no alternating current through the load and therefore no a-c voltage drop across it.

When, however, alternating voltage is applied to the grid, alternating current flows in the plate circuit and an alternating voltage appears across the plate resistor. If conditions are properly chosen the value of the alternating voltage across the load will be greater than the alternating voltage applied to the grid. The tube then is amplifying.

For the moment let us simulate an alternating grid voltage by merely increasing and decreasing the grid voltage about its average or bias (E_c) value. For each value of grid voltage there will be a new plate current, and by plotting these values we shall obtain a curve like that in Fig. 2–12. Note that the tube has both direct and alternating voltages in the input and direct and alternating currents in the output circuit. The grid current is practically zero because the grid is always maintained at a potential negative with respect to the cathode and electrons will not be collected in any quantity by the grid. At some instants the grid is more negative than the bias voltage and at others less negative. The grid voltage, e_g, and plate current, i_p, rise and fall in unison, more current flowing when the grid is less negative and less current flowing when the grid is more negative. This is exactly what occurs when an alternating voltage is applied to the grid-cathode circuit.

The plate current can be analyzed in two ways: (1) it is made up of a pulsating current, increasing and decreasing above and below some median value in unison with the grid-voltage changes; (2) it is made up of two currents, (*a*) a d-c component whose value is fixed by the steady plate and grid voltages, and (*b*) an a-c component whose peak or rms value is determined by the circuit conditions.

Thus in Fig. 2–12 a d-c meter will read 6.6 ma and an a-c meter will read $1.4 \times 1/\sqrt{2}$ or 1.0 ma. If a 60-cycle a-c voltage is applied to the grid the a-c component of the plate current will have a frequency of 60 cycles.

The point *P* on the curve is called the **operating point.** It is determined by the fixed values of grid bias and plate voltage.

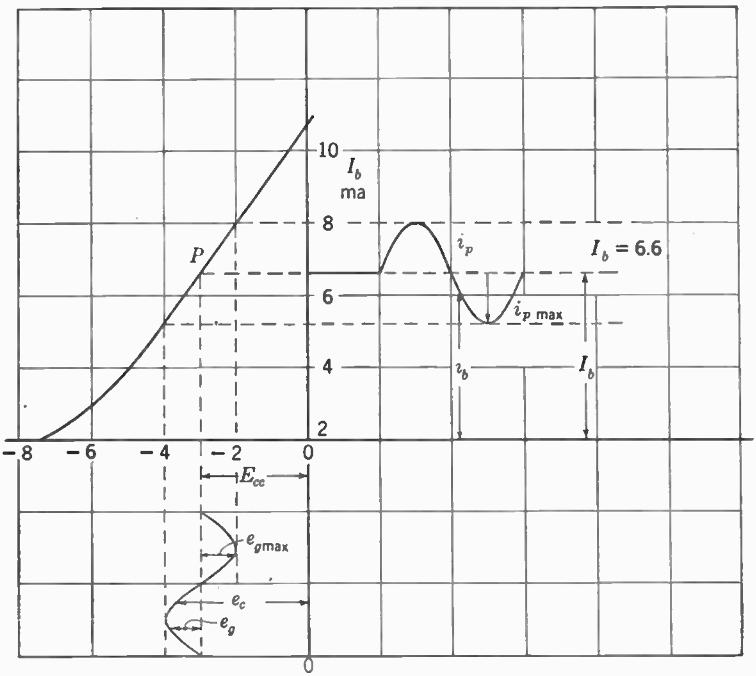


FIG. 2-12. Alternating voltages superimposed on the steady (bias) grid voltage produce a pulsating plate current which rises and falls from its steady (d-c) value.

Choice of operating point. Two conditions determine the placement of this point on the tube characteristic curve: (1) the amount of power the tube can safely dissipate, and (2) the amount of distortion of the input wave form that can be tolerated in the plate circuit. The power dissipated is controlled solely by the plate voltage and the plate current and is the product of these factors. The limit to this power is determined by the construction of the tube. The distortion factor will be considered later.

Tube voltages. All tube voltages are measured with reference to the cathode; that is, the grid bias is the voltage which exists between grid and cathode and the plate voltage is the voltage measured between plate and cathode. This is less than the battery voltage (E_{bb}) by the drop in the plate resistor load. With filament-type tubes operated from direct current, the most negative end of the filament is considered the reference point, and grid and plate voltages are measured with reference to this point.

The low-potential ends of the grid and plate circuits are connected to the cathode of an indirectly heated tube, to the negative terminal of a filament-type tube when operated from direct current, and to the center tap of the filament transformer winding or to the center of a resistance across the filament when operated from alternating current. These low potential leads are called the "common return" of the grid and plate circuits. Electrons flowing in the plate circuit return to the cathode through this lead.

Tube nomenclature. Since several voltages and currents are present in tube circuits, a problem in nomenclature arises. In

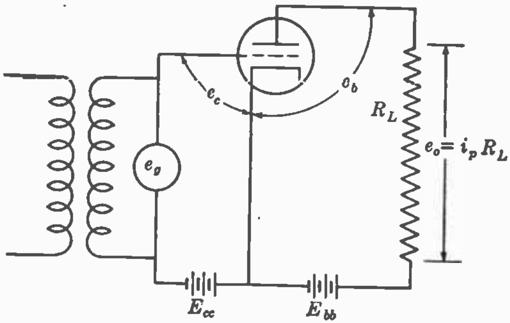


FIG. 3–12. Lower-case letters (e_g, i_p, etc.) indicate a-c values; capital letters (E_{cc}, E_{bb}, etc.) indicate d-c values.

general, lower-case letters indicate instantaneous values; capital letters indicate d-c or rms values. In Fig. 3–12 these several voltages and currents are indicated. The a-c values are the instantaneous values unless otherwise indicated. Thus e_g is the instantaneous alternating voltage applied to the grid; E_{cc} is the grid battery voltage; E_c is the actual direct voltage between grid and cathode; e_c is the instantaneous voltage measured from grid to cathode. This last value, e_c, is not the same as e_g; it is the algebraic sum of e_g and E_c. (See also Fig. 10–12.)

Resistance load. Figure 1–12 is the basic amplifier circuit. If conditions are properly adjusted the wave form of the alternating voltage measured across the resistor will be similar to that of the input alternating voltage, and the instantaneous values of this output voltage will be greater than the corresponding values of the input applied voltage. These conditions exist when (*a*) the C bias and the input a-c voltage are such that the alternating plate current has values which lie on the straight part of the characteristic curve only, and (*b*) the load resistance R_L is high compared to the tube resistance.

Let us look at the circuit in Fig. 1–12 rather critically. The plate voltage E_b as measured by a voltmeter connected from the plate to the cathode is no longer the voltage E_{bb} across the B battery. It is less than this value by the voltage drop in the resistor R_L, i.e., the I_bR_L voltage drop caused by the plate current. The voltage actually at the plate then is $E_b = E_{bb} - I_bR_L$.

If, for example, the B battery voltage equals 180 volts, R_L equals 100,000 ohms, I_b equals 0.5 ma, $I_b \times R_L$ equals $0.0005 \times 100{,}000$ or 50 volts, and E_b equals $180 - 50$ or 130 volts.

Now it can be seen that any variation in I_b causes a variation in the IR drop across R_L, and the plate voltage E_b must change accordingly. Any variation in the grid bias will cause a variation not only in the plate current of the tube but also in the plate voltage (E_b).

We must now have a new kind of characteristic curve to indicate the fact that the plate voltage is no longer constant but varies as the plate current varies. The difference between the plate battery voltage and the voltage actually measurable from plate to cathode depends upon the load resistor, the difference being greater the higher the value of the load resistance.

Dynamic characteristic curves. A series of "dynamic" curves is shown in Fig. 4–12. They were taken by placing resistances in series with the plate battery and power tube, maintaining the voltage on the plate equal to 180 when the C bias was −38.5

by increasing the plate battery voltage to compensate the loss in voltage across R_L. It will be noted that these curves are much flatter and longer than the static curves, indicating that plate-current variations are much smaller in magnitude under the same grid-voltage variations. The transconductance of the circuit is not as high as the value for the tube alone, but the slope

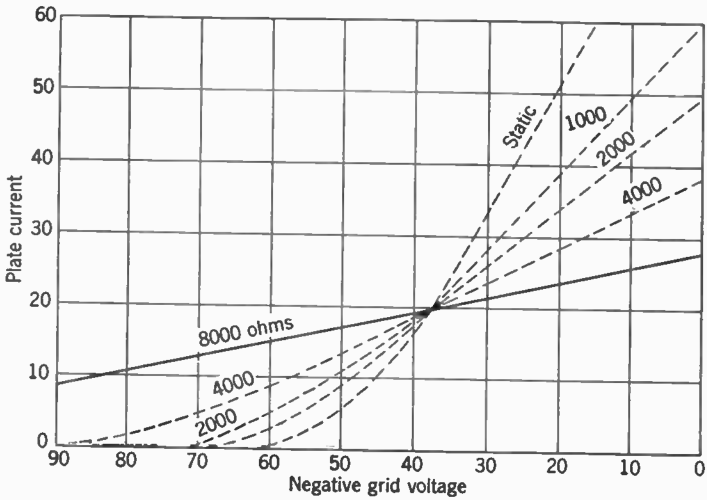


FIG. 4–12. Dynamic characteristic curves show plate current as a function of grid voltage when a load resistance is in the plate circuit.

of the curve shows the alternating current which will flow through the resistor when a given a-c voltage is applied to the grid, and these curves are therefore more useful than the static curves.

Experiment 1–12. Connect as in Fig. 1–12 a tube of the 6J5 or similar type, a plate battery, and a resistance. Short-circuit the terminals for e_i. Measure and plot the current in the plate circuit as the grid voltage is varied. Then change the resistor and repeat. Use values of 8,000, 16,000, 24,000, and 48,000 ohms.

Magnitude of the amplified voltage. There are several variable factors in a one-stage resistance amplifier, as shown in Fig. 1–12. The grid bias, E_{cc}, the plate battery voltage, E_{bb}, the load resistance R_L, all may be changed as desired. With a fixed value of E_{cc} and E_{bb}, the plate resistance of the tube r_p is fixed. If we measure the alternating voltage across a load resistance with a given value of alternating voltage on the grid we shall determine the voltage amplification of the circuit. Then, if we change the load resistance and again measure the output voltage across R_L, we shall get data for a curve which shows that the amplification increases as the load resistance increases, but which finally comes very near a certain fixed value numerically equal to the amplification factor of the tube. Increasing the load resistance beyond this point has little effect on the amplification.

Because each change in R_L will change the plate current, we must adjust the plate battery each time so that the voltage actually on the plate (E_{bb} minus the voltage drop along the load resistance) is the same.

The amplification that will be realized in a laboratory experiment of this kind may be calculated from the formula

$$\frac{\mu R_L}{R_L + r_p} = A \quad \text{(voltage amplification)} \tag{1}$$

and the alternating plate current from

$$\frac{\mu e_g}{R_L + r_p} = i_p \tag{2}$$

and the a-c voltage across the load from

$$\frac{\mu e_g R_L}{R_L + r_p} = A e_g \tag{3}$$

where e_g equals grid voltage.

The maximum voltage amplification takes place when the external resistance, R_L, is infinite. Practically, however, 75 per cent of the maximum amplification is obtained when R_L is three times as great as r_p. For example, if r_p is 10,000 ohms, a load resistance of 30,000 ohms will produce an amplification equal numerically to 75 per cent of the μ of the tube. The maximum possible value is the amplification factor (μ) of the tube. Thus, if a tube especially made for resistance-coupled amplifiers having an amplification factor of 30 is used, the maximum possible amplification will be 30, but under actual operating conditions the voltage amplification, A, of the complete circuit is about 20.

Power output. The a-c power in the load resistance is calculated as follows:

$$P = i_p{}^2 R_L$$

$$i_p = \frac{\mu e_g}{R_L + r_p}$$

$$P = \frac{(\mu e_g)^2 R_L}{R_L + r_p{}^2}$$

This is a maximum when the two resistances are equal, that is, when

$$r_p = R_L$$

Then the power

$$P = \frac{\mu^2 e_g{}^2}{4 r_p}$$

where e_g = rms input voltage; and

$$P = \frac{\mu^2 e_g{}^2}{8 r_p}$$

where e_g = peak or maximum voltage.

This power, which is fed into the load resistance, must come from the plate battery, because the tube itself generates no power—it acts merely as a transformer or valve which takes small voltages and currents on its input and turns out larger voltages and currents to its output circuit. The power in the input circuit must come from the circuit to which the tube is attached. The power in the output must come from the batteries. The tube therefore releases power from the batteries in a form which is an exact replica of the power utilized in the input circuit to which the tube is attached. The tube itself consumes power from the battery. This power heats the plate and is wasted. It is equal to $I_b E_b$. Note that the power output is proportional to the square of the input voltage. Doubling the input voltage increases the output power four times.

Power amplification. Ordinarily the grid circuit takes no current from the input. If, however, the grid voltages are fed through a resistance so that the voltage drop across the resistance represents the applied input, then some power is consumed in this resistor. The power output divided by this input power represents the power amplification.

Example 1–12. Four milliampheres of current at 1000 cycles is fed through a 10,000-ohm resistance in series with the grid of a tube and its C battery. In the output is a 2000-ohm resistor. The amplification factor of the tube is 3; its internal plate resistance is 2000 ohms. What is the power output, what is the voltage across the output, what is the voltage and power amplification, and what power is lost on the internal resistance of the tube?

Let the input voltage $e_i = i_i R_i = 0.004 \times 10{,}000 = 40$ volts rms

$$\text{voltage amplification } A = \frac{\mu R_L}{R_L + r_p}$$

$$= \frac{3 \times 2000}{4000} = 1.5 \text{ times}$$

The output voltage $e_L = e_i \times A = 40 \times 1.5 = 60$ volts rms.
Power output

$$P_o = \frac{\mu^2 e_i^2}{4r_p}$$

$$= \frac{9 \times 40^2}{4 \times 2000} = 1.80 \text{ watts}$$

Power input, P_i,

$$i_i^2 R_i = (0.004)^2 \times 10{,}000$$

$$= 16 \times 10^{-6} \times 10{,}000$$

$$= 0.16 \text{ watt}$$

Power amplification $= P_o \div P_i = 1.8/0.16 = 11.25$ times.

Since i_p flows through both R_L and r_p and since $r_p = R_L$, a-c power lost in r_p = power lost in R_L = 1.8 watts.

Total power taken from batteries = 3.6 watts.

$$\text{Efficiency} = \frac{\text{Useful power}}{\text{Total power}} = \frac{1.8}{3.6} = 50 \text{ per cent}$$

Problem 1–12. Assume a tube with a load resistance of 12,000 ohms and an a-c grid voltage of 3 volts (peak). The μ of the tube is 8; its r_p is 12,000 ohms. What is the peak value of the alternating plate current? (Use formula 2.) What is the rms value? What a-c power is developed in the load resistance? What voltage appears there?

Problem 2–12. What is the voltage amplification in the above circuit? (Use formula 1.) How much would it be increased if the load resistance were increased to 36,000 ohms? How much would this change the power output?

Problem 3–12. The plate resistance of a certain "high-μ" tube is 60,000 ohms, and its amplification factor is 20. What load resistor value in ohms must be used to obtain a voltage amplification of 15?

Problem 4–12. A certain tube has an internal plate resistance of 5000 ohms, its amplification factor is 8; it is worked into a load of 5000 ohms. Plot the output power in the load resistance as the input a-c voltage is increased from 1 to 8 volts (peak).

Problem 5–12. What is the voltage amplification in Problem 4–12?

Problem 6–12. The power output of a tube when worked into a load whose resistance is equal to the tube resistance (when the input grid voltages are peak volts) may be written as $\frac{\mu^2}{8r_p} \times e_g^2$. If we divide this expression by e_g^2 we shall have a figure which gives us the power output in watts per (volt input)^2. Make a table of such values for representative tubes used at the present time, getting the data from manufacturers' information.

Problem 7–12. The normal C bias for a certain tube is 9 volts. What is the largest rms voltage that may be applied to its grid before the grid goes positive?

Amplifier overloading. The conditions for undistorted amplification are: (*a*) the C bias and magnitude of the a-c input voltage must be such that only the straight part of the tube characteristic curve is employed, and (*b*) the load resistance R_L must be large with respect to the internal resistance of the tube r_p. Let us examine these conditions. Suppose first that the C bias is too great, as in Fig. 5–12, so that the operating point is down on the curve of the characteristic. Suppose that, at the start, there is no alternating voltage on the grid. The plate current I_b will be as indicated on the drawing. Now let the a-c voltage be impressed on the grid. The instantaneous plate current i_p will increase when the effective negative bias is decreased by the superposition of the positive half of the input a-c wave. Similarly, i_p will decrease from the steady no-input-signal value as the grid bias is increased by having added to the steady C bias the negative half of the a-c input wave. Note, however, that, at the most negative portion of the input voltage cycle, no current at all flows in the plate circuit. The plate current has been cut off.

The current flows longer on the positive portion of the cycle than on the negative part, and the average value will be some-

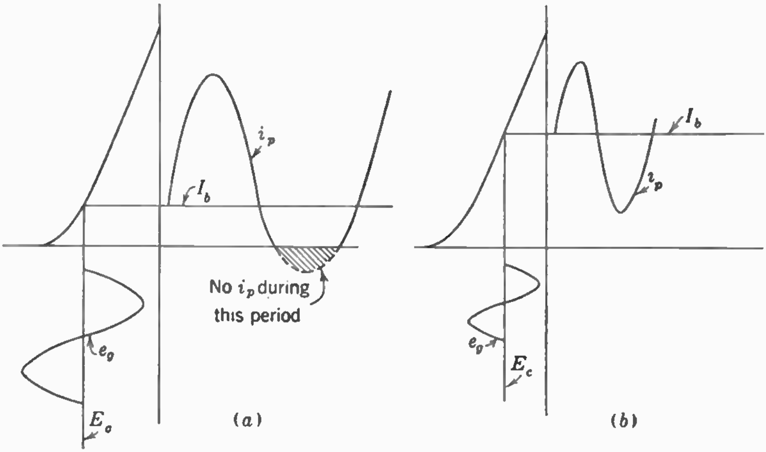


FIG. 5–12. Wrong bias (*a*) produces an output current and voltage that is not like the input voltage. In (*b*) the bias is correct.

thing between zero and the maximum value attained. This will be greater than the current which flows when no input voltage is applied. A meter in the plate circuit will show an increased current when a-c voltage is applied to the grid under these conditions compared with the current that flows without an input signal (or without *excitation*, as the engineers say).

This is a sure sign of distortion—an increase or decrease in plate current as read by a d-c meter when the grid is excited. A glance at the plate-current curve in Fig. 5–12 is all that is necessary to prove that distortion is present: the plate-current wave does not resemble the grid-voltage exciting wave. In this case the remedy is to decrease the C bias, permitting more no-signal plate current to flow. (This discussion is true of class A amplifiers only. In some amplifiers this type of distortion is intentional.)

Distortion may be caused if the C bias is correct, say at the center of the E_c-I_b curve, but if the input a-c voltage is so great that the grid is still forced so negative, at times, that no instantaneous plate current flows.

The C bias may be too low so that, on the positive half cycles of input voltage, the grid is forced positive. The remedy is either to reduce the voltage applied as excitation or to increase the C bias. If the C bias is increased, the negative part of the input voltage may force the operating point down onto a curved part of the characteristic and distortion again occurs. The remedy now is to increase the plate voltage. This may cause such a large steady plate current to flow that the plate of the tube overheats.

Thus there are a number of conditions that must be met in designing and operating an amplifier correctly.

Distortion due to curved characteristic. When a load is in the plate circuit the tube characteristic becomes flatter and its straight part longer, but at the lower part of the curve there is still a bend and throughout the E_c-I_b graph there is considerable curvature. Large input voltages, therefore, may cause distortion because of these curved parts of the characteristic.

Although the maximum power output is obtained from a tube when the load resistance is equal to the internal tube resistance, the maximum *undistorted* power output is attained when the load resistance, R_L, is twice as great as the tube resistance, r_p. Some power is sacrificed under these conditions but, as the curve in Fig. 6–12 shows, the loss is not great. If the load resistance is lower than the tube resistance distortion due to curvature of the characteristic is certain to occur. Even with a load equal to twice the tube resistance there is curvature, and input voltages high enough to drive the plate current too low must not be applied. The power output when R_L equals $2r_p$ is

$$P = \frac{2}{9} \times \frac{\mu^2 e_g{}^2}{r_p}$$

where e_g is rms input voltage.

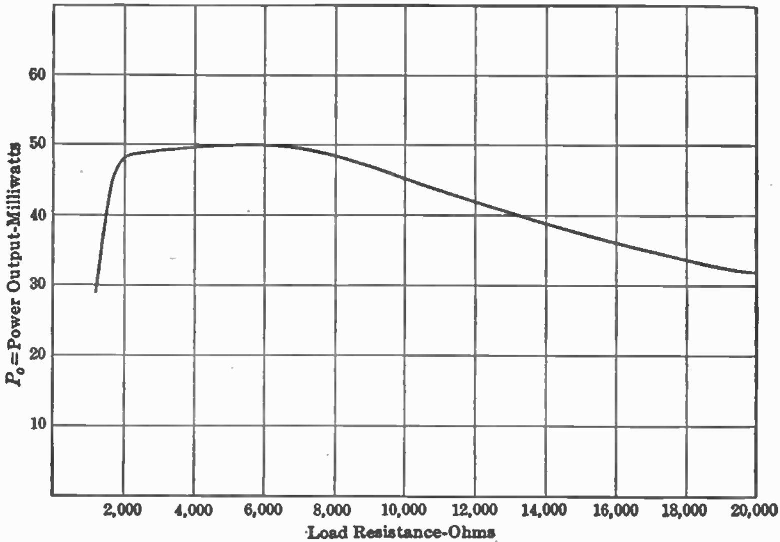


FIG. 6–12. Relation between load resistance, 5000 ohms, and power output. Maximum power is transferred from tube to load when load and tube resistances are equal.

The load line. The amount of power output, the amount of distortion created, and other important factors can be determined from the routine known as plotting the load line. To do this a family of E_b-I_b curves is necessary. The value of the load resistance to be used must be known.

The load line gives the division of voltage across the load and the tube for all values of plate current as produced by the varying exciting grid voltages. There are several ways of plotting this load line, but the simplest is as follows.

All the voltage of the plate battery is divided between the tube internal resistance, r_p, and the load resistance, R_L. If the tube is considered a short circuit, then the entire voltage appears across the load and the current flowing is simply the voltage divided by the resistance as given by Ohm's law. This value of current, which is the maximum value the plate current can ever attain in the given circuit, constitutes one point of the load line which is a straight line. The other point is found by this consideration: If no current flows in the plate circuit, there is no voltage drop across the load, R_L, and therefore the total battery voltage appears across the tube. This is the maximum value of plate voltage that can occur.

In Fig. 7–12 is a family of curves for a typical tube. With a load of 8000 ohms, a battery voltage of 552 volts, one point is

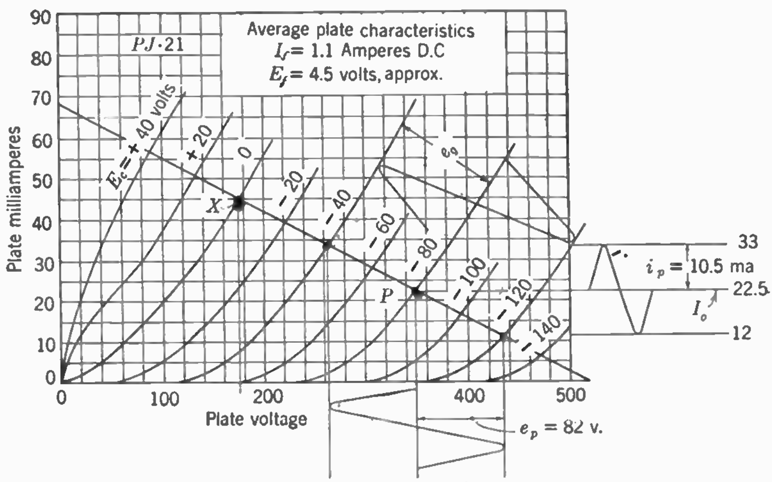


FIG. 7–12. The straight line between I_b = 69 ma and E_b = 552 volts is the "load line." The value of I_b for any combination of E_b and E_c for a load resistance of 8000 ohms may be found from it. The slope of the line $E_b \div I_b$ = 8000 ohms.

given directly, 552 volts and zero plate current. The other point of the load line is given by the maximum current that would flow if the tube were short-circuited. This maximum current is given by Ohm's law, $I = E \div R$, or $552 \div 8000 = 69$ milliamperes. The load line is drawn between these two points.

Any point on the line indicates the plate current for the corresponding value of plate voltage. Thus in Fig. 7–12 with 265 volts actually between plate and cathode the plate current will be 33 ma, and the corresponding grid bias is −40. If the point represented by these conditions is chosen as the operating point, this value of plate current (33 ma) will flow and will not change when the tube is "excited" (that is, when an alternating voltage is applied) unless distortion is present.

Now what can be learned from this procedure?

At any given plate voltage (E_b) [that is, battery voltage (E_{bb}) minus the drop in the resistance (I_bR_L)] the plate current will be given by this line. The actual voltage on the plate at any value of plate current will be given by the distance along the plate voltage axis to the right of the zero voltage point. Thus if the grid is so excited that a plate current of 16 ma flows, the plate voltage will be 400 and, if the grid is driven to zero voltage, the plate current will be 45 ma.

The difference between the battery voltage and the plate voltage at any instant represents the voltage appearing across the load resistance R_L.

Now suppose that we impress upon the grid an alternating voltage of 40 volts peak value. The grid voltage actually on the tube, therefore, will vary 40 volts up and down from the −80-volt value, or from −40 to −120. The instantaneous plate current will vary from 33 to 12 ma, with a peak value of 10.5 ma. The instantaneous plate voltage varies up and down from the 350-volt value, with the upper and lower limits given by the values of plate voltage at which the load line crosses the grid voltage lines −120 and −40. These limits in plate voltage are 432 and 262, and the peak value of the alternating plate voltage is approximately 82 volts.

The power into the load resistance is the useful tube output. It may be calculated from the formula:

$$P = \tfrac{1}{8}[(e_{p_{\max}} - e_{p_{\min}}) \times (i_{p_{\max}} - i_{p_{\min}})]$$

$$= \tfrac{1}{8}(432 - 262) \times (0.033 - 0.12)$$

$$= 0.445 \text{ watt or } 445 \text{ mw}$$

This is the a-c power conveyed to the load by the tube. The steady d-c power lost in the load resistance whether a-c voltage is fed to the grid or not, is, of course, $I_bR_L = (0.0225)^2 \times 8000 = 4$ watts.

Harmonic distortion calculation. The amount of distortion produced in such an amplifier may be calculated from this load-line curve. The distortion is produced when the non-linear parts of the tube characteristic curve are used, when the load resistance is wrongly calculated (and used) for the tube in question, and when the a-c input to the grid is incorrect with respect to the steady grid bias. The distortion increases as the input excitation to the grid (a-c voltage) increases, and it is very low for low inputs and may become quite evident to the ear if the grid is over-excited.

The second harmonic distortion may be calculated from:

$$\frac{\dfrac{i_{p_{\max}} + i_{p_{\min}}}{2} - I_b}{i_{p_{\max}} - i_{p_{\min}}} \times 100\%$$

where I_b is the steady plate current when no signal is applied to the grid. In Fig. 7–12 where the excitation is low compared with the grid bias, distortion will be very low. Suppose, however, that the grid is excited with an a-c voltage whose peak value is equal to the grid bias, −80 volts. Substitute in the above expression the values from Fig. 7–12.

$$\frac{\dfrac{0.045 + 0.0065}{2} - 0.0225}{0.045 - 0.0065} \times 100\% = \frac{0.02575 - 0.0225}{0.0385}$$

$$= \frac{0.00325}{0.0385} = \frac{3}{38.5} = 0.078 \times 100 = 7.8\%$$

This means that, for every signal put on the grid having a peak value of 80 volts, there will be in the output a second harmonic of this input signal frequency equal to 7.8 per cent of the value of the fundamental. Since it is generally considered that 5 per cent second harmonic distortion is readily audible to the average listener, this amount of distortion is too great for high-quality reproduction. The solution is to reduce the input voltage, to increase the plate voltage, to increase the load resistance, or to try some combination of these possibilities.

Problem 8–12. Let the reader choose other load lines, such as one representing a low resistance, say 4000 ohms as in Fig. 7–12, and calculate the power output and percentage second-harmonic distortion for several values of input voltage. A curve might be plotted showing percentage distortion as a function of the load resistance for a given input a-c voltage, or as a function of input voltage with a constant load resistance.

Note on Fig. 7–12 that an alternating voltage e_g is represented as being impressed on the tube. This has a peak value of 40 volts and produces a plate-current variation of $i_p = 10.5$ ma peak. These current variations produce plate-voltage variations (bottom of diagram) of 82 volts peak.

If the plate-current variations are equal above and below the 22.5-ma line, then no distortion is being created; but, if the bias point or any other condition is improperly chosen, the plate current may rise more above 22.5 ma than it decreases below this value. Now the plate-current curve will not have the same form as the grid-voltage curve, and the output is not a replica of the input. The principal component of the distortion is the second harmonic of the fundamental, that is, a new alternating current of twice the frequency of the original.

Power diagrams. To review the tube with a resistance load as a power amplifier, and to gather a few additional facts, let us draw the E_b-I_b curves as shown in Fig. 8–12, which is a purely theoretical case—the curves and values of current and voltage represent no particular tube now available; they were chosen at random. Let us assume that the voltage actually on the plate is 160 volts when the steady plate current is 20 ma. That is, $E_b = 160$, and $I_b = 20$ ma. Let us assume a load of 5000 ohms, and draw the load line through the operating point and the zero-plate-current point; we find that the bias on the grid to maintain such a plate current is 20 volts. Let us assume that on the grid is put a sine wave of voltage with a maximum value of 10, so that the actual grid voltage varies between -10 and -30 volts. The corresponding I_b variations are from 30 to 10 ma.

We note that the load line crosses the battery voltage line at about 260 volts. This, then, is the battery voltage necessary to insure that the plate of the tube gets its 160 volts. In other words, there is a drop of 100 volts (5000 × 0.02) in the resistance of the load. Under normal conditions of no a-c input to the

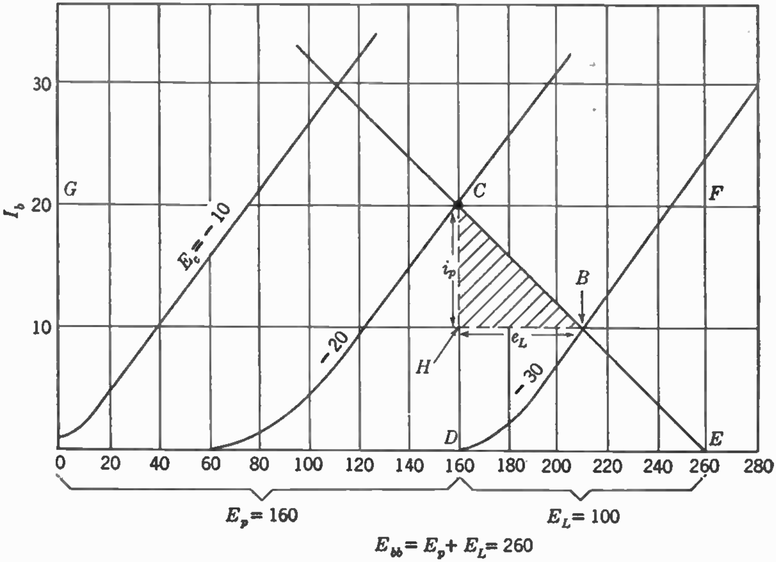


FIG. 8–12. Power diagram for an amplifier with resistance load of 5000 ohms.

grid, the voltage across the tube is 160 volts, across the load is 100 volts, and the plate current is 20 ma.

Now the d-c power lost in the load is the product of the voltage across the resistance and the current through it, and the power lost on the plate of the tube is the product of the voltage across the tube and the current flowing. Thus,

$$\text{Power lost in load} = E_L \times I_b = DE \times CD$$

$$\text{Power lost in tube} = E_p \times I_b = OD \times CD$$

and the total power supplied by the battery must be the sum of these powers, that is

$$\text{Power from battery} = E_{bb} \times I_b = OE \times CD$$

When a 10-volt a-c input is applied to the tube so that the grid operates about its mean or average value of -20 volts, the peak value of the alternating current through the load is given by the line HC and the peak value of the a-c voltage across the load is given by the line HB. Let us call these values of current and voltage

$$e_L = \text{peak a-c voltage across load} = HB$$

$$i_p = \text{peak alternating current through load} = HC$$

Since a-c power in a resistance circuit is the product of the rms current and the voltage, to obtain the a-c power in the load we must first get the rms values of the above values and then multiply them

$$e_{L_{\text{rms}}} = e_{L_{\text{max}}} \times \frac{1}{\sqrt{2}}$$

$$i_{p\ \text{rms}} = i_{p_{\text{max}}} \times \frac{1}{\sqrt{2}}$$

whence a-c power in load $= e_L i_p$ (rms)

$$= e_L \times i_p \times \frac{1}{\sqrt{2}} \times \frac{1}{\sqrt{2}}$$

$$= \frac{e_L \times i_p}{2}$$

$$= \frac{HB \times HC}{2} = \text{area of triangle } HCB$$

This power is dissipated in the load resistance in addition to the d-c power lost there due to the voltage drop across it and the current through it. Since the average current drawn from the battery has not changed, the power taken from the battery has not changed—and yet the load has an additional amount of power used up in it. Where does this power come from? Clearly it must come from the power used up on the plate of the tube. *When an a-c voltage is placed on the grid, then a-c power is* *developed in the load, less power is wasted on the tube plate, and the tube will actually run cooler when it is delivering power to the load than when standing idle, that is, with no a-c input grid voltage.*

Let us see what these values of power are. We can take them directly from the graph in Fig. 8–12.

D-c power lost in load (no a-c grid voltage)

$$= DE \times CD = (260 - 160) \times (20 \text{ ma}) = 100 \times 0.02 = 2.0 \text{ watts}$$

D-c power lost in tube (no a-c grid voltage)

$$= OD \times CD = (160) \times (20 \text{ ma}) = 160 \times 0.02 = 3.2 \text{ watts}$$

D-c power taken from battery

$$= OE \times CD = 260 \times 0.02 = 5.2 \text{ watts}$$

Max a-c voltage across load (a-c grid voltage = 10)

$$= HB = 210 - 160 = 50 \text{ volts}$$

Max alternating current through load (a-c grid voltage = 10)

$$= HC = 10 \text{ ma} = 0.01 \text{ amp}$$

A-c power in load (a-c grid voltage = 10)

$$\tfrac{1}{2} \times HB \times HC = 0.5 \times 50 \times 0.01 = 0.25 \text{ watt} = 250 \text{ mw}$$

which is subtracted from the d-c power in tube.

The formula for the power output of a tube is

$$\text{Power output} = \frac{\mu^2 e_g{}^2 R_L}{(r_p + R_L)^2}$$

and to check the above calculations we must find the tube constants from the curves in Fig. 8–12. The plate resistance of the tube is simply the reciprocal of the slope of the E_b-I_b line, and noting that a change of 120 − 40 volts on the $E_g = -10$ line produces a change in plate current of 32 to 10 ma we calculate

$$r_p = \frac{120 - 40}{0.032 - 0.010} = \frac{80}{0.022} = 3630 \text{ ohms}$$

From the two curves, $E_c = -10$ and $E_c = -20$ we ascertain that a change of 10 volts on the grid produces a change in plate current of 32 to 10 ma when the plate voltage is 120. From these facts we calculate

$$g_m = \frac{0.032 - 0.010}{10} = 2200 \text{ micromhos}$$

And, since

$$\mu = r_p \times g_m = 3630 \times 2200 \times 10^{-6}$$

$$\mu = 8$$

whence power output

$$= \frac{(8 \times 7.07)^2 \times 5000}{(5000 + 3630)^2}$$

$$= 0.262 \text{ watt} = 262 \text{ mw}$$

This value agrees closely enough with our data secured from the characteristic curves in Fig. 8–12.

From a collection of the E_b-I_b curves of any tube we may obtain all the necessary data upon which to build an amplifier and calculate its power output, the losses in the various tubes, the percentage distortion, etc.

Equivalent tube circuit. Since a change in plate voltage may be replaced by a smaller change in grid voltage multiplied by the mu of the tube, we may replace the entire tube by a fictitious generator whose voltage is μE_g and whose internal resistance is equal to the resistance of the tube. In fact in all problems the tube is so considered, and is indicated symbolically as in Fig. 9–12.

Problem 9–12. With a 20-volt bias the plate current of a tube under a given value of plate voltage is 55 ma, and when the bias is increased to 30 volts the plate current is reduced to 28 ma. If, however, at this value of grid bias (−30) the plate voltage is increased from 180 to 210 volts the plate current comes back to its original value, 55 ma. What is the amplification factor of the tube?

Problem 10–12. The amplification factor of a tube is 8, and when the grid is −3 volts the plate current is 3 ma. If the bias is reduced to zero the current increases to 8 ma. Both these current values were read when the plate voltage was 90 volts. How much would the plate voltage have

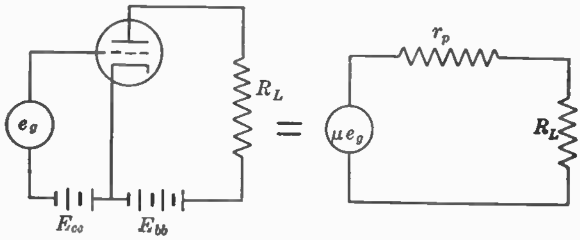


FIG. 9–12. A tube and its load may be replaced for purposes of calculations by a generator with an output of μe_g and a resistance of r_p.

to be reduced (at zero grid bias) to bring back the current to its 3-ma value?

Problem 11–12. Changing the plate voltage of a power tube from 100 to 300 volts changes the plate current from 10 to 55 ma. If the bias is zero at 55 ma what must be done to it to reduce the current to 10 ma if the amplification factor is 7.8?

Phase of e_g, e_p, and i_p. When the grid voltage increases in a positive direction, the plate current increases. When this happens a greater voltage drop occurs along the load resistance and less voltage will be measured from plate to cathode. Thus the plate voltage decreases when the grid voltage increases in a positive direction. These two voltages, e_g and e_p, are out of phase.

The end of the plate load resistor farthest from the plate battery is negative with respect to the end connected to the battery. The negative end of the load resistor becomes more negative the greater the current drawn through it from the plate battery.

Therefore, as the grid voltage increases in a positive direction, the voltage across the load increases in a negative direction. This amounts to a phase difference of 180 degrees between the grid voltage and the output voltage.

If, therefore, a sine wave or square wave which starts from zero and rises in a *positive* direction is applied to the grid, a sine wave or a square wave appears across the plate load which starts from zero and rises in a *negative* direction. A negative grid pulse, on the contrary, produces a positive pulse in the plate circuit.

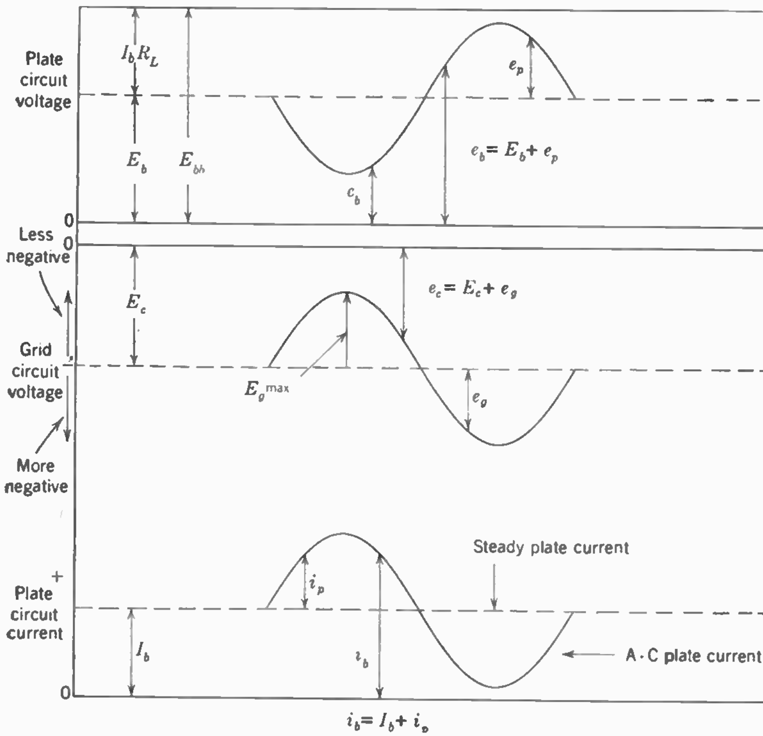


Fig. 10–12. Phase of grid, plate, and load voltage. As e_g becomes more negative e_p becomes less negative. The load voltage, e_L, also differs from the grid voltage by 180 degrees.

Screen-grid tubes. The grid and plate represent two metallic conductors more or less parallel to each other like the plates of a condenser. Furthermore, capacitance exists between wires connected to the electrodes and to the tube prongs. Capacitance exists between the prongs themselves. The total capacitance existing between grid and plate is appreciable and may have quite a low reactance if the frequency of operation is sufficiently high. Current will be conducted from a point of high potential (plate) to a point of low potential (grid) through this reactance. The voltage produced by this current will add to or subtract from the grid voltage, depending upon the phase of the voltage thus "fed back" from output to input.

If the amplification of the tube and its circuit is 10, and if one-tenth of the output voltage gets back to the grid in some way, as much voltage may be impressed on the grid from the output as was placed there by the driving circuit. This additional voltage is amplified in the tube and is again impressed upon the input. Soon the tube and circuit will break into uncontrollable oscillation and the tube becomes worthless as an amplifier. The grid loses control so far as amplifying and reproducing the original excitation is concerned.

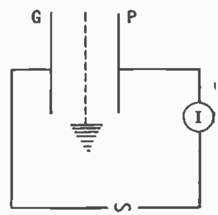


FIG. 11–12. A grounded shield between plate and grid reduces current flow from plate to grid.

If now, another grid is placed between the control grid and the plate and is maintained at zero a-c potential by connecting a condenser between it and ground, the capacitance between grid and plate is reduced. The feedback from plate circuit to grid circuit is reduced and much greater amplification without instability is possible. Actually the screen grid is maintained at a positive d-c potential and so tends to neutralize the space charge and accelerate the electrons.

Some electrons strike the screen and form a screen current, but the greater proportion get through to the plate. Since the screen current subtracts from the total possible plate current the efficiency is impaired, but by proper design the effect of the grid upon the plate current is as great as if the screen were not present.

Not only does the screen protect the grid from reaction from the output circuit but it produces a further change in the tube characteristic. Since the plate is now shielded from the grid, it follows that it is shielded to some extent from the cathode. Plate voltage, therefore, is not so effective in controlling plate current as it is in a triode. On the other hand, the effect of the grid is not impaired (except by those few electrons which are lost on the way to the plate). The ratio of the effectiveness of the grid voltage to plate voltage is increased, and the amplification factor increases proportionately. Screen-grid tubes may have amplification factors as high as several hundred.

Since the plate resistance of the tube is the ratio of a change in plate voltage to the corresponding change in plate current, it follows that the plate resistance may become very large since large changes in plate voltage produce only small changes in plate current.

The screen-grid tube (tetrode), therefore, has a high dynamic plate resistance (1 megohm), high amplification factor, and

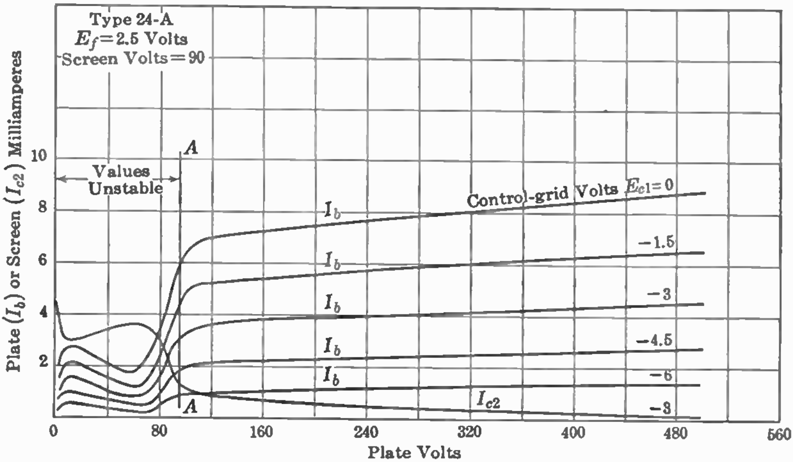


FIG. 12–12. Characteristic curves of typical screen-grid tube.

average mutual conductance. It is very useful in voltage amplifiers, but it must be employed with care so that no output voltage gets back to the input except as is desired for one purpose or another.

In circuit operation the screen is connected to a positive potential, and the alternating voltage components in the screen circuit are by-passed to the cathode by a condenser connected close to the screen terminal.

Characteristic curves of the screen-grid tube are shown in Fig. 12–12. Note the fairly flat portion of the curves to the right of the plate voltage = 90 line. This is the useful region of the characteristic. To the left of this 90-volt line something very interesting happens. Over a small region the plate current actually decreases as the plate voltage is increased. How does this happen?

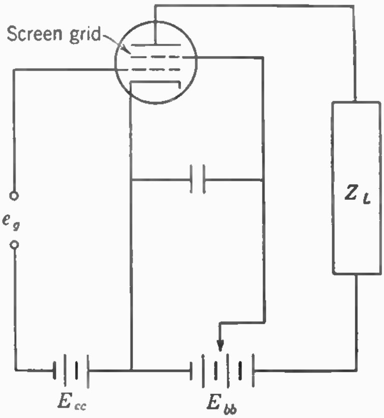


FIG. 13–12. Screen grid has a positive d-c potential with respect to cathode and is by-passed to cathode to provide low-impedance path for alternating currents.

Consider the screen grid biased at +90 volts. An electron leaves the cathode and is impelled onward by this positive voltage on the screen. It gets through the mesh of the screen and may strike the plate. If the plate is positive enough, say 10 volts or so, the velocity of the electron as it hits the plate may be high enough to knock an electron out of the plate. Since the screen is at a higher positive potential than the plate, it has greater attraction for the electron than the plate, and the electron will naturally turn around and go back to the screen. This new electron is called a secondary electron, and the emission of electrons under these conditions is known as **secondary emission.**

If the plate voltage is raised higher than that of the screen, the electrons will prefer to go to the plate and the characteristic curve straightens out like an ordinary law-abiding tube. When tubes are operated with a-c input voltages, the actual voltage on the plate varies from instant to instant, under some circumstances going from zero to near the battery voltage impressed upon the plate-cathode path. A circuit operated so that the plate voltage might be lower than the screen voltage would be unstable, and, therefore, the useful operating region of the E_b-I_b characteristic is limited. This means that the voltage output and power output of the tube are less than if this unstable region of the characteristic did not exist.

Pentode tubes. How can this unwanted falling-plate-current region be eliminated? By simply placing another shield or screen in the tube, this one being put between the screen grid and the plate. It is connected to the cathode and therefore is negative with respect to the plate. Now an electron may strike the plate hard enough to knock some electrons out of it, but these electrons find the zero-potential screen (negative compared with the plate) an insuperable barrier and cannot go through it to get back to the screen grid. They are turned around and fall back into the field of the plate. This third grid is called **a suppressor grid** because it suppresses secondary emission. Note the curves in Fig. 14–12 and compare them to those of the screen-grid tube of Fig. 12–12. The tube itself has become a five-element tube, or pentode, and as a class the pentode has taken

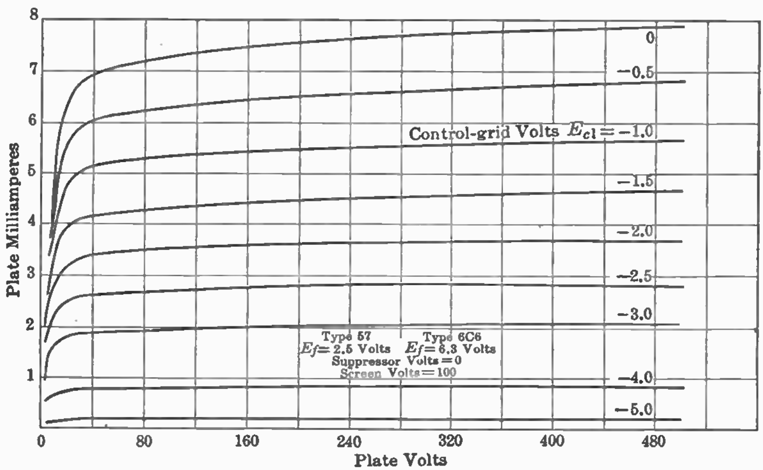


FIG. 14–12. Characteristics of a typical pentode tube used for voltage amplification.

over functions formerly performed by screen-grid tubes or tetrodes. Pentode power amplifiers such as the 6F6 produce a considerable output with low grid excitation voltages; voltage amplifiers like the 6J7 produce very high amplification compared with that possible in triodes.

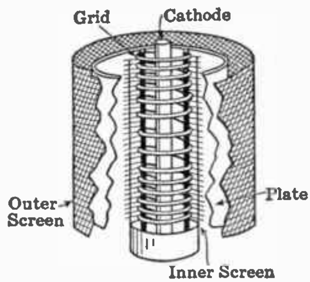


Fig. 15–12. Note how the grid winding has variable pitch from top to bottom.

Variable-mu tubes. Early receivers utilizing screen-grid tubes suffered from many troubles. Among others was the production of cross modulation ("cross talk") by strong undesired signals. These tubes had such a short range of grid voltage over which they could work that a strong signal would force the grid voltage to the point where the alternating plate current would cut off. This resulted in severe distortion, evidenced by blurbs or gasps of undesired modulation crashing through the desired program, or by hum entering from the power supply.

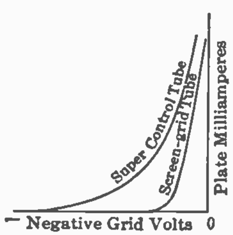


Fig. 16–12. Difference between variable-mu and sharp cut-off tube. Variable-mu tubes are said to have a "remote" cut-off.

Late in 1930 a new screen-grid tube made its appearance and became quite important in the following year. This was the variable-mu or **super control** tube. It had a very long, fairly flat characteristic at the bottom of its E_c-I_b curve. In other words, strong negative voltages on the grid did not force the plate current to zero, and cross modulation was prevented.

Because of the very long even characteristic, the tube was nicely adapted for automatic volume control, for i-f and r-f amplifiers, and for recording signal strength, or loud-speaker output, or other measurements where large ranges of current or voltage, etc., were to be measured.

In this tube the grid is wound, not regularly or evenly along the length of the winding space, but with a wide spacing at one portion and close spacing at another. The E_c-I_b characteristic may be controlled in manufacture by varying the pitch of the winding or the manner in which the wide and close spacing is put on the grid-wire supports.

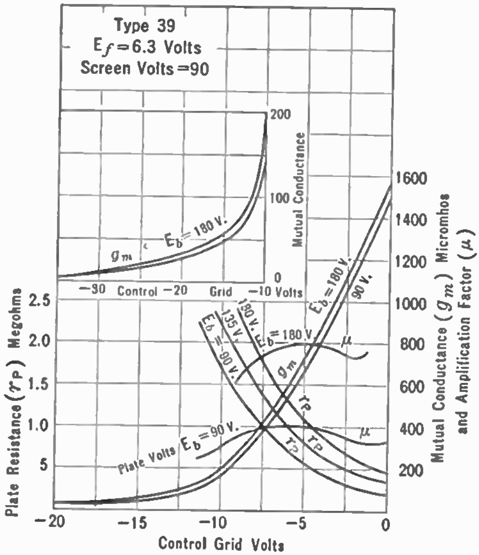


FIG. 17–12. Characteristics of variable-mu or super control tube.

The super control tube is also known as a "remote cut-off" tube.

Beam power tubes. Although the pentode is an advance over the screen-grid tetrode, there is need for improvement. At the lower plate voltages, there is still enough curvature to the characteristic to produce distortion unless the tube is limited severely in its output. It would be an advantage to straighten out the E_b-I_b curves. This is done in the beam power tubes in a most ingenious way.

Instead of placing an actual gridlike electrode within the tube structure, electrons are forced to travel in prescribed paths

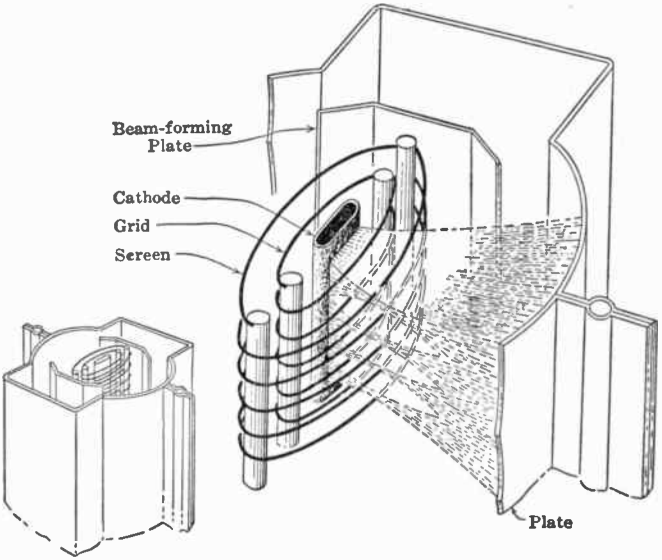


FIG. 18–12. Construction of beam power tube showing beam effect on electron flow.

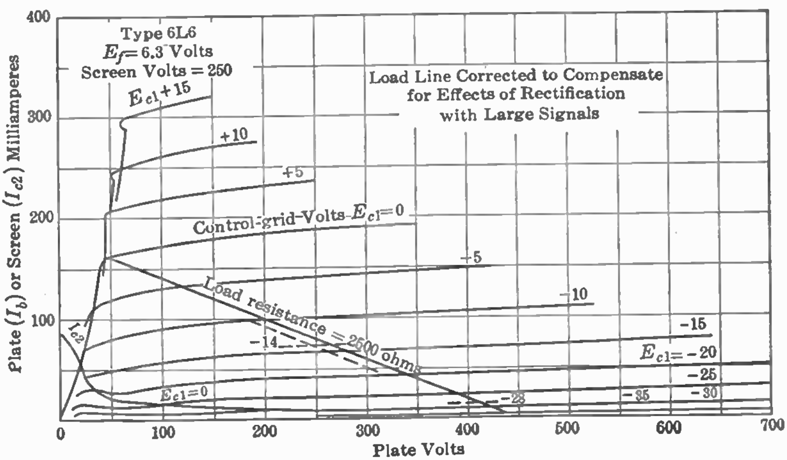


FIG. 19–12. Plate voltage-plate current characteristics of beam-type power tube.

and with such density that a "virtual" low-voltage electron-repelling screen (called a suppressor) is in the desired region. The screen, therefore, which repels secondary electrons back to the plate does not exist in the form of wires, but rather in the form of a collection of other electrons. The second grid, corresponding to a screen-grid tube, is spiral-wound like the control grid, and each turn is shaded from the cathode by a turn of the control grid. For this reason electrons do not often hit the screen and therefore very little screen current flows. The beam power tube is very efficient since so little loss of current exists in the screen circuit, and it has very high sensitivity, which means that a large amount of power is delivered for a small input grid voltage.

Grid nomenclature. In tubes having several grids, it is customary to speak of them according to their position from the cathode. Thus the grid nearest the cathode is the No. 1 grid, the next grid is called the No. 2 grid, and so on. Grid voltages are labeled accordingly, E_{c1} being the bias on the grid nearest the cathode.

Power output for pentode and beam power tubes. The general procedure for obtaining power output and distortion from pentodes and beam power tubes is much the same as for triodes just described. The calculations are made from a special E_b-I_b family. In Fig. 20–12 from a point A just above the knee of the zero-bias curve draw arbitrarily selected load lines to the zero plate current axis. These lines should be on both sides of the operating point P whose position is located by the desired operating plate voltage, E_o, and one-half the maximum signal plate current. Along any load line AA_1 measure the distance AO_1. Along this load line lay off line O_1A_1 equal in length to AO_1.

For best operation the change in bias from A to O_1 and from O_1 to A_1 should be nearly equal. If they are not, select some other load line arbitrarily, until one is found which makes the equal-bias change a fact. The load resistance for the tube may then be found by

$$\text{Load resistance } R_L = \frac{E_{max} - E_{min}}{I_{max} - I_{min}}$$

This value of load resistance may be substituted in the following formula and the power output obtained.

$$\text{Power output} = \frac{[I_{max} - I_{min} + 1.41(I_x - I_y)]^2 R_L}{32}$$

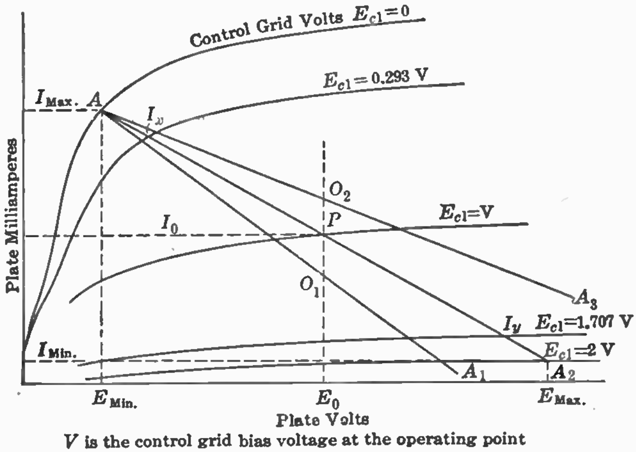


Fig. 20–12. Method of determining correct load resistance for a pentode power tube.

In these formulas, I is in amperes, R is in ohms, E is in volts, and W is in watts.

Distortion calculation from pentodes and beam power tubes. The following formulas will be useful in calculating distortion from tubes of the pentode or beam power output type.

$$\text{Second harmonic distortion} = \frac{I_{max} + I_{min} - 2I_o}{I_{max} - I_{min} + 1.41(I_x - I_y)} \times 100\%$$

$$\text{Third harmonic distortion} = \frac{I_{max} - I_{min} - 1.41(I_x - I_y)}{I_{max} - I_{min} + 1.41(I_x - I_y)} \times 100\%$$

Gain from high-mu tubes. Since the plate resistance of screen-grid tubes may be as high as 1 megohm, a better method must be found for estimating the voltage amplification than the formulas used with low-mu tubes. Such a formula is

$$A_{sg} = g_m R_{eq}$$

where g_m is the transconductance (in mhos) of the tube at the operating values of grid, screen-grid, and plate voltages.
R_{eq} is the equivalent resistance of r_p and R_L in parallel.

Since the tube resistance is very much greater than any load resistance likely to be used with the tube, and since the effective resistance of a small resistance shunted by a large resistance is not very much different from the small resistance unshunted, a good approximation of the voltage gain of a screen-grid tube, or of any similar tube having high amplification factor and high plate resistance, is

$$A_{sg} = g_m R_L$$

where R_L is the load resistance.

Problem 12–12. A tube with a transconductance of 1250 micromhos works into a load of 100,000 ohms. The tube internal resistance (r_p) is 0.8 megohm. Disregarding the shunting effect of the tube upon the load resistance, calculate the voltage gain. Then calculate the gain taking into account the fact that the tube resistance (r_p) shunts the load resistance (R_L).

An amplifier of three such stages delivers 1 volt. What is the input voltage?

Problem 13–12. In the plate circuit of a tube with a plate resistance of 10,000 ohms and an amplification factor of 10 is an inductance of 20 henries and a Q of 10. What alternating voltage e_o will appear across this inductance if 10 volts at 100 cycles is placed on the grid? NOTE: $e_o = i_b/\sqrt{R_L^2 + X_L^2}$ and $i_b = \mu e_g/\sqrt{(r_p + R_L)^2 + X_L^2}$.

Problem 14–12. In a receiver to be operated from a-c or d-c lines are one tube requiring 25 volts for the filament and four others requiring 6.3 volts each. They require 0.3 amp each, and the filaments are in series. If the receiver is to be operated from 115-volt circuits, what value of resistance is to be used between line and the filaments?

Problem 15–12. What loss in power is experienced in operating a tube under the conditions of maximum undistorted power as compared with the maximum output power? Assuming any reasonable tube values, calculate power under both conditions and express the loss in percentage of the maximum power.

Multi-unit tubes. It is possible to put two or more tubes on one mechanical structure and to put the combination in one envelope. Some multi-unit tubes are entirely distinct with separate cathodes and have means of preventing electrons from straying from their desired paths; other tubes share a cathode; and in still other tubes extra grid structures are added to exercise additional control functions, as in converter tubes.

In the 6L7, for example, there are seven electrodes. This tube is employed as a mixer tube in a superheterodyne circuit. It mixes signals coming from the antenna with signals generated in an oscillator within the tube. These two signals are at different frequencies, and in the output of the tube signals of still another frequency occur. For descriptions of these tubes see the material on superheterodyne receivers.

Putting two tubes in one envelope conserves space, shell material, and bases, but a circuit which needs two tubes rarely works as well if the tubes are combined in this way.

Metal tubes. The tube elements may be mounted within a glass or a metallic envelope. Modern receivers use both types. The metal-envelope tubes are smaller and better shielded than those in glass envelopes. They are more expensive to make, which may explain why they have not superseded the glass types. Fundamentally they perform exactly as their counterparts in glass envelopes.

Several kinds of bases and sockets are utilized, but, since the tube operations are not controlled by the kind of base, this matter will not be discussed here.

Cathode-heater connections. In heater-type tubes, it is common practice to connect the cathode to the center tap of a transformer winding supplying heater voltage or to the center of a 50-ohm resistor across this winding. It is sometimes desirable to have the heater somewhat positive with respect to the cathode to prevent any electrons from flowing from heater to cathode and thereby producing a-c hum. This voltage is of the order of 10 volts. In general, however, it is desirable to have no high potential between cathode and heater.

Tube operation. Tubes may be operated from batteries or from d-c or a-c lines. The heater-type tube may be operated either on alternating or on direct current without danger of hum appearing in the output. Filament-type tubes cannot be operated on alternating current except in the final output stage, since otherwise hum produced in the filament will be amplified by succeeding stages and cause noise in the output.

If several tubes are to be operated in a circuit, the electron-emitting filaments or the heater filaments may be wired in series or in parallel. If the source of filament power is of low voltage,

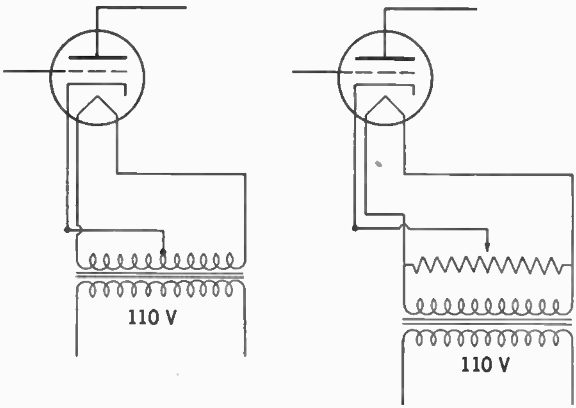


FIG. 21–12. Cathode of heater-type tubes is connected to filament circuit in two ways as shown. This minimizes hum.

such as a battery or a step-down transformer, the filaments are wired in parallel. Each tube gets the same voltage, and the current required from the source is the sum of the currents taken by the individual tubes. If the source of filament power is of high voltage, a house lighting circuit, for example, the tubes are frequently wired in series. Each tube, then, gets the same current, and the total voltage required to heat the filaments properly is the sum of the voltages appearing across each tube.

Filament voltages vary from 1.4 to 117. The 1.4-volt tubes are designed to operate directly from a dry cell: the 2-volt tubes operate directly from a single storage cell; the 6.3-volt tubes may be operated directly across a 6-volt storage battery. If these tubes are to operate from higher-voltage sources some means must be provided to limit the flow of current through the filaments. Thus if a 1.4-volt tube is to be operated from a 2-volt battery, a series resistance must be provided. If several tubes are to be operated with their filaments in series across a line whose voltage is greater than the sum of the tube voltages required, a voltage-dropping resistor must be employed.

The required resistance may be calculated from this formula:

$$\text{Required resistance} = \frac{\text{Supply volts} - \text{Voltage required by tube}}{\text{Current required by the tubes}}$$

Thus if two 32's, two 30's, and two 31's are to be operated from two dry batteries in series, the resistor is equal to $(3 - 2) \div 0.62$, approximately 1.6 ohms.

If one 6SA7, one 6SK7, one 6B8, one 25A6, and one 25Z6 are to be operated from a 117-volt line, the resistance equals $(117 - 68.9) \div 0.3$, or 160 ohms.

Grid-bias arrangements. A tube operated from batteries may gets its grid bias from a C battery. Batteries, however, are inconvenient, and every means are taken to avoid their use except where they are the only practical source of power.

One of the simplest ways of placing a steady negative voltage on the grid is by means of a resistor through which the plate current of the tube flows. The voltage drop ($I \times R$) along this resistor is of such a direction that the negative end can be connected to the grid and thus give the grid a negative bias with respect to the filament or cathode. This resistor is usually connected between the negative end of the plate supply voltage and the cathode, although the voltage drop along a resistor between battery and filament in battery-operated circuits may be used. Thus in Fig. 22–12 may be seen the two methods of providing bias. In one case the plate current flowing through the "cathode-bias resistor" provides the voltage; in the other the resistance needed to lower the battery voltage to the voltage required by the tube filament is utilized.

Although any current flowing through any resistance can be used, and therefore the plate currents of several tubes flowing through a common resistor may be utilized as bias, it is common practice to have each tube provide its own bias and thus to isolate its plate current from that of other tubes.

In series filament circuits, advantage may be taken of the voltage drop across tube filaments to bias one or more grids. Thus in Fig. 23–12 the filaments are wired in series. The same current flows through each filament, and the total voltage required

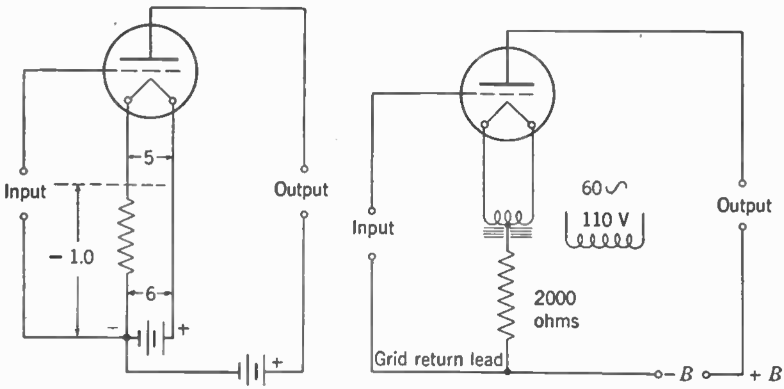


FIG. 22–12. Methods of biasing the grid negative.

is the sum of the individual tube filament voltages, plus the voltage drop across the 10-ohm resistors. If each tube requires a quarter ampere, then this current (250 ma) must be provided and the voltage is 5 times 5 or 25 volts (since each tube filament requires 5 volts) plus the voltage drop across the bias resistor for the first and second tubes. The current enters one end of the string of filaments and leaves at the other.

As we go along the series of filaments there is a series of voltage drops which may be used as biases.

For example, suppose that the final tube (No. 5) is a power tube requiring a C bias of 10 volts. We note that the negative terminal of the filament of tube 3 is 10 volts more negative than the negative terminal of tube 5 because there is a 5-volt drop across the filament of tube 4 and an additional drop of 5 volts across tube 3. We attach the grid circuit to this point in the circuit. If tube 4 requires only 5 volts C bias we can attach its grid circuit to the same point in the circuit because so far as tube 4 is concerned this point is only 5 volts negative. If the detector tube 3 requires a C bias of positive 5 volts we can attach its grid circuit to the positive leg of its own filament.

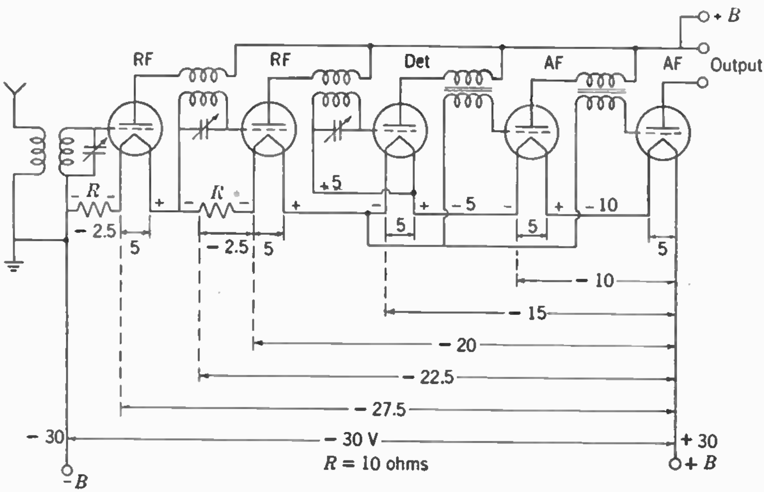


FIG. 23–12. Typical series-filament circuit in which voltage drops across filaments are used as C bias.

Suppose that the first two tubes in the set do not require 5 volts bias but some value less than this. All that is necessary is to insert a resistor in the circuit ahead of that tube so that the negative drop in voltage in this resistance will be utilized. Thus in Fig. 23–12 the 250 ma through a resistance of 10 ohms will produce a drop of 2.5 volts. The grid circuit should be attached as shown.

CHAPTER 13

RECTIFIERS AND POWER SUPPLY APPARATUS

Among the many useful functions of vacuum tubes is the conversion of alternating currents into direct currents for charging batteries or furnishing power for radio receivers or transmitters at the voltages desired. Battery chargers require high currents at low voltages; radio apparatus requires high voltages at low currents. High-vacuum rectifier tubes for x-ray apparatus will operate safely at voltages as high as several hundred thousand. Gaseous rectifiers will pass several thousand amperes for electroplating and welding and other applications, and three-element gaseous rectifiers perform very useful industrial functions.

The essential characteristic of the rectifier is that it permits current to flow in only one direction. It is only when the plate is positive with respect to the cathode that current will flow through the tube.

High-vacuum rectifiers. When high-voltage, low-current power is desired high-vacuum tubes are ordinarily used. If greater current is necessary gaseous rectifiers are required. If still greater currents are needed, mercury-pool rectifiers are employed. Figure 1–13 shows characteristic curves of receiver-type rectifier tubes giving the voltage drop across the tube for various currents flowing through it. Note that the drop varies with current for the high-vacuum tubes and is a constant value of about 15 volts for the gaseous tube.

Mercury-vapor tubes. When greater currents are needed than can be safely handled by high-vacuum tubes, or when greater efficiency is required, mercury-vapor tubes must be employed. The voltage drop across mercury-vapor tubes is essentially constant regardless of the current taken. They are specifically useful where currents are in excess of 100 ma.

Since the voltage drop in the tube is constant, some means must be provided to limit the current flowing through it, in the event of a short circuit in the load. Otherwise the cathode of the tube will be ruined. Gas tubes also create radio interference

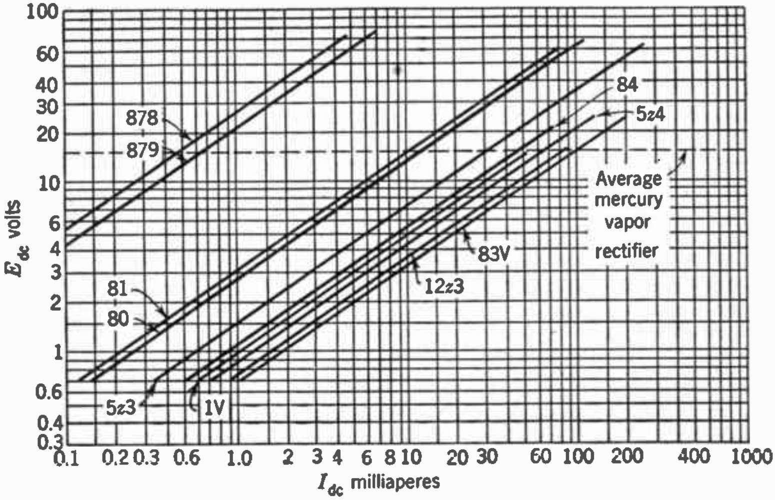


FIG. 1–13. Voltage drop versus current drain of several high-vacuum rectifiers. Note the dotted line at 15 volts representing the drop across a typical mercury-vapor rectifier.

at times. This trouble is due to the fact that they start conducting abruptly on the half cycle which makes the plate positive, and this steep wavefront or abrupt starting resembles turning on and off an electric switch. A choke coil of about 1 mh connected close to the plate and placed within a shield which also contains the tube will help reduce the difficulty. Condensers connected from the outer terminals of the transformer secondary windings to the center tap, plus the use of chokes. will usually eliminate most of the trouble of this sort.

To prevent excessive current flow in the event of a short circuit, resistors are often placed in the plate leads to produce a voltage drop. Fuses are essential in the transformer primary winding since during a short circuit the sole current-limiting factor is the resistance of the transformer.

Mercury-vapor tubes have heavy filaments in order to supply a very large number of electrons. Considerable time is required for these filaments to heat up to the required emitting temperature. If the plate voltage is applied and current taken before the cathode has reached the required temperature, a high voltage drop near the cathode may accelerate the positive ions resulting from ionization to sufficient speed to damage the cathode coating by impact. Time delay relays are employed to keep the load circuits open until the cathode has been heated sufficiently.

Because of their good regulation, mercury-vapor or gaseous rectifiers are employed in circuits where considerable fluctuation of output current is experienced. In radio receiver circuits the current taken from the power supply system is fairly constant; but in a transmitter, for example, where the output is keyed, the current requirements vary from instant to instant. If the internal drop within the rectifier tube varies with current output (as in high-vacuum tubes) the voltage supplied to the transmitter varies from instant to instant. This will not be true of the gaseous tubes, except for the variation in voltage caused by transformer *IR* drop. Vapor and gaseous rectifiers are used wherever the currents are high enough to make the voltage drop of a high-vacuum tube uneconomical.

Controlled rectifiers. A most important addition to the vacuum-tube family has been made in the thyratron, which is a gaseous rectifier with a control grid. The characteristics of such a tube present many interesting points. For example, if the grid is made negative by the proper amount no current whatever will flow to the plate even though it is positive with respect to the cathode. At any given plate voltage, however, there is a value of grid voltage which will permit current to flow—and when this occurs the current is limited only by the external circuit, not by the tube itself.

In other words, all the current flows, or none of it. There is another interesting characteristic. Once the current begins to flow, the grid has lost control over it and negative voltages on the grid make no difference—the current still flows. The only way to stop the flow of electrons from cathode to plate is to remove the positive voltage from the plate.

Thus the tube acts like a relay and can be used to key a radio transmitter or for any other power control purpose.

Now suppose that we place a-c voltages on the plate and a fixed bias on the tube. Whenever the plate has reached the

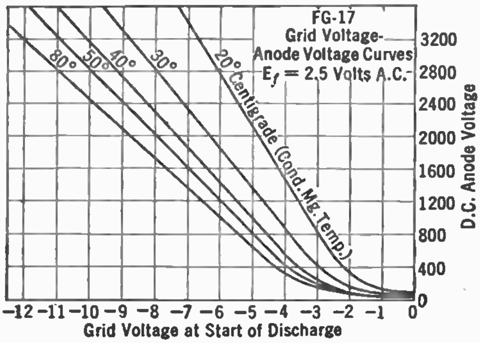


FIG. 2–13. Characteristics of mercury-vapor thyratron. Note that the temperature of the tube affects the starting voltage.

point in the positive half cycle such that its positive voltage is high enough to overcome the inhibiting effect of the grid voltage, current will flow. Then, at the end of the cycle, the plate becomes negative with respect to the cathode and current flow stops. By adjusting the bias on the grid the point in the cycle at which current starts may be controlled—but it will flow throughout the rest of that half cycle.

Now if alternating voltage is placed on the grid as well as the plate and if the phase between grid and plate is adjusted, any amount of average current may be taken from the tube, from zero to the maximum cathode emission permissible. The thyratron is a gaseous rectifier, with low internal voltage drop, which can supply many amperes of current under the control of the user by means of the variable phase between voltages on grid and anode. A high-vacuum tube acts like a rheostat and a rectifier in series. A thyratron acts like a latching relay plus a rectifier plus an opposing battery.

Relay tube. The OA4G is a gaseous triode useful for relay services. In an envelope filled with an inert gas or vapor at low pressure are a cathode, an anode, and a grid. The cathode is cold; that is, there is no heated filament or heater to supply electrons. The grid is a starter electrode, to initiate and control the gaseous discharge. A relatively small voltage on the starter electrode starts the discharge between starter electrode and cathode. This discharge ionizes the inert gas and enables a discharge between cathode and anode to take place. This is the discharge which performs the work required of the tube, such as closing a relay.

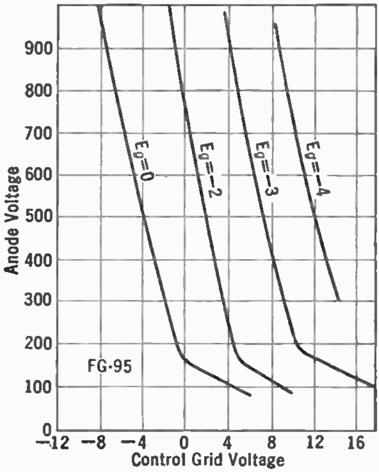


Fig. 3-13. Characteristics of shield-grid thyratron. Note that various starting voltages are possible, depending upon voltage of shield grid.

The first use to which this tube was put was remote operation of a radio receiver. A very small amount of energy is required to start the initiating discharge. This energy was transmitted over the power lines to the tube which was located within the radio receiver itself. The remote-control energy was generated at radio frequencies, and the power lines were utilized as the carrier of this controlling energy.

In the remote-control function, the starter electrode is maintained at a voltage just below that required for breakdown. In Fig. 4-13 this voltage is supplied from the power lines by R_1 and R_2. A tuned circuit with a high Q made up of L and C has a considerable voltage across it when a carrier of the proper frequency is impressed on the line at the distant point. This voltage increases the potential between cathode and starter electrode, and the increase of potential starts the discharge which flows between cathode and anode and closes the relay. The tube ceases to pass current when the carrier is removed, since the tube is supplied with alternating current.

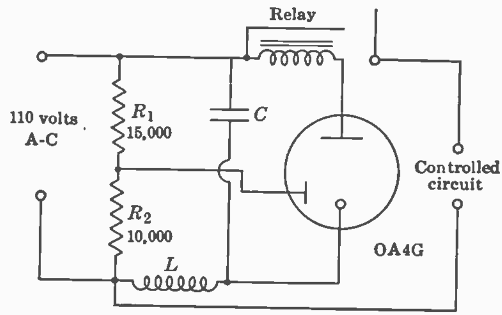


FIG. 4-13. Use of gas tube in a remote-control circuit.

Deionization time. Once the grid of the thyratron has permitted current to flow, the grid is surrounded by positive ions. These are one of the two products of ionization, that process by which electrons are knocked out of mercury (or gaseous) molecules to leave the molecule positively charged. The positive ions neutralize the negative voltage on the grid, preventing it from exercising any further control. When, however, the plate voltage is cut off, the ions recombine with electrons and become neutral molecules again. A few microseconds are required for this recombination process, known as the deionization time. If the tube is operated on 60-cycle power the plate voltage is cut off each half cycle when the plate becomes negative with respect to the cathode. Within the few microseconds required for deionization after the plate becomes negative, the grid has regained control and is now ready to permit or prevent further plate current at the will of the engineer. The deionization time in mercury-vapor tubes is of the order of 100 to 1000 microseconds. The use of thyratrons of this general type is therefore limited to frequencies in the audio range.

RECTIFIER CIRCUITS

Half-wave rectifier. The simplest rectifier circuit is shown in Fig. 5–13. The alternating a-c power source is connected in series with the tube and the load. Current flows through the circuit on the half cycle which makes the plate positive, but not on the other half cycle. If I_m is the maximum or peak value of the current passed through the conduction half cycle, the average value (which would be read by a d-c meter) is $0.318I_m$.

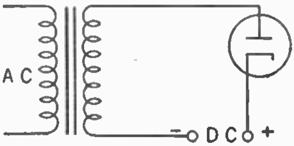


Fig. 5–13. Simple rectifier circuit.

The half-wave rectifier has two disadvantages: (1) The output contains an a-c component of the same frequency as the power source. If this frequency is low, say 60 cycles, it is difficult to filter or smooth it out so that the load has only direct current flowing through it. (2) The a-c power supply is furnishing power only half the time so that the peak load is high compared to the average load. These disadvantages limit its use to low-current loads, such as cathode-ray-tube circuits, and to applications where there is no transformer. In this case the unit using the tube can be operated on either alternating or direct current.

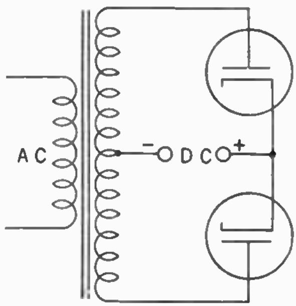


Fig. 6–13. Full-wave rectifier which passes current on both halves of the a-c cycle.

Full-wave rectifier. If two tubes are connected as in Fig. 6–13 a full-wave rectifier results. Now each tube conducts on each alternate half cycle so that current flows through the load all the time. Under these conditions the a-c component in the load has twice the frequency of the power source and it is easier to filter. The ratio between peak power taken from the transformer and the average power is lower.

Voltage-doubler circuits. It is possible to have two tubes in a circuit which will produce a d-c output voltage with a value equal to approximately twice the peak voltage of the a-c line from which the circuit derives its energy. The circuit in Fig. 7–13 requires two rectifiers with their outputs connected in series. This type of rectifier is most valuable in places where

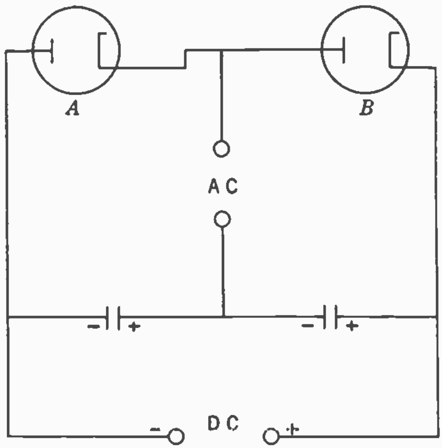


FIG. 7–13. Use of two rectifiers in a voltage-doubling circuit.

it is impossible or undesirable to use a power transformer. It is possible to obtain a useful output of 240 volts at 40 ma from a 117-volt a-c line.

For the moment consider only tube *B*. When the plate is positive, current flows through the tube and charges the condenser in its output. Now consider tube *A*. When the plate is positive, current flows through the tube and charges its condenser. Since the two tubes are connected opposite to each other across the line, the two condensers are charged in the same direction, so that the positive terminal of one connects to the negative terminal of the other. If no current is taken from them, as by a load, they will charge to the full voltage of the a-c input line. When current is drawn by the load, the terminal voltage drops somewhat.

In practice, a special tube such as the 25Z5 or 25Z6 is used. This has two separate diodes within one envelope. The heater

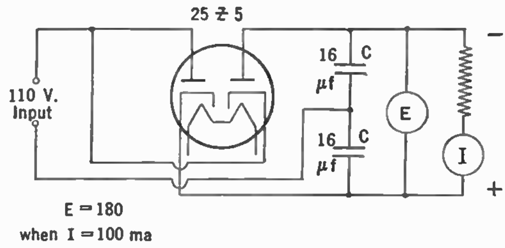


FIG. 8–13. Use of a rectifier with two plates and two cathodes to double the input voltage.

requires 25 volts and is connected in series with the heaters of the other tubes in the receiver and across the line. A typical circuit is shown in Fig. 8–13.

This circuit has two disadvantages: it cannot be grounded, nor can the d-c output be connected to one side of the a-c line.

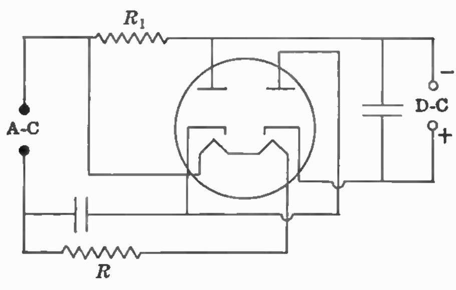


FIG. 9–13. Voltage doubler in which one terminal is at line or ground potential.

Hum may be produced because of the high voltage between heaters and cathodes. A circuit which overcomes the difficulty is shown in Fig. 9–13. One side of the output is connected to the line. It is called a half-wave doubler because rectified current flows to the load only on alternate half cycles of the input voltage. One of the diodes charges a condenser which discharges in series with the line and through the other diode on the next half cycle. The regulation is not as good as the circuit in Fig. 8–13.

Voltage-doubler tubes have other interesting applications. Since the cathodes and anodes are separate, they really constitute two complete rectifiers. Thus one half of the tube might be used to supply plate voltage and the other half to supply grid-bias voltage; or one side might supply plate voltage and the other half supply energy for a dynamic loud-speaker field coil.

Rectifier tube ratings. Vacuum rectifiers have fairly high internal resistance which varies with the current taken by the load. Gaseous and vapor rectifiers have a rather low constant voltage drop across them. Because of this low voltage drop, much more current can flow through them for the same power loss in the tube itself than can flow through high-vacuum tubes. Mercury-vapor rectifiers have characteristics which depend to some extent upon the temperature of the tube. Plate-current flow should be delayed until the cathode has had time to warm up to the proper temperature. In addition to this time delay, the mercury vapor in tubes employing it must be allowed time to reach the proper temperature. High-vacuum tubes are, in a measure, self-protecting against excessive current drain, but gaseous and vapor tubes must be protected against damage from this cause.

The **maximum peak inverse voltage** is the highest peak voltage that may be applied safely to the tube in the direction opposite to that in which it conducts current (plate negative). In single-phase half-wave circuits with sine-wave input and condenser input to the filter, the peak inverse voltage may be as high as 2.8 times the rms value of the applied secondary voltage.

The **maximum peak plate current** is the highest instantaneous current that a rectifier tube can safely stand in the direction in which it is designed to pass current. If a condenser is used as the input to the filter, the peak current often is many times the average load current.

The **maximum average plate current** is the highest value of average current that should be allowed to pass through the tube. This may be read on a d-c meter in the load circuit.

Dry disk rectifiers. Certain combinations of metals, such as copper sulphide and magnesium, or copper and copper oxide, or iron and selenium, have the useful property of passing current in only one direction. They may be used as rectifiers. The copper oxide rectifier is typical in its simplicity and reliability. It consists of a sheet of copper on one side of which has been formed a coat of cuprous oxide (Cu_2O). Properly made, this combination has relatively low resistance in the direction oxide-to-copper, with very high resistance in the reverse direction.

The units are generally made in the form of washers, of 1½-inch outside diameter. These washers are then assembled in any desired series and parallel arrangement on mounting bolts. Soft metal washers are placed between the oxide layer and the adjacent metal surface for the purpose of improving the contact with the oxide. The surface of the oxide is graphitized for the same reason.

This rectifier operates electronically and not electrolytically. Rectification commences instantly on the application of voltage, with no forming or transient condition interposed. Current is carried uniformly over the available area. Furthermore, the rectifying elements can be paralleled to any extent. Operation in series presents no difficulties, as the disks divide the voltage with approximate uniformity.

Dry disk rectifiers may be used either half wave or full wave. The bridge connection is common since it simplifies the transformer design and permits units to be operated direct from the a-c line without intervening transformer if so desired. The quality of the d-c wave obtained is good so that filtering is readily accomplished. Battery charging, battery elimination, magnet operation, loud-speaker excitation, etc., in fact almost any d-c application, can be successfully handled by these simple rectifiers. They are often used in measuring instruments. Thus a d-c movement can utilize the rectified current produced by a copper oxide disk when alternating current is applied. The disk is kept as small as possible so that it will have low electrostatic capacitance.

FILTER CIRCUITS

Since the rectifier passes current only when the anode is positive, and since alternating voltages are applied to it, the output is not a smooth flow of current but a pulsating one. Although this current may be applied directly to some uses, nearly always the pulsations must be smoothed out so that a steady direct current results. The smoothing takes place in a **filter** made up of shunt capacitances and series inductances or series resistances.

Let us consider the simplest situation, a rectifier shunted by a condenser with the load resistance across the condenser. If 60-cycle alternating current is applied to the rectifier circuit, for $\frac{1}{60}$ second the anode of the rectifier is positive and current flows into the condenser, charging it. Let us suppose that the condenser is completely charged in this time; that is, the voltage of the condenser is equal to the peak alternating voltage applied to the rectifier or some value less than this voltage. Now when the polarity of voltage applied to the circuit changes, that is, on the next half cycle when the anode is negative with respect to the cathode, the tube will not pass current, and the charge in the condenser begins to flow out through the load. If the rate at which the charge leaves the condenser is slower than the rate at which the condenser is recharged by the rectifier, the terminal voltage of the condenser will not fall to zero at each change in polarity of the applied alternating voltage, but to some positive value greater than zero.

In other words, the time constant of the *RC* circuit must be such that the time required to discharge the condenser is greater than the interval between the periods during which the rectifier is non-conducting. Under these conditions the terminal voltage of the condenser will look somewhat as shown in Fig. 10–13.

A mechanical analogy to the storage of charge in a capacitor at relatively short time intervals is a pump. If the handle is worked slowly the water all flows out during the downward motion of the handle. No water flows while the handle is being raised in preparation for the next downward motion. If, however, the handle is worked at a faster rate so that the complete handle cycle takes less time, water will flow continuously since

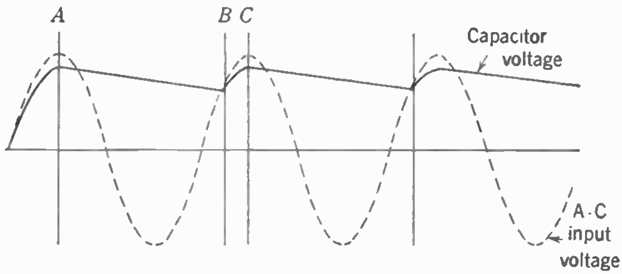


FIG. 10–13. Effect of using an *RC* circuit with a time constant greater than the time during which the rectifier is not passing current. The condenser maintains part of its charge between cycles of charging current.

it takes longer for the water to flow out of the pump than it takes to raise the pump handle.

Actually, of course, the situation is not as simple as this. For one thing, after the peak of the positive voltage wave has been applied to the rectifier circuit, the applied voltage begins to fall so that the condenser will be at a higher potential than the applied voltage. Furthermore, at the same time that current is taken by the condenser, current is also flowing through the resistor. When the voltage drop across the resistor due to the charged condenser equals the instantaneous input supply voltage, the rectifier ceases to pass current. The rectifier current, therefore, is sharply peaked with respect to time, indicating that the rectifier must be able to supply large peak current output.

If the rectifier is a full-wave system, the time between charges of the condenser is roughly half that of a single-wave system so that the voltage drop between rechargings will be less. The variation in the voltage in the full-wave rectifier occurs at twice the rate or frequency of that of a half-wave system. This "ripple" voltage is easier to filter because of its higher frequency.

On the half cycle in which the tube does not conduct, the voltage across the tube is equal to the line voltage plus the voltage to which the condenser has been charged. For this reason the tube must be insulated to withstand voltages approximately twice the output d-c voltage. Furthermore, since the tube conducts only during a short period near the voltage peaks, the transformer supplies power to the rectifier only part of the time. More complex systems have been devised in which the transformers and tubes are employed much more of the time than in a simple half- or full-wave rectifier.

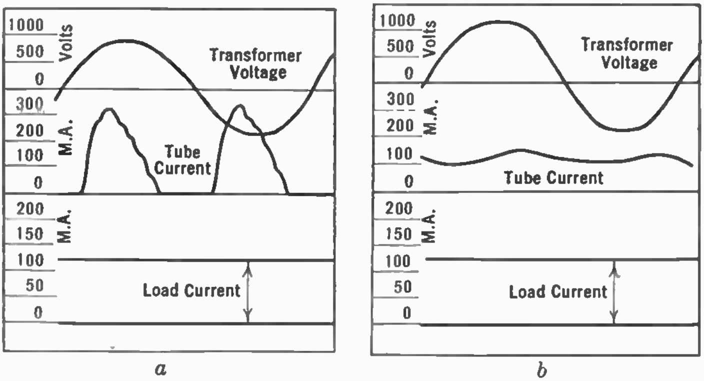


FIG. 11–13. (*a*) High instantaneous current required from rectifier with capacitive input. (*b*) Reduced tube current in an inductance input filter.

There are two general types of filter, depending upon whether a condenser or an inductor occurs first. The relations between the voltages and currents existing in the two types of filter system are shown in Fig. 11–13. Note that, with the inductor input, the peak currents required by the filter are much less than with the condenser input. For light load currents, the condenser system is almost always used since the output voltage is higher with a given input voltage than with an inductance input filter.

Where high currents are required, as in a transmitter, the inductor comes first. With this circuit the regulation * is better.

The connection which places an inductance first (Fig. 12–13) increases the life of the tube, permits the use of a tube whose emission has fallen below the point at which the tube would be useless in a capacity input filter, gives better regulation at current drains above 20 ma, reduces the filament emission required, and reduces the heating of the tube.

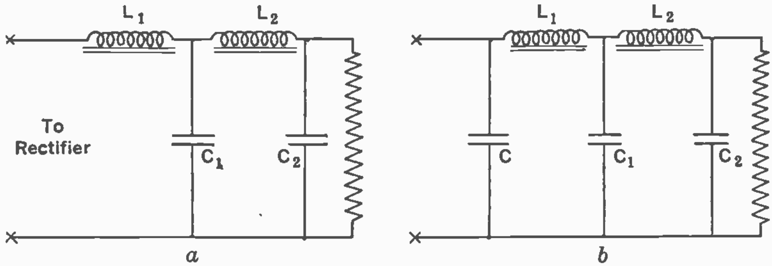


FIG. 12–13. Two types of filter: (*a*) inductance input; (*b*) capacity input.

Utilizing the natural property of a condenser to maintain constant voltage and the natural property of an inductance to maintain constant current, a proper combination of L and C in as many sections or stages as needed will make the output of a rectifier as free from ripple voltage as necessary.

On the half cycle in which the rectifier does not pass current, the voltage to which the condenser has been charged tends to maintain the flow of current through the load. When the condenser voltage starts to fall, the natural property of the series inductances is to prevent a fall of voltage by maintaining a flow of current. This is due to Lenz's law—a back emf is produced by the inductance when the current through it decreases. The polarity of this emf is in such a direction that it tends to force current to continue at its original value.

* Regulation is a term used to indicate the rate at which output voltage drops as the output current increases. Poor regulation indicates a high internal resistance and causes high terminal voltage drop with high current drain.

Filter design. The output of the rectifier into the filter may be considered a direct current plus an a-c component (ripple). The series inductances in the filter offer high impedance to the ripple and none to the direct current; the shunt capacitances offer little impedance to the ripple and much to the direct current. A proper combination of L and C will produce an output as free from ripple as is desired. If a single inductance and condenser do not reduce the ripple enough, another LC "section" may be added to the filter.

The condenser-input filter will have a higher d-c output than the inductance input, all other things being the same. It will require that the tube pass much higher peak currents. With a given tube, therefore, the condenser-input type of filter will enable less load current to be drawn from it without danger to the tube. Receiver power supply systems nearly always employ the condenser input; transmitters use the inductance input.

In Fig. 13–13, showing the inductance input, the ratio of the ripple voltage (alternating voltage across load) to the alternat-

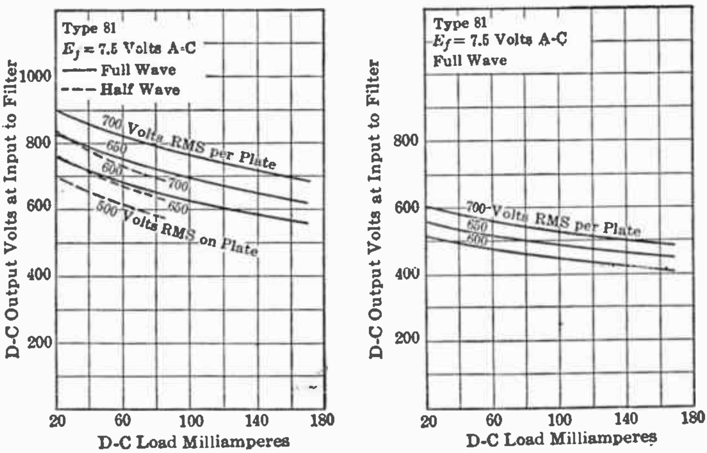


Fig. 13–13. Comparison of condenser-input (*left*) and choke-input (*right*) filter systems. Voltage drops due to resistance of filter chokes are not considered and must be added to these regulation curves if complete regulation data are desired.

ing voltage applied to the input to the filter is $1/\omega^2 LC$. Thus

$$\frac{\text{Ripple voltage across load}}{\text{Alternating voltage to input}} = \frac{1}{\omega^2 LC}$$

From this equation the value of the LC product for any frequency and any desired attenuation of the alternating voltage may be determined. If the filter has two sections the relation becomes $1/\omega^4 L^2 C^2$, and for a circuit in which the 120-cycle ripple is the strongest (as in full-wave rectifiers) the formulas become

$$\frac{\text{Ripple}}{\text{Input a-c voltage}} = \frac{1}{0.57LC}$$

where L = henries for 1 stage.
C = microfarads.

$$= \frac{1}{(0.57LC)^2} \text{ for 2 stages}$$

In other words, the filtering action produced by one stage is multiplied by itself when two stages are used and when the L and C components in the two stages are identical.

Example 1–13. A choke of 10 henries and a condenser of 10 μf are employed. What is the ratio between the ripple voltage and the applied input alternating voltage?

$$\frac{\text{Ripple voltage}}{\text{Applied voltage}} = \frac{1}{0.57 \times 10 \times 10} = \frac{1}{57}$$

or the ripple will have $\frac{1}{57}$ of the voltage applied to the input to the filter or roughly 2 per cent. Now if two stages are employed the ratio will become

$$\frac{\text{Ripple voltage}}{\text{Applied voltage}} = \frac{1}{0.57 \times 0.57 \times 10^2 \times 10^2}$$

$$= \frac{1}{0.3250 \times 100 \times 100}$$

$$= \frac{1}{3250}$$

Since the filtering depends upon the product of L and C it seems that any set of values would be useful provided that the product remained the same. In practice this is not true, owing to the problem of regulation. There is a minimum value of the inductance beyond which it is not possible to go. The function of the inductor (choke) is to prevent change in the current flowing. When the rectifier ceases supplying current to the filter, during the half cycles in which the plate is negative, the effect of the inductance is to continue the flow of current. If the inductance is too low to produce a continual current, the condensers will completely discharge between half cycles, the current peaks drawn by the rectifier tube will be high, and the output load voltage will vary widely with load current.

The minimum value of inductance which should be used in a 60-cycle half-wave rectifier is $R_L \div 25.6$, and for a full-wave rectifier it is $R_L \div 1130$, where R_L is the effective load resistance which must include the resistance of the load plus the resistance of the chokes. It may be seen at once that the minimum inductance required for a full-wave circuit is very much smaller than for a half-wave circuit.

Swinging choke. If the output current required from a rectifier-filter system varies widely, for example in a keyed transmitter, the voltage supplied will vary because at high currents there is a greater voltage drop in the system than at low currents. Much greater freedom from regulation troubles will be had if a gaseous rectifier is used than if a high-vacuum tube is employed as rectifier for the simple reason that the voltage drop across the tube itself is relatively independent of the current passed through it.

Another method of keeping the output voltage more independent of the load current is to use a "swinging choke" as the input to the filter. A swinging choke has high inductance at low currents through it and low inductance at high currents through it. When the inductance is low, as it is at high current drains, the filter resembles a condenser-input system with high output voltage. When the inductance is high, as when the current taken from it is low, the filter resembles an inductance-input filter.

The "swinging" effect of the choke is a matter of design, so that the inductance varies widely with the current through it. The design involves the dimensions of the air gap in the iron core. With a very small air gap the core approaches saturation at high values of current, but at low currents the choke will have a high inductance.

Choke coils used in filters should always be rated at so many henries inductance with so many milliamperes of direct current flowing through them. A typical swinging choke might be rated at 6 henries at full or rated current and 15 henries at 10 per cent of the full current.

Multi-section filters. If the ripple, that is the alternating voltage component in the filter output, is too high for the purpose at hand, it may be reduced by employing additional filter sections. These sections may be identical with the preliminary section or they may be different. Rectifier-filter design does not lend itself to mathematical rigor, and most power supply systems are designed by trial or on the basis of past experience. The values of inductance and capacitance are usually greater than necessary so that more than sufficient filtering is secured. Naturally if the output of the filter is to supply the plate circuit of an amplifier tube whose output is to be still further amplified (a microphone amplifier, for example), the ripple must be much smaller than if the amplifier output is fed directly to a loud speaker. In the first case the ripple in the plate voltage will be amplified along with the microphone voltages and will appear in the final output as a hum. The ripple should be as low as 0.005 per cent. If the desired audio voltages are already raised to a high level the ripple voltages will produce a hum that is not bothersome. Here it may be as high as 1 per cent of the direct voltage. In general, however, more than sufficient filtering is employed.

Bleeder. A resistor, called a "bleeder," is usually connected across the output of the filter. It has a resistance value high compared to the load resistance. In practice a current of about $\frac{1}{10}$ of the load current is permitted to pass through it. The use of this shunt resistance improves the regulation of the circuit; that is, it prevents high no-load voltages across the output and stabilizes the output voltage as a function of load current. It also discharges the condensers when the power is removed from the input to the rectifier and prevents shock if the operator accidentally comes in contact with any part of the system.

The voltage divider. The maximum voltage available is usually applied to the plates of the tubes. Sometimes the power tubes require higher voltages than the other tubes, but radio receivers are usually designed to use tubes which take 250 volts on the plates. The screens require lower voltages. Grid-bias voltages must be negative.

The screen and lower positive voltages are usually supplied from the maximum voltage by means of a voltage-dropping resistor with additional filtering sometimes provided for the lower voltages. The bias resistors usually appear in the cathodes of the individual tubes but may be placed in the negative lead of the power supply system. The bleeder may have taps on it so that it can be used as a voltage divider.

Engineering the voltage divider. The lower the resistance of the voltage divider, the better will be the regulation of the entire device, but the greater the load on the tube. In general, a voltage divider is engineered as follows. In Fig. 14–13, suppose that the current flowing through the resistance R_1 is 20 ma. This is known as the "bleed" or "waste" current. It flows whether or not there are any loads connected at the several taps. At the point *B*, a voltage of 45 and a current of 2.5 ma are desired. The value of resistance is, according to Ohm's law, $E \div I = 45 \div 0.02$ or $R_1 = 2250$ ohms. This 2.5-ma current is added to the 20 ma taken by the lowest resistance. Thus through R_2 flows 22.5 ma, and since 90 volts is desired at point *C* the resistance R_2 will be 45 $(90 - 45) \div 22.5$ or 2000 ohms. If the load which requires 90 volts takes a total of 15 ma, the final resistance will be $(180 - 90) \div 37.5$ or $90 \div 37.5$ or 2400 ohms. The entire resistance will be 6650 ohms, with taps at 2400, 2000, and 2250 ohms. The greatest amount of power must be dissipated by the 2400-ohm resistor, and so, if the entire resistance up to R_4 is wound with wire large enough and on a frame that can dissipate the heat corresponding to 4 watts, there will be no trouble. Voltage dividers are available which consist of a single winding of resistance wire on a heat-resisting form. By means of several sliders the correct voltages can be obtained easily.

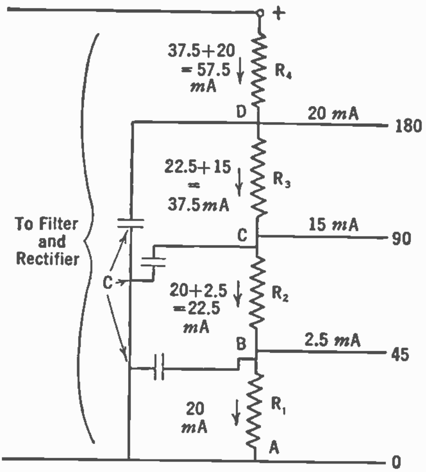


FIG. 14–13. Designing the voltage divider.

Problem 1–13. The maximum voltage needed with the voltage divider of Fig. 14–13 is 180 volts, but the output of the filter is 200 volts under load. What is the value of resistance R_4 to reduce this voltage to 180, and what must be its power dissipation in watts?

Problem 2–13. Using a tube with a filament voltage of 7.5 volts, operated from direct current, the proper C bias is 50 volts. The plate current then is 25 ma. What is the voltage that must appear across the C-bias resistor when the tube is operated from alternating current and when this resistor is connected to the center of a resistanceless transformer winding?

Problem 3–13. What power in watts is dissipated in each of the resistances in Fig. 14–13 under the conditions of Problem 1–13?

Problem 4–13. The resistance of the filter is 300 ohms. What must be the output of the rectifier if the output of the filter is 250 volts and if 40 ma flows through the voltage-divider resistance R_4?

Grid bias requirements. The steady negative voltage required by the grid of an amplifier must be "clean"; that is, it must be steady in value regardless of the alternating voltages placed upon the grid circuit. In some special cases this requirement is not necessary. The source of bias voltage should be adequately by-passed and should have good regulation for use with class B audio amplifiers, class B linear radio amplifiers, and grid-bias-modulated class C amplifiers. The regulation is unimportant with class A audio amplifiers (since there is no grid current) or with class C plate-modulated amplifiers.

Grid-bias sources. Methods using a C battery between the lower end of the grid-cathode input circuit (the "grid return" circuit) and the cathode, and methods utilizing the voltage drop along a resistance in a filament lead of battery-operated circuits, have been described. The most common source of bias is a resistance placed in the cathode circuit where the plate and screen (if any) currents must flow through it. The voltage drop across this resistance is employed as bias. But there are other possibilities, all of which find application at one time or another.

For example, if the cathode of the tube in question is connected to a point up from the most negative end of the voltage divider of a rectifier-filter system, then points below this connection are negative with respect to it and this negative voltage may be employed as bias. Of course a rectifier-filter system or a motor-generator can be employed to produce direct voltages for use as bias, just as plate voltages are supplied, but somewhat more care must be taken to filter the output.

In some circuits the grid is permitted, and often forced, to become positive with respect to the cathode. On the half cycles of excitation, the grid will draw current. This current can be forced to flow through a resistance, and the voltage drop along this resistance can be used as bias.

Cathode-resistor bias is called "self-bias" since the voltage supplied to the grid is a function of the plate current. If for some reason the plate current should increase, the voltage drop along the cathode resistor increases. As a result the grid bias increases, thus limiting the plate current. On the other hand a fixed bias, such as that supplied by a battery or by a separate

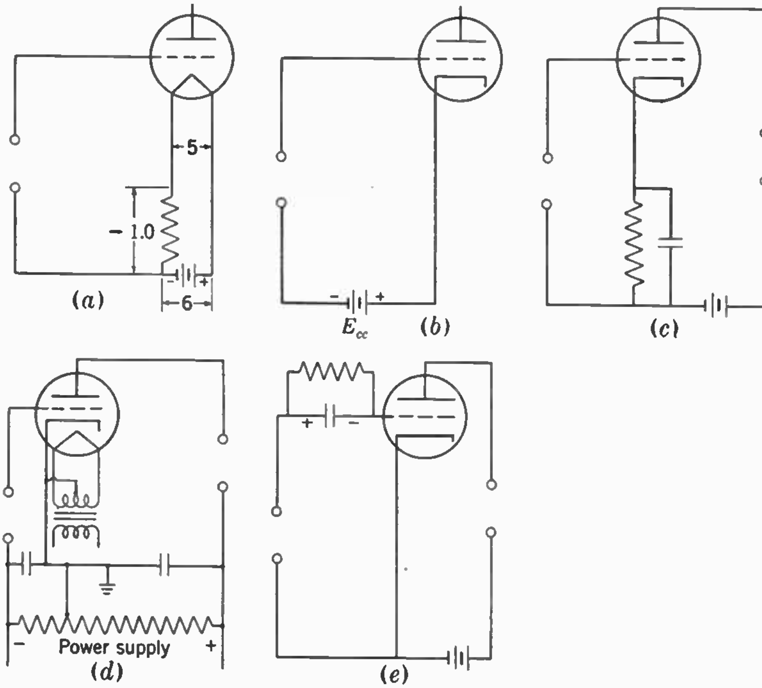


Fig. 15–13. Methods of supplying grid bias.

(*a*) Voltage drop along a resistor in filament lead. Here the grid is 1 volt negative with respect to the negative end of the filament.

(*b*) Conventional C battery.

(*c*) Cathode bias produced by plate current flowing through a resistor.

(*d*) Tap on power-supply voltage divider. Here the grid is negative with respect to the cathode by the voltage drop to the left of the cathode connection.

(*e*) Grid leak and condenser. Grid must go positive at some portion of the excitation cycle so it collects electrons. This current produces a voltage drop along the grid leak resistor.

power supply system, has no such self-regulating feature since the voltage supplied by it is independent of plate current.

More power can be obtained from an audio amplifier operated with fixed bias than from one operated with cathode bias, if the same plate voltage and the same excitation are employed.

This is because, with cathode bias, large input voltages produce greater average plate currents which increase the tube bias and limit the plate current. With fixed bias, this is not true and greater excitation produces greater output.

The resistance required to supply cathode bias may be found from

$$\text{Resistance (ohms)} = \frac{\text{Bias voltage desired} \times 1000}{\text{Cathode current (ma)}}$$

where cathode current is equal to the sum of the plate and screen currents.

Ballast tubes. A current-regulator tube consists of a metallic wire filament, which has a large temperature coefficient of resistance, immersed in a gas. If the voltage across the wire increases, the temperature of the wire changes, increasing the resistance so that the current through the wire changes very little. Such a tube is often placed in series with a load, such as a radio receiver power transformer, so that, as the line voltage changes, the tube with its ballasting action prevents any change in the voltage supplied to the transformer, that is, the change in voltage is largely absorbed by the ballast tube. For example, the B4 ballast tube passes 1.24 amp at 105 volts and 1.36 amp at 125 volts. A change of 19 per cent in voltage produces a change of only 9.7 per cent in current. This current flows through the primary of the power transformer, or other load, and has half the variation of the line voltage.

Tubes of this type are available in several voltage ranges, for example, 5–8, 3–10, 15–21, 40–60, and 105–125 volts. Ordinarily the transformer must be designed to operate from a voltage lower than that of the line so that the voltage required by the tube plus that required by the primary of the transformer equals the line voltage. The tube operates slowly and cannot regulate rapid fluctuations.

Voltage regulator tubes. The ballast tube is a *current* regulator. Another useful tube is the *voltage* regulator, which consists generally of two electrodes in a gas. A discharge occurs between the electrodes. The voltage across the discharge is approximately constant for a considerable variation in discharge current. In this characteristic it resembles the mercury-vapor and gaseous rectifier tubes already described.

The voltage-regulator tube is placed in parallel with the load so that variations in the applied voltage are smoothed out with the result that the load has a constant voltage applied to it. The tube may be employed to reduce ripple voltages or to improve regulation when the load current changes. The tube operates quickly. It is designed for a definite operating voltage, and the current flowing through it must remain within the maximum and minimum values.

Surge protector tubes. Surge protector tubes are two-electrode gaseous discharge tubes which are connected across a line or a load for protection against excessive voltages. When the voltage exceeds the breakdown voltage of the tube, the discharge takes place, and, as is characteristic with such discharges, the voltage across the terminals is more or less constant.

Regulator tube circuit design. In Fig. 16–13 is shown a ballast tube connected in series with the primary of a power trans-

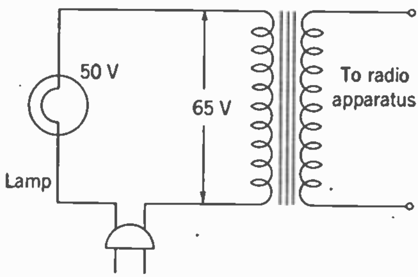


FIG. 16–13. Use of current regulator or ballast tube.

former. The circuit is designed so that 65 volts appears across the primary of the transformer, and 50 volts across the tube. When the line voltage increases, more drop occurs across the tube but the current through it stays constant so that the drop across the transformer remains fixed.

A circuit with a voltage regulator is shown in Fig. 17–13. The 874 tube maintains a voltage drop across its terminals of 90 volts for a current range from 10 to 50 ma. If the load requires 20 ma at 90 volts the value of the resistance in series with the tube may be calculated as follows. Allow a current of 20 ma for the tube. This total current flowing through the resistance should reduce the terminal voltage to 90 volts. Thus if the applied voltage is 150 the required resistance is $R = E \div I = 60 \div 0.040 = 1500$ ohms.

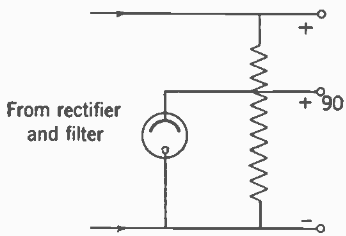


FIG. 17–13. Circuit using voltage regulator tube, such as the 874, to maintain 90 volts across terminals marked.

Since the voltage drop across the tube remains almost constant regardless of current changes, an increase in the applied terminal voltage results in an increased voltage drop across the resistance. On the other hand, an increase in load current results in a decrease in tube current so that the load voltage applied to the load remains constant.

Vibrator power supply systems. Since radio transmitters and receivers require d-c voltages higher than can conveniently be obtained from batteries, the advent of automobile and other forms of mobile radio apparatus forced engineers to develop special forms of power supply systems for these purposes. Early forms of motor-generators or rotary converters, which took 6- to 24-volt power from storage batteries and transformed it into higher-voltage direct current, were followed by most ingenious vibrator systems which accomplished the same end at less expense and required much less space.

The simplest method of securing high direct voltage from a source of low direct voltage is as follows: A switch periodically reverses the direction of flow of current through the primary of a transformer, the secondary of which furnishes alternating voltage of the required amplitude to a rectifier-filter system. Of course the current through the secondary reverses its direction whenever the primary current reverses, and the purpose of the rectifier-filter is to secure current which does not reverse its direction of flow. Imagine, now, a second switch arranged between the secondary of the transformer and the load (receiver or transmitter). This switch is so arranged that when the current in the secondary tends to flow from right to left, let us say, instead of left to right the switch changes a set of contacts so that the current continues to flow from left to right. These two switches, the primary and the secondary, can be mechanically coupled together so that they operate in synchronism. In this manner the rectifier tube can be eliminated. The filter, however, is still necessary.

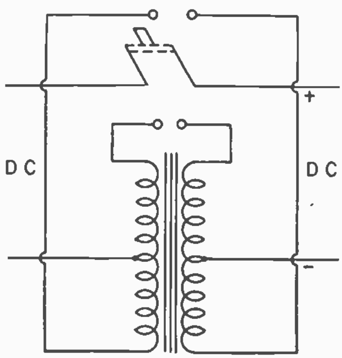


Fig. 18–13. Elements of vibrator power supply. As the switch is moved up and down, the direction of current flow through the windings reverses. The switch contacts on the right reverse the direction of secondary current in synchronism with the reversals in the primary so that the polarity of the output terminals does not change. See also Fig. 11–6.

In practice the switches are vibrating reeds which carry contacts and which are attracted toward one or the other of a stationary set of contacts by an electromagnet. In this manner the desired reversals in current flow are obtained. Such a mechanical vibrating switch is called **a vibrator;** it can be arranged to act as a half-wave rectifier, as a full-wave rectifier, or simply as a current reverser for the primary circuit without the additional provision of reversing the secondary current. Vibrators are available which will operate from 6 volts, as from a storage battery, or from higher voltages, such as a 24-volt farm lighting system, or from 110 volts direct current, or other voltages. The transformer, of course, serves its usual purpose of changing the low voltage to a high voltage. The secondary switch acts as a rectifier.

The transformer must be engineered especially for this service since the primary switch does not produce a pure sine wave of voltage. The contacts on the vibrator must be carefully designed and made and maintained in proper condition; careful consideration must have been given to the time of contact in the

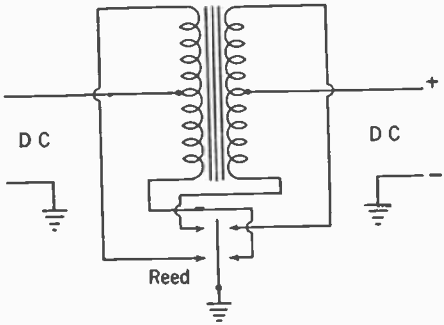


FIG. 19–13. Reed vibrated by electromagnet takes place of switch in Fig. 18–13 to produce high-voltage direct current from low-voltage direct current.

primary cycle of current reversal, the time of contact on the secondary side, and the proper use of condensers to produce the desired wave form. Much research has gone into the vibrator method of transforming one kind of power to another, leading to the present successful mobile radio transmitters and receivers.

A very good description of the vibrating systems will be found in the *MYE Technical Manual*, published by P. R. Mallory & Co. in 1942.

CHAPTER 14

AUDIO AMPLIFIERS

Amplifiers may be classified according to: (1) frequency range, i.e., audio, radio, video; (2) voltage or power amplification; (3) mode of operation (class A, B, etc.); (4) circuit connections.

An audio amplifier may be designed to amplify all audio frequencies, say from 20 up to 20,000 cycles. Sometimes it may be called upon to amplify only a limited band in this region, say 250 to 2500 cycles as for telephone communication, or a narrow band in the vicinity of 400 cycles for aircraft signaling. A radio-frequency amplifier usually handles only a narrow range of radio frequencies and is, therefore, said to be a selective amplifier. The frequencies may be anything up to hundreds of megacycles. A video amplifier is called upon to cover a very wide range, perhaps several megacycles wide. A d-c amplifier is useful in amplifying very low frequencies as well as direct currents and voltages.

Voltage and power amplifiers. The difference between a voltage and a power amplifier is only a matter of degree. Both amplify voltage, and both amplify power. In general the voltage amplifier is arranged so that the load resistance differs greatly from the plate resistance of the tube. Tubes with high internal resistance are generally employed. The fact that the d-c power of the battery or other plate-voltage source is converted to a-c power at low efficiency does not matter, as the purpose is to increase an input a-c voltage to a higher output a-c voltage—as high in fact as is consistent with circuit stability.

A power amplifier, on the other hand, has as its prime requisite the delivery of a-c power to a load, and whether or not there is an increase in the voltage from input to output is immaterial. D-c power is to be converted to a-c power, and efficiency of conversion is a matter of concern. The maximum power transfer will take place between a tube and a load when the impedance of the load is equal to the internal resistance of the tube. Even here it is customary to make the load impedance somewhat greater than the tube resistance, in the interest of lower distortion. Tubes designed primarily for power amplification, therefore, are usually of low internal resistance, and low amplification factor naturally results.

In pentodes and beam power tubes the plates merely act as collectors of electrons and the plate voltage has very little control over plate current. Load impedances about one-tenth that of the tube resistance are commonly used. The tubes have high "sensitivity"; that is, they produce high power output for little alternating grid voltage input.

Amplifiers are also classified according to their mode of operation, i.e., whether the grid is allowed to go positive or not, and according to the portion of the input a-c cycle during which current flows in the plate circuit.

Mode of operation. A class A amplifier is so operated that the operating point is approximately at the center of the I_b-E_b characteristic curve. The exciting grid voltage is such that the grid does not go positive nor is the lower bend of the characteristic curve utilized. Current flows in the plate circuit for the entire a-c cycle, and the wave form of the output is similar to the input grid voltage; distortion and efficiency are low.

A class B amplifier is operated so that the operating point is much lower on the characteristic curve, almost at the plate-current cut-off point. Plate current therefore flows only during a portion of the cycle. The grid is driven positive. Distortion is high, and efficiency is fairly high. If a reasonably undistorted output wave form is desired, two tubes must be used in a "push-pull" circuit, or a tuned circuit must be employed in the output.

In a class C amplifier, the grid is biased so that the plate current is definitely cut off, plate current flowing only during a portion of the input cycle when the grid is driven less negative than the cut-off value. Efficiency is high, and the output wave form is so distorted that this mode of operation cannot be used for audio amplifiers. In amplifiers having a resonant circuit in the output, undistorted modulation of a radio-frequency wave may be produced.

A class AB amplifier is somewhere between class A and class B. The bias is greater than with class A; and therefore higher

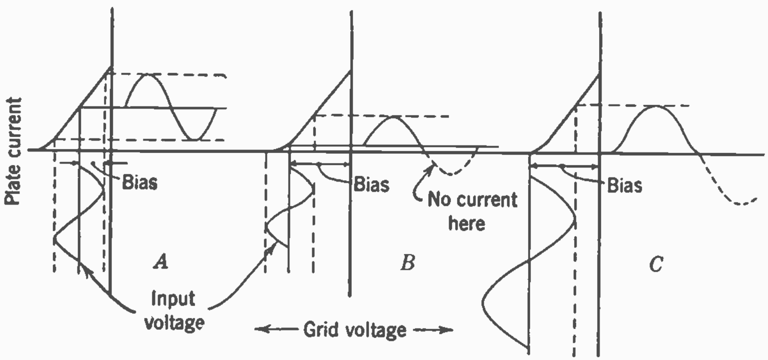


Fig. 1-14. Class A, B, and C amplifiers. Note bias values, portion of characteristic employed, and form of plate current.

plate voltages may be employed. The excitation can be greater than with class A; and higher power output is obtainable although with somewhat greater distortion than with class A. Because a non-linear portion of the characteristic curve is used, distortion is produced. Class AB amplifiers, therefore, are not operated singly but only in push-pull, a mode of operation which cancels out the even harmonics.

Requirements of audio amplifiers. An amplifier to operate at audio frequencies must have the following requirements:

1. The amplification versus frequency curve must be sufficiently flat for the job to be done.

2. The distortion (harmonic production) in the output must not be greater than is satisfactory or tolerable.

3. The amplification must be sufficient to deliver the required power from the given input voltage.

4. The noise and hum must be lower than some predetermined limit.

5. The amplification should not vary much with ordinary changes in filament temperature, plate voltages, etc.

6. The amplifier must work out of a certain impedance and transmit its power to another (usually different) impedance.

Multistage amplifiers. A single stage of amplification is rarely capable of satisfying the above requirements. Often it is possible to get high amplification in a single stage but at the expense of a poor frequency characteristic. Therefore, means must be provided for using the output of one stage to drive the input of another. Preliminary stages in a multistage amplifier (**cascade** amplifier) are designed to produce a maximum of voltage amplification with little regard to the efficiency of power amplification. The final stage is usually the power amplifier and is engineered to deliver the desired power to the load with as high efficiency as is practical.

Resistance-capacity coupled amplifier. The simplest method of connecting two amplifiers together is by means of a resistance-capacity network. This circuit is shown in Fig. 2–14. Voltage

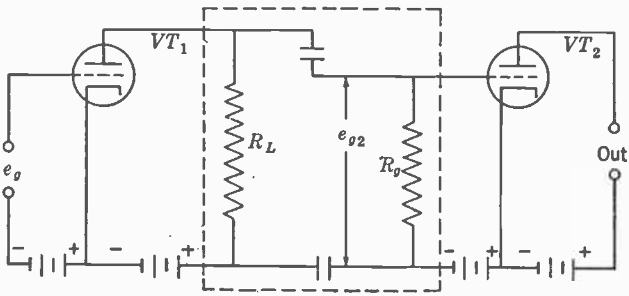


FIG. 2–14. A two-stage resistance-capacitance coupled amplifier.

fed to the grid of the first tube reappears in amplified form across the plate resistor, R_L. This voltage is used to drive the grid of the second tube.

The equivalent circuit of a single triode stage coupled to another following tube by a resistance-capacity network is shown in Fig. 3–14. Here R_L is the load resistance of the first tube, R_g is the grid resistance of the following tube, and C is a condenser designed to keep the d-c plate voltage of the first tube from being impressed upon the grid of the second tube. This condenser must permit no direct current to flow but must allow all alternating currents to flow. Alternating voltages developed across R_L produce voltage drops across the condenser and across R_g. The voltage across C is lost, since it is not impressed upon the following tube. The reactance of C, therefore, must be low, and, since the reactance of a condenser is highest at low frequencies, the reactance of this condenser has a controlling effect upon the low-frequency amplification of the circuit.

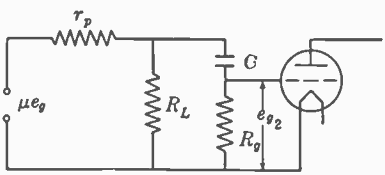


Fig. 3–14. The equivalent of one stage of the amplifier of Fig. 2–14.

The maximum amplification from a tube will be obtained when the maximum possible load impedance is in the plate circuit. C and R_g, shunted across R_L, tend to lower the impedance of R_L and consequently to lower the amplification that can be obtained. The series impedance of C and R_g must be high compared to R_L. Since the reactance of C must be small, it follows that R_g must be high.

In a triode amplifier of the resistance-capacity type, the effective load impedance in the plate circuit of the first tube—made up of R_L, C, R_g, and the inherent capacitances in the wiring, inside the tubes, between tube and socket terminals, etc.—should be several times the plate resistance (r_p) of the first tube. In amplifiers using pentodes or tubes with very high internal plate resistance, the load impedance must simply be as high as possible or convenient. In this case the amplification is equal to the product of this impedance and the transconductance of the tube. Thus

$$\text{Amplification} = g_m \times Z_0$$

where Z_0 is the effective load into which the tube works.

It is common practice to use values of C of the order of 0.01 μf and R_g values of 0.5 to several megohms.

Middle-frequency gain. At frequencies of the order of 1000 cycles, the reactance of the condenser in comparison to the resistance of R_g is small and has little effect upon the amplification. The maximum possible amplification is now obtained. With triodes it will be equal to

$$\text{Amplification} = \frac{\mu R_{\text{eq}}}{R_{\text{eq}} + r_p}$$

where R_{eq} is the equivalent resistance of R_L shunted by R_g. With pentodes or other high-resistance tubes the amplification will be $g_m Z_0$.

Low-frequency gain. At frequencies lower than those considered above, the reactance of C becomes higher and must be taken into account. Across this reactance will appear a voltage that cannot be applied to the following tube. Of the total alternating voltage developed by the first tube, therefore, some is lost across C and that across R_g is all that can be usefully applied. The amplification now depends upon the relative values of X_c and R_g since the larger X_c is compared to R_g the more voltage is lost across C and the less appears across R_g. These two elements must be taken into account together. If R_g is large, C can be small; if R_g is low, C must be larger.

Actually, if X_c is equal to R_g, both expressed in ohms, the amplification will be 70 per cent of the maximum possible. If X_c is ⅓ R_g, the amplification will be 95 per cent of the maximum attainable.

If C is not a good condenser and permits some direct current to flow through it, the grid of the second stage will be unfavorably biased and the entire system may become unstable. This will be caused by the direct current flowing into C which will charge the condenser at a rate that may be faster than it can leak off so that the system blocks and nothing gets through. Or the condenser may discharge at an audible rate, or at a rate of one or two cycles per second, so that the desired signals are modulated with the voltages produced by the periodic charge and discharge of C. If this occurs at a low but audible rate, the noise set up is called "motorboating."

High-frequency gain. Loss in gain at high frequencies arises from other causes. Of course the reactance of C is small at high frequencies; but by the same argument the reactance of any inherent capacitances shunted across R_L or R_g is small and thus tends to reduce the load impedance into which the tube works. These capacitances are made up of the wiring capacitance, socket and base capacitance, plus the inherent capacitances existing within the tube.

The higher the gain at middle frequencies, the more important do these stray capacitances become since any loss of gain due to them is *relatively* more important if the gain is high than if the gain is low.

The actual amplification at a frequency where the reactance of shunting capacitances becomes effective compared to the amplification at, say, 1000 cycles, may be found from

$$\frac{\text{Gain at high frequency}}{\text{Gain at 1000 cycles}} = \frac{1}{\sqrt{1 + (R/X)^2}}$$

where R is the effective resistance of the grid leak, the coupling resistance, and the plate resistance all in parallel, and X is the reactance of the stray capacitances shunting the three resistances in parallel.

At the frequency at which the stray shunting capacitances have a reactance equal to three times the value in ohms of the effective resistance of the three resistances in parallel, the amplification will be 95 per cent of the amplification at, say, 1000 cycles. If the frequency is raised until the reactance is one-third the effective resistance of the circuit, the amplification will be 70 per cent of the maximum attainable at, say, 1000 cycles.

Problem 1–14. Suppose that the amplification of a single pentode stage working into 100,000 ohms at 1000 cycles is 100. The grid leak is ½ megohm; the plate resistance of the tube is 1 megohm. Stray capacitances plus tube capacitances add up to 25 $\mu\mu$f. What is the voltage gain at 3 megacycles?

Over-all amplification. Let us consider a three-stage amplifier, having resistances, tubes, etc., as shown in Fig. 4–14.

The first two tubes have an amplification factor of 30; the final or power tube has a μ of 3. What is the over-all voltage amplification? If the values of R_L, C, and R_g are properly chosen, the amplification per stage will be 75 per cent of the μ of the tube, and what the first tube amplifies will be amplified again by the second; that is, if an input of 0.1 volt is applied

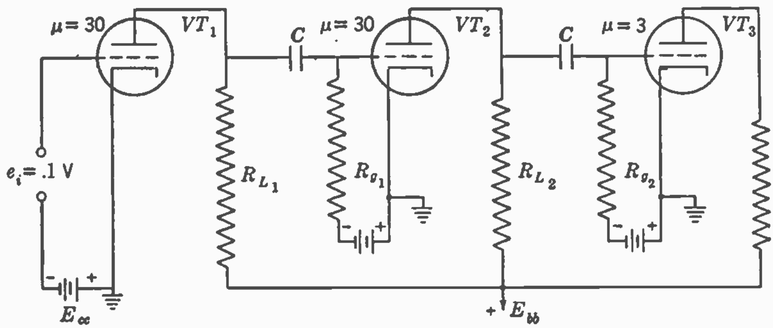


FIG. 4–14. A three-stage amplifier coupled with *RC* intertube networks.

to the grid-filament input of the first tube a voltage of $0.1 \times 0.75 \times 30$ or 2.25 volts will be applied to the second, and a voltage of $2.25 \times 30 \times 0.75$ or 50.5 volts to the third input; and if the resistance into which this tube works has twice the value of the plate resistance of the tube, two-thirds of the total a-c plate voltage of the last tube will be applied to the load. That is, the load will have across it a voltage equal to $50 \times 3 \times \frac{2}{3}$ or 100 volts.

The over-all voltage amplification in such a case will be $100 \div 0.1$, or 1000. This is equal numerically to the product of the voltage gain of the individual stages. Thus, if three stages have gains of A_1, A_2, and A_3, the first two will have a combined gain of $A_1 \times A_2$ and all three will have a gain of $A_1 \times A_2 \times A_3$. In this example it will be $(22.5)^2 \times 3 \times \frac{2}{3} = 1000$ (approximately).

Plate-voltage requirements. One of the objections to the resistance-capacitance coupled amplifier using triodes is the excessive plate voltages that are necessary. The voltage on the plate of a tube in whose plate circuit is a high resistance is not the voltage of the supply but this voltage minus the voltage drop across the resistor in the plate circuit. We want the a-c voltage drop across this resistance to be high but the d-c voltage drop to be low. As an example, suppose that a tube has a plate resistance, r_p, of 60,000 ohms when the plate voltage, E_p, is 100 volts. To get 75 per cent of the μ of the tube as the over-all amplification from grid input voltage to the voltage output across the resistor, we need a resistance in the plate circuit of 180,000 ohms. Suppose that under these conditions the plate current is 1 ma. What supply voltage is necessary if 100 volts is required on the plate?

The plate voltage may be found from

$$E_p = E_{bb} - I_pR_L$$

$$100 = E_{bb} - 0.001 \times 180{,}000$$

$$E_{bb} = 280 \text{ volts}$$

Now the tube will amplify with fewer volts on the plate than this, but the difficulty lies in the following fact. The plate resistance of a tube is a function of the plate voltage; that is, at low plate voltages the plate resistance is high, which in turn necessitates a higher value of load resistance to get the 75 per cent of the μ of the tube, which means that the voltage drop will be high across this resistor, and so on. A high-voltage plate supply is always needed.

To get 100 volts on the plate of this tube requires a plate supply of 280 volts. What can be done about it? One solution is to use a tetrode or pentode in which the amplification is not so dependent on plate voltage, and the other is to use a coupling system of high reactance but low resistance.

Inductance-load amplifier. Suppose that instead of the resistor in the plate circuit we substitute a low-resistance inductor with a high value of reactance at the frequencies at which the amplifier is to work, for example, a 100-henry choke coil with a d-c resistance of 1000 ohms. The circuit looks like Fig. 5–14. Now the direct plate current encounters no appreciable opposition in the 1000-ohm resistance. The alternating current, however, must flow through the high inductance and the resistance in series. Across this impedance $\sqrt{R^2 + \omega^2 L^2}$ will appear the amplified a-c voltage which may be used to drive another amplifier stage if desired.

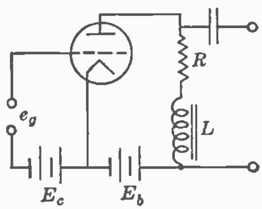


FIG. 5–14. Inductance-capacitance or impedance-coupled amplifier.

Such an amplifier is commonly called an **impedance** amplifier. The loss in d-c voltage between the plate supply and the plate is 1 volt per 1000 ohms per milliampere of current, and if the plate current is 2 ma only 102 volts of supply voltage will be required to put 100 volts on the plate of the tube. Nowadays, impedance-coupled amplifiers are used only where power tubes having high plate current are employed. The desire to lower the voltage loss can be accomplished here by lowering the d-c resistance, for example, by means of a choke of large wire.

Problem 2–14. An amplifier is to be used at audio frequencies. What must be the impedance of the choke coil at 100 cycles to secure an amplification of 7.5 from a tube whose μ is 10 and whose plate resistance is 15,000 ohms? What will be the amplification at 1000 cycles?

Problem 3–14. If the resistance of the choke coil used in Problem 2–14 is 500 ohms, what is the inductance required?

Problem 4–14. A tube has a load resistance of 100,000 ohms. The plate current is 2 ma and the plate voltage required is 90 volts. What plate supply voltage is necessary?

Problem 5–14. A tube draws 0.5 ma from the B batteries whose voltage is 180. A coupling resistance of 100,000 ohms is used. What is the voltage actually on the plate?

Miller effect. Not only do the tubes and wiring capacitances play havoc with high frequencies, but the effect of the grid-plate capacitance of the second tube is multiplied by the μ of the tube—a very interesting and unfortunate fact.

The capacitances which cause trouble at high frequencies are:

1. Grid-cathode capacitance, C_{gk}.
2. Plate-cathode capacitance, C_{pk}.
3. Grid-plate capacitance, C_{gp}.
4. Stray capacitances in wiring, etc.

The capacitance C_g across the input to the tube is equal to

$$C_g = C_{gk} + C_{gp}(\mu_o + 1)$$

where μ_o is the effective amplification in the circuit, and the other factors are as itemized above and shown in Fig. 6–14.

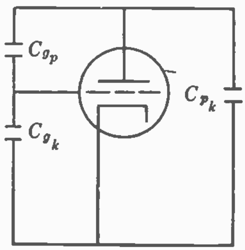


FIG. 6–14. Tube interelectrode capacitances.

These values for the UX-240, a tube adapted for resistance-coupled amplifiers, are $C_{pk} = 1.5$ $\mu\mu$f, $C_{gk} = 3$ to 4 $\mu\mu$f, and $C_{gp} = 8.8$ $\mu\mu$f. In practice C_g may vary from 20 to as high as 300 $\mu\mu$f, depending upon the intrinsic values of its components and the amplification of the system. That is, if the $C_{gk} = 3.0$ $\mu\mu$f, and $C_{gp} = 8.0$ $\mu\mu$f, and the effective amplification of the system is 20, C_g becomes 3 plus 8×21 or 171 $\mu\mu$f—which is an appreciable capacitance and not at all what one would expect. The effective capacitance across the input to the tube is a function of the μ of the tube; thus the greater the amplification factor the more trouble one gets into because of this capacitance-multiplying effect. Low-μ tubes are not so troubled, but their amplification is low; and so the amplifier designer and experimenter must compromise between a good frequency characteristic and a high gain. If the gain is good the high frequencies are discriminated against, and if the characteristic is very good the over-all gain is likely to be low.

Problem 6–14. The μ of a tube is 8, its grid-cathode capacitance is 6 $\mu\mu$f, its plate-grid capacitance is 8.0 $\mu\mu$f. The effective amplification of the circuit is 7, and stray capacitances across the cathode and grid circuit amount to 2 $\mu\mu$f. What is the effective shunting capacitance C_g?

Problem 7–14. Suppose that a 1-megohm resistor is in the plate circuit of a tube and that 10 volts is developed across it. If 100,000 ohms is shunted across it, the effective impedance becomes 0.91×10^5, and if the current is the same the voltage developed has been reduced to 0.91 volt. Now, suppose that 10 volts is developed across 100,000 ohms and that another 100,000-ohm resistor is shunted across it. What is the resultant reduction in voltage?

Problem 8–14. A tube with an amplification factor of 30 and a plate resistance of 150,000 ohms is used with a half-megohm coupling resistance. Suppose that in construction a path of soldering flux is placed across the terminals of this resistor accidentally so that the half megohm is effectively shunted by 100,000 ohms. What is the resultant amplification? What is the amplification without the shunting resistance? What is the percentage loss due to the flux?

Tuned-inductance amplifier. In the resistance-load and the inductance-load amplifier the maximum amplification is attained when the impedance in the plate circuit is a maximum. If, then, we desire to receive only a narrow band of frequencies, say around 1000 cycles, we can get greater amplification by tuning the inductance with a shunt condenser so that we have an anti-resonant circuit in the plate circuit of the tube, as shown in Fig. 7–14. If an inductance of 0.1 henry is tuned with a 0.254-μf condenser the resonant frequency will be about 1000 cycles. If the coil has a resistance of 100 ohms the impedance of the anti-resonant circuit will be roughly 3950 ohms whereas the reactance of the coil alone will be only 628 ohms. Tuning the coil therefore increases the load impedance in the plate circuit by about 6.3 times.

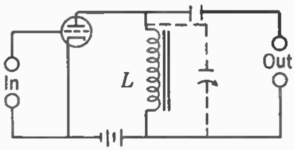


FIG. 7–14. Tuned inductance load impedance.

The untuned inductance-coupled amplifier may be used when a fairly wide band is to be amplified; if the inductance is tuned the band width will be smaller, and the higher the Q of the coil, the narrower will be the band amplified.

Practical high-gain amplifiers. The introduction of pentode amplifier tubes has made it very simple to get high voltage gains in single- or two-stage resistance-capacitance coupled amplifiers. Since the gain in a single stage may be 100 or more and since the gain of several stages is the product of the individual stage gains, two tubes plus the appropriate coupling system will produce a voltage amplification of 100 × 100, or 10,000. As much amplification can be secured with a single pentode amplifier as is possible with two stages of triode resistance-coupled amplification, with corresponding simplicity of apparatus and saving of space.

The over-all frequency characteristic is also the product of the characteristics of the individual stages. Thus, where the

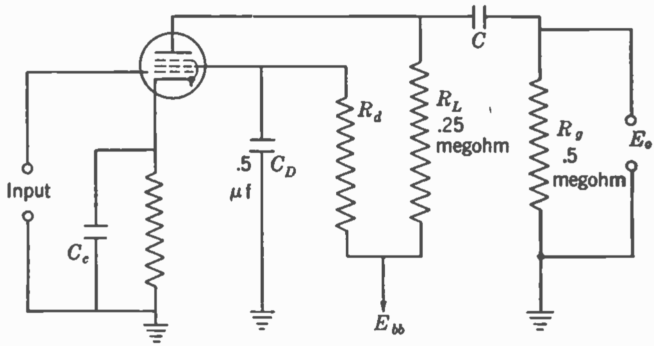


FIG. 8–14. Typical resistance-coupled amplifier using a pentode tube.

curve is flat, all frequencies being amplified alike, the resultant curve of the complete amplifier will be flat, but, at the lower and higher frequencies where the amplification tends to fall off, the characteristic of a two-stage amplifier will be twice as bad as that of a single-stage amplifier. It is often advisable to sacrifice high gain per stage to secure a flatter response per stage.

A typical single-stage amplifier is shown in Fig. 8–14. Higher gain will be secured with higher values of R_L but at the same time the gain at higher frequencies falls off sooner. For example, if R_L is 0.1 megohm the response will be good to 20,000 cycles; but if it is 0.5 megohm the response will begin to fall off at 5000 cycles. The peak gain, however, will be higher with the 0.5-megohm resistor.

The peak voltage output may be estimated as roughly one-fourth of the plate battery voltage if the grid leak resistance to the following stage is of the order of 0.5 to 1.0 megohm and R_L is 0.25 megohm. For the average pentode amplifier stage the distortion under these conditions will be about 3 per cent.

Amplifier instability. The larger the coupling capacitor, the better the low-frequency response. But there are difficulties in this direction. If, for any reason, the grid of the following stage collects some electrons, the coupling condenser becomes charged.

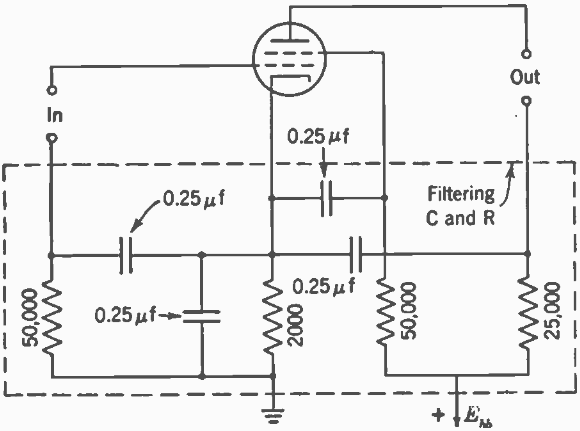


FIG. 9-14. Addition of shunt capacitances from screen, grid and plate load to cathode and series resistances for filtering.

This charge must leak off to ground through the grid leak resistance. If the product of C and R_g is so great as to cause appreciable time for this charge to leak off, the second tube may "block," that is, it may cease operating because of plate current cut-off. A practical value for this RC product is 0.05, and therefore there is a practical limit to the low-frequency response that is compatible with high gain and good stability.

If two or more stages secure their plate and screen potentials from a common source of positive voltage, as from a rectifier-filter system, care must be taken to see that each of these individual circuits is filtered so that none of the a-c components flow through the power supply system. This filtering is accomplished by connecting by-pass condensers to cathode directly from the screen and from the lower end of the plate load. Generally additional series impedance is used in these connections to force the a-c components to stay out of the power supply system. Otherwise instability results since some of the output alternating current can produce a voltage across the power supply circuit which can be re-introduced into a previous plate circuit.

The transformer-coupled amplifier. The output a-c voltage across a resistance or untuned inductance load can never be higher than the input grid voltage multiplied by the μ of the

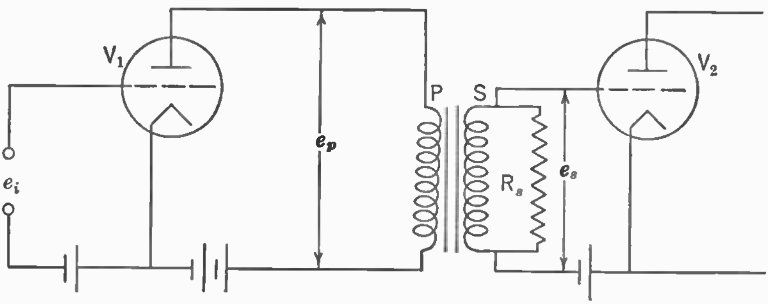


FIG. 10–14. Transformer-coupled amplifier.

tube, and can attain that value only when the resistance or impedance is high compared with the plate resistance. Suppose, however, that we use a transformer, as in Fig. 10–14. The voltage across the secondary will be increased by the turns ratio of the windings, and so the voltage developed in the plate circuit of the tube may not only be passed on to a following tube multiplied by the μ of this tube but may also be multiplied by the turns ratio.

If the secondary circuit takes no current or power the greatest voltage will appear across the secondary when a very high turns ratio is used, but this is not true when power is taken—and some always is.

The maximum power in the secondary circuit will be obtained when the turns ratio is given by the expression

$$N^2 = \frac{R_s}{R_p}$$

where R_s and R_p are the resistances between which the transformer works. With such a turns ratio the voltage across the secondary is given by

$$e_s \text{ (max)} = \frac{\mu_1 e_{g1} N}{2} \qquad [1]$$

All this assumes that a perfect transformer is used, that is, one which has no d-c resistance, no magnetic leakage, and infinite primary and secondary reactances. If the resistance of the load across the secondary is 1 megohm, and the plate resistance of the tube out of which the transformer works is 10,000 ohms, the proper turns ratio

$$N = \sqrt{\frac{1{,}000{,}000}{10{,}000}} = 10$$

and the voltage gain from equation 1 is

$$\frac{e_s}{e_{g1}} = \mu_1 \times 5 = 40 \text{ if } \mu_1 = 8$$

The foregoing mathematics discloses several interesting points. In the first place it is possible to get much greater voltage amplification by means of a transformer than can be obtained with the same tube and either resistance or inductance output. In the second place, for every ratio of resistance across the secondary and primary of the transformer there is a certain turns ratio which will produce the maximum voltage step-up. This means that, once the resistances on either side of the transformer are determined, the turns ratio for maximum voltage gain is fixed. This, in turn, means that a transformer can be used as an impedance-adjusting device, that is, a coupling device between two circuits of different impedance, one of which acts as a source and the other is the recipient of a voltage.

Whenever the impedances of the two sides of the transformer are not the values indicated by the expression above, there is a loss in secondary power, a transmission loss, engineers say. Whenever two circuits are to be coupled together with the least loss in voltage or power the proper turns ratio transformer must be used, that is, $N^2 = Z_s/Z_p$, where Z_p and Z_s are the impedances between which the transformer works.

The advantage of the transformer. The transformer has the great advantage that it can contribute toward the voltage amplification, and can contribute toward the maximum power output when the load resistance or impedance differs from the tube resistance. In addition, high plate supply voltages are not necessary.

A transformer can be constructed so that it has little loss. Transformer losses are due to the d-c resistance of the windings, to the fact that perfect coupling between primary and secondary is not attained, and to iron losses. The value of the use of transformers is of two sorts: first, transformers may contribute toward the voltage gain by producing a voltage step-up; second, they may be used to "match" the load resistance to the tube resistance, or vice versa. Let us suppose that the tube resistance is 10,000 ohms, and that the secondary load resistance is 100,000 ohms. A resistance of this value interposed in the plate circuit of a tube will not have the maximum power developed in it. But if we use the proper transformer such that N^2 equals 10, this 100,000 ohms will look, to the tube, like a resistance of 100,000 $\times$ 0.1 or 10,000 ohms—the condition for maximum voltage and maximum power.

The transformer can always be forgotten if we substitute for it and its secondary load, Z_p, this load divided by N^2. From the secondary side, the transformer looks like the primary circuit impedance, Z_p, multiplied by N^2. These statements are true provided that a step-up in voltage occurs in going from primary to secondary. If a step-down occurs, these expressions must be reversed.

Example 1–14. A transformer with a turns ratio of 3 connects a 10,000-ohm tube with a load which has a frequency characteristic such that at 100 cycles the load has one-tenth of the impedance it has at 1000 cycles. The load impedance at 100 cycles is 90,000 ohms. The μ of the tube is 8. What are the voltage amplification and the power output at 100 and 1000 cycles?

The transformer may be dispensed with in the calculation if we transfer the secondary load directly into the plate circuit of the tube by multiplying it by $1/N^2$, that is, 1/9.

Thus at 100 cycles the impedance in the plate circuit is

$$90{,}000 \times \tfrac{1}{9} = 10{,}000 \text{ ohms}$$

and the voltage amplification is

$$\frac{\mu \times R_L}{R_L + r_p} = \frac{8 \times 10{,}000}{20{,}000} = 4$$

and the power

$$P_o = \frac{\mu^2 e_g{}^2 R_L}{(R_L + r_p)^2}$$

$$= e_g{}^2 \times 1600 \times 10^{-6} \text{ watt}$$

At 1000 cycles

$$R_L = 900{,}000 \times \tfrac{1}{9} = 100{,}000 \text{ ohms}$$

and

$$G = \frac{8 \times 100{,}000}{110{,}000} = 7.3 \qquad \text{and} \qquad P_o = e_g{}^2 \times 533 \times 10^{-6} \text{ watt}$$

Because of the ease of securing high gain with uniformity over wide bands by means of high-resistance tubes and resistance-capacitance coupling, transformers are not so widely used as coupling agents as formerly. By taking special care in construction to keep down the distributed capacitance and by the use of special alloy steels for the core material, it is possible to make transformer-coupled amplifiers with reasonably high gain quite uniform in response over a 60- to 20,000-cycle range. Triodes are required to secure low plate resistances, which enable the use of interstage transformers of turns ratios greater than unity.

Transformers are a necessity when the amplifier is finally connected to its load since the load is almost always lower in impedance than the output tube or tubes. These are known as **output transformers** and are designed to connect a single power tube to a loud speaker or to a 500-ohm line, or push-pull tubes to speaker or to line. They may have appreciable direct current in the primary and must cover a wide frequency range with high alternating currents in both primary and secondary. It is not an easy matter to make a good output transformer, especially of the single-tube type.

Problem 9-14. A tube whose plate resistance is 12,000 ohms works into a resistance, R_L, of 600,000 ohms. What is the proper turns ratio of the output transformer, and what is the voltage gain if $\mu = 8$? If $e_{g1} = 1$ volt what voltage will appear across R_L?

Problem 10-14. One tube whose plate resistance is 25,000 ohms is connected to another whose grid circuit has a resistance of 400,000 ohms. What is the turns ratio of the transformer for maximum voltage amplification?

Problem 11-14. An output transformer is to be used to connect a loud speaker to a tube. The impedance of the loud speaker at the desired frequency is 4000 ohms; the tube has a resistance of 2000 ohms. What is the proper value of N?

Problem 12-14. A telephone line has an impedance of 600 ohms. The a-c voltages in the plate circuit of a 6000-ohm tube are to be transferred to this line. What must be done to effect such a transfer with least power loss?

Problem 13-14. Assuming no loss in the transformer, what is the power transmitted to a 6000-ohm load from a 2000-ohm tube when they are coupled by a transformer whose load winding has 2.24 times as many turns as the primary? Assume $\mu = 3$, $e_g = 10$ volts rms. Use $P = \frac{(\mu e_g)^2 R_L}{(R_L + r_p)^2}$. What would be the value of N for maximum power in the load? What would be the power then? What value of N would deliver maximum undistorted power output into the load? What is this value of power?

The push-pull amplifier. Where considerable sound output is desired with maximum tone fidelity, as in high-quality receivers or in public-address systems, two power-amplifier tubes are connected in a push-pull circuit shown in Fig. 11-14. The advantages of this circuit are several, as outlined below.

Distortion due to second harmonics is reduced to a low figure compared to a single tube; the power output is at least twice that obtainable with a single tube; because of the elimination of troublesome second harmonics the input voltage can be raised somewhat with correspondingly greater output; there is no direct current in the output transformer which can therefore have less iron; there are no hum voltages in the output, and therefore less filtering of the plate supply is necessary; there are no currents corresponding to the frequency of the signal in the center tap which supplies plate voltage, and therefore these signal currents cannot get into the plate supply system.

Push-pull operation is not limited to any type of tube. Either low-μ triodes of the 2A3 type, or pentodes like the 6F6 or 6L6, may be operated in this manner. Frequently a combination of the push-pull and parallel operation of tubes is utilized; two tubes in parallel are on each side of the amplifier, making four tubes in all. The two parallel tubes give twice the output of a single tube, and the push-pull setup of these two sets of parallel tubes multiplies the output by another factor of two or more. Therefore at least four times the power output is secured. The same result could be attained by substituting a single tube of double the rating for the parallel tubes, but such single tubes

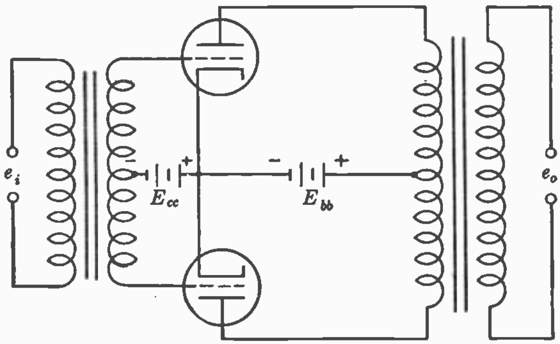


FIG. 11–14. Push-pull amplifier.

usually require high plate voltages, are large, and increase the expense of the power supply considerably.

The push-pull amplifier circuit is shown in Fig. 11–14. It consists of a center-tapped input transformer, two tubes of identical characteristics, and an output transformer, or a choke, with a center tap. When a voltage is induced into the secondary winding of the input transformer, one grid becomes less negative a certain amount and the other grid becomes more negative the same amount.

If the C bias which is attached to the center of the input winding is 35 volts and if across the entire secondary winding appears a peak a-c voltage of 20, one tube has its negative grid voltage increased by 10 volts, or to 45 volts, and the other decreased by 10 volts, or to minus 25 volts. Therefore one plate current increases and the other decreases. These variations in alternating plate current flow through the output winding and may be transferred to the load. These plate current variations are out of phase by 180 degrees, one increasing, the other decreasing a like amount. One tube "pushes" current through the output, the other "pulls" current through it. Suppose both tubes pushed at the same time. What would happen? If they both pushed the same amount and at the same time no current would flow through the output winding because the two currents would neutralize each other.

If, then, we can cause the second harmonic currents to be in phase, that is, to push or pull at the same instant, and the same amount, they will not get into the load, and distortion due to these extra currents will not appear in the loud speaker.

Push-pull characteristics. Distortion is caused in an amplifier because the tube characteristic is not straight but has a curve in it near the bottom. This curve is such that even-numbered harmonic currents will either push or pull simultaneously. Whereas the fundamental and odd harmonics decrease in one tube and increase in the other as the grid is excited and so cause an effective voltage to appear across the output transformer windings, the even harmonics decrease and increase together and effectively balance each other out across the transformer. They do not appear in the load.

All this hinges on the supposition that the two tubes are alike, that is, have identical characteristics. If they do not have identical characteristics, some even harmonics will appear in the output.

Because the input voltage across such an amplifier is divided into two parts this amplifier requires twice the input voltage to give the same output power—unless the turns ratio of the input transformer is doubled, which is difficult to carry out in practice. Because of the push-pull arrangement, however, considerable overloading can be tolerated before the third harmonics (which are not canceled out) become objectionable. There is another advantage—the output transformer or choke need not have such a large core because of the fact that the direct currents in the two halves are flowing in opposite directions, and, since the two windings are closely coupled, the resultant magnetization of the core is very low. Since the two windings are connected, "series aiding" so far as alternating currents are concerned, the total inductance is increased. Not only less iron but also less copper is necessary. The proper match between the tubes and the load can be obtained by means of such a transformer.

The tube resistance of such an amplifier is double that of the single tube; therefore when worked into a low-impedance load an output transformer must be used so that maximum undistorted power is fed to the load. In other words, the plate load must be matched to the plate resistance of the amplifier by means of an appropriate output transformer.

The push-pull amplifier, then, is a device for eliminating the even-harmonic distortion which occurs when tubes are worked too far down on the curved part of their characteristic. Since the distortion is inherently less, the tubes can be worked harder, having greater input voltages impressed on them, with consequently greater output.

Class B amplification. Suppose, for example, that we increase the steady C bias to the two push-pull tubes so that the plate current is near zero. Now there will be little or no plate-current flow until the input a-c signal overcomes this d-c bias. For this reason there will be an appreciable fraction of the a-c cycle in which no plate current flows.

It is also possible to drive the grids of these tubes positive so that they actually consume power. This power must be supplied by the circuit from which the two tubes are driven.

Such an amplifier is known as class B. Very high efficiency can be secured from small tubes with medium plate voltage. The two tubes must be well matched, and frequently are in the same envelope.

Two tubes must be used; they are connected just as any other push-pull amplifier. Sometimes no C bias is used, the tubes being specially built with a high μ so that with no bias the plate current is small. When the excitation is applied little or no current flows in the plate circuits except on the half cycles which make the grid positive.

Where the ordinary push-pull circuit, connected and biased for class A operation, draws plate current in both halves at all times, one tube taking more current when the other takes less from the plate supply source, class B tubes draw current in spurts. When one tube grid is going negative the other is going positive. The latter tube draws current; the other tube does not.

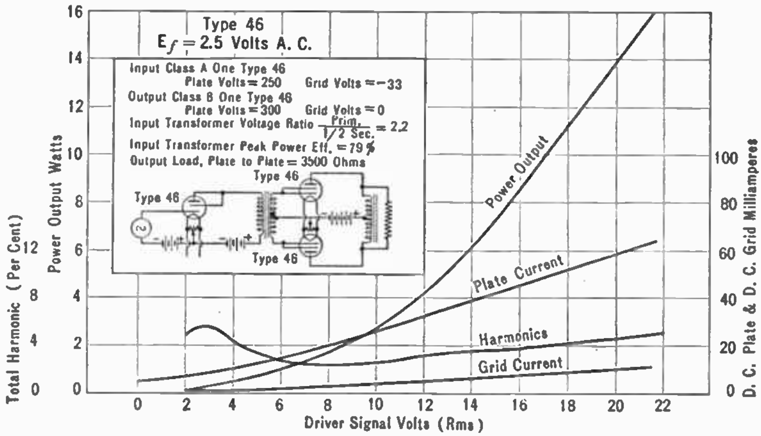


FIG. 12–14. Characteristics of class B amplification.

The amount of current each tube draws depends upon the exciting grid voltage. The power output of such a pair of tubes may be very high; the distortion may be made quite low by careful design. Since the current taken from the plate supply system varies with the exciting voltage applied to the tube grids, the voltage regulation of this supply must be very good; that is, the terminal voltage of the filter must not drop appreciably when a sudden large current is drawn from it. For this reason a mercury-vapor rectifier is usually employed.

Since the grids of the two tubes are driven positive, they will draw current. This flow of current represents power, and this power must be supplied by the previous tube. The previous tube is usually called a "driver" and is connected to the class B tubes through a *step-down* input transformer. Distortion present in the driver stage will go through the power stage to appear in the output. Therefore this stage must be carefully designed. It may be "single-ended" or push-pull.

Phase inverters. A push-pull amplifier requires that the two tubes have their grids excited 180 degrees out of phase with each other, one increasing in voltage while the other decreases. It is easy to feed such an amplifier from a single-ended amplifier by means of an input transformer since the secondary can be divided in the center. But, if the preliminary tube is a single tube and the coupling is to be by resistance, how is one to split the output so that the two grids are excited in opposite directions?

Very clever circuits known as phase inversion circuits have been developed for this purpose. In Fig. 13–14 is a simple circuit in which the single-ended amplifier A feeds excitation to a power tube through a conventional RC network. Part of the output of this driver stage A is tapped and is used to drive the

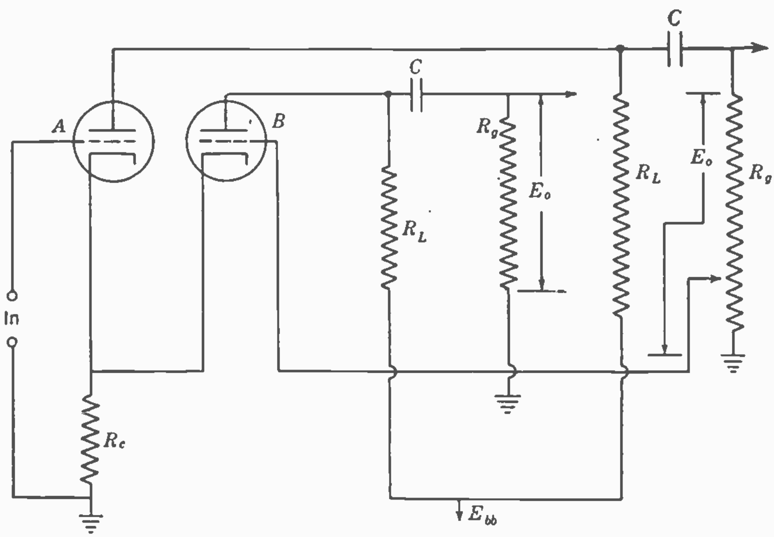


FIG. 13–14. Phase inverter (B) tube. The outputs of A and B are used to drive two push-pull tubes.

grid of a phase inverter tube *B* whose output is employed to excite the grid of the second of the two push-pull stages. Since the plate and grid alternating voltages of tube *A* are out of phase by 180 degrees, the voltage derived from the output of *A* and applied to the grid of *B* is out of phase with the grid of *A* by the required 180 degrees. The voltage taken from the output of the driver must be such that when amplified by tube *B* it becomes equal to the voltage applied to the first of the push-pull tubes. This output voltage of the phase inverter then is used to drive the second push-pull tube.

Feedback in amplifiers. When part of the output voltage of an amplifier is re-introduced to the input it is said to be "fed back" and the mechanism is called "feedback." Feedback has both advantages and disadvantages, depending upon how it is done.

For example, let us suppose that an amplifier has a gain of 10 and that one-tenth of the output is fed back in phase with the input. Now there is as much voltage on the grid due to feedback as there is due to excitation from the external source. This additional voltage is amplified and reappears in the output. Again 10 per cent is fed back to the input. After a few cycles of such a process the amplifier is overloaded and it will probably break into uncontrollable oscillation. The amplifier, then, is said to "sing" or howl because it will probably emit an audible tone whose frequency is determined by the constants of the elements of which it is constructed.

If, on the other hand, the 10 per cent of the output is fed back into the input out of phase with the original excitation voltage, there is now no effective excitation, and no matter how much is put into the amplifier, nothing comes out.

In the first case the process is called positive feedback or regeneration, and in the second case it is called **inverse** or **negative feedback** or **degeneration.** These extreme situations would not occur in practice except by bad design. The effect of regeneration, even when under control, is to exaggerate any non-uniformity in the frequency-gain characteristic since the feedback is on a percentage basis. Thus any increase in gain at some frequency would produce more excitation at the input than the gain at some other frequency. This increase would be emphasized and reamplified, and thus the characteristic curve of such an amplifier would get worse the more regeneration was permitted. The frequency at which the maximum increase in amplification, due to regeneration, occurs is the frequency at which the amplifier will finally howl unless the regeneration is kept under control.

Application of negative feedback. Let us feed a voltage wave (e_g) to the input of an amplifier. If some non-linearity exists

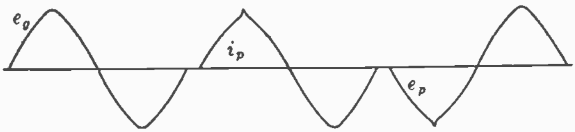


FIG. 14–14. When a voltage wave e_g goes through an amplifier, distortion is produced, so that the plate current i_p is unsymmetrical. Feeding back to the input some of this distorted output (e_p) may reduce or overcome the distortion.

in the amplifier, the plate current will not look like e_g. Let us suppose that it looks somewhat like Fig. 14–14. A hump that did not exist in the original signal has been added to the input wave form. There has been wave-form distortion. Now let us feed back to the input some of the output signal e_o. If this feedback is in a positive direction, the hump will be increased; but if it is added to the input in a negative direction the hump will be decreased. The application of negative feedback has reduced distortion.

Negative feedback will not reduce distortion produced previous to the amplifier in which the feedback is applied; it will only correct what occurs in the amplifier itself.

The use of inverse feedback has the following purposes: (1) to reduce frequency and wave-form distortion; (2) to reduce variation in amplifier gain caused by variations in power supply voltage, variations in circuit components, and variations due to substituting new tubes for old ones; and (3) to reduce noise and hum produced in the amplifier. Also, some types of feedback reduce the effective resistance of the amplifier tube so that the load which the tube works into is effectively increased. This tends to reduce distortion since it straightens out the tube characteristic. In addition, if the tube works into a load of varying impedance, such as a loud speaker, these variations have less effect upon the frequency characteristic and produce less distortion.

Negative feedback will not correct distortion occurring ahead of the point at which the feedback is applied. It may be applied, however, over several stages so that an entire multistage amplifier may secure the benefits.

Practical aspects of feedback. If the tube already has a low resistance, for example the 2A3 power tube, little advantage is to be gained by the use of feedback, for the low plate resistance, 800 ohms, keeps distortion due to loud-speaker-impedance variations to a low value.

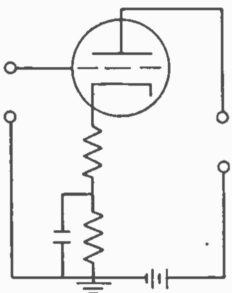


Fig. 15–14. Feedback circuit in which part of the cathode bias resistance is not by-passed.

If the cathode bias resistor of an amplifier is not by-passed, negative feedback is produced, owing to the fact that alternating currents in the plate circuit flow through this resistor and there produce alternating voltages that are applied to the input in series with the original signal. They are opposite in phase to the input signal. This form of feedback will reduce distortion but will increase the plate resistance of the tube. Sometimes part of the cathode resistor is by-passed and part is not. In this way the degree of feedback can be controlled.

Application of feedback reduces the amplification to the same degree that it reduces distortion.

Feedback, therefore, in the proper phase and proper degree will improve the frequency characteristic and the stability of an amplifier and will reduce distortion and hum and noise which come from the amplifier itself. At the same time the effective amplification of the amplifier is reduced by the same extent as the advantages are secured. A perfect amplifier, then, would have no distortion and would be absolutely stable, but it would not amplify. We must temper our improvements with some discretion. Amplification with modern tubes is easy to get—easier than freedom from distortion. The loss in gain can easily be made up.

There are many useful feedback circuits, all ingenious. Usually, part of the output voltage of a tube is deliberately intro-

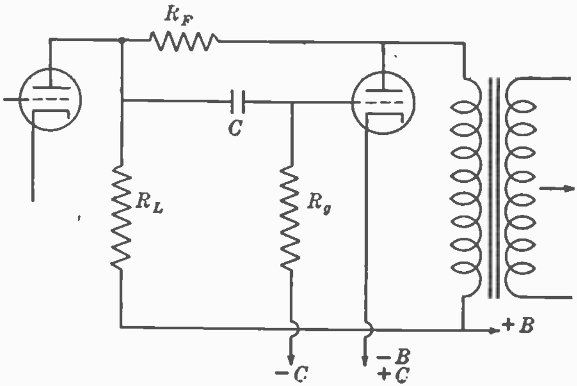


FIG. 16–14. Feedback introduced by R_F.

duced in the input by means of a resistance and capacitance as shown in Fig. 16–14.

Cathode follower. A circuit much used in wide-band circuits, such as video amplifiers, is the cathode follower shown in Fig. 17–14. Here the load impedance is in the cathode circuit and as such is also part of the input circuit. Thus there is a degenerative feedback. The advantages of this circuit over the conventional method of taking the output voltage from a resistor in the plate circuit are: (1) one of the load terminals can be grounded; (2) the impedance presented to the circuit from which the circuit derives its driving voltage is higher; (3) the output impedance is lower and can, therefore, be matched to low-impedance transmission lines with greater ease, and also can have better high-frequency characteristics.

The voltage output, that is the alternating voltage developed across the load resistance, is almost equal to the input voltage applied to the tube. This means, simply, that there is no amplification in the circuit, which merely serves as a coupling device or transformer between a high-impedance source and a low-impedance load.

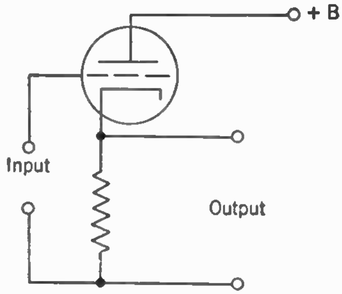


Fig. 17–14. Cathode follower circuit.

Regeneration in amplifiers. Feedback in the wrong sense may occur in several ways. Insufficient shielding between input and output of a high-gain amplifier may permit voltage to get back to the input. This may cause instability or even singing. Plate and screen currents (a-c) must not be permitted to flow through the positive voltage supply common to all tubes. This may be prevented by adequate filtering (called **decoupling**), as shown in Fig. 18–14. Here shunt capacitances from the lower parts of the grid and plate circuits are connected directly to the cathode

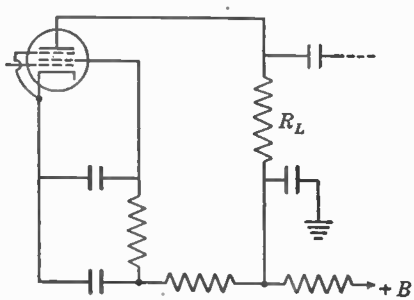


Fig. 18–14. Use of series resistance and shunt capacitance to keep alternating current out of a power supply. See also Fig. 9–14.

of the tube; series resistances are connected between these points and the voltage sources. Alternating currents in these circuits find an easy path back to cathode and a difficult path to cathode through the voltage source. Therefore they are kept out of the plate or bias voltage supply system.

The decibel. An amplifier may be described in several ways. For example, we may state that it has a voltage amplification, or gain, of 160 times. This tells us nothing about the amount of power this amplifier can transmit to an antenna or to a loud speaker. Or we can state that the amplifier has an output of 160 watts. This tells us nothing about the voltage required on the input to deliver this amount of power to the load.

The voltage amplification and power output are useful facts to have; but when we wish to compare two amplifiers, or two sounds of different intensity, we must look into the matter a bit more closely. Let us suppose that we have an amplifier which is delivering 800 watts to some sort of load and that this amount of power is not enough. We must increase it. By how much must we increase 800 watts so that the ear can tell the difference? Suppose that we add 100 watts. This is quite a bit of power; but the ear would not be able to note any difference. Let us add another 100 watts. Now the ear tells us that we have increased the volume of sound somewhat over the original value of 800 watts.

It is a fact that a given volume of sound must be increased approximately 25 per cent before the ear can note any difference. Having made this increase, we must increase the new value another 25 per cent before any further increase in sound can be detected. Through a wide range in sound intensities, we must add equal increments to each succeeding intensity before the human ear will tell us that a change has been made.

Would 10,000 people shouting sound 100 times as loud as 100 people shouting? The answer is no—the greater sound will actually impress us as being only about 20 times as loud.

If two sounds bear an intensity ratio to each other of 1.25, the ear will tell us that one is louder than the other. Now this is true for a very wide range of absolute values; that is, if two sounds of 100 and 125 watts are compared by the ear, the same difference in loudness will be noted as if the two sounds were actually 1.00 watt and 1.25 watts, in spite of the fact that in one case we have made an increase of 25 watts and in the other we have added only 0.25 watt. Clearly it is not the absolute values that count, it is the ratio.

Now we might use this ratio of 1.25 as our loudness unit. To determine the relative loudness of two sounds, we would have to determine the number of times 1.25 would have to be multiplied by itself to go into the ratio of the power outputs of the two amplifiers. Then we would know that one was a certain number of times louder than the other. This is a bit awkward, and so a simpler system has been evolved.

The decibel is a unit which approximately expresses this important ratio of 1.25. Each time we increase the power of an amplifier, or the loudness of a sound, by a ratio of 1.25, we say that we have added 1 decibel. Here we have one important advantage of the decibel as a convenient unit—we can add decibels, whereas actual powers or loudnesses must be multiplied.

If two sound intensities P_1 and P_2 (or the power outputs of two amplifiers) are to be compared according to the ability of the ear to detect intensity differences, we may determine the number of decibels which expresses the relative value of the two intensities by

$$N_{db} = 10 \log_{10} \frac{P_1}{P_2}$$

where P_1 is greater than P_2.

The factor 10 comes into this picture because the original unit devised was the **bel,** which is the logarithm of 10 to the base 10 and which represents the audible difference between two powers one of which is 10 times as great as the other. The decibel is one-tenth of a bel, and, therefore, there are 10 times as many decibels in any expression involving the relation between two sound intensities as there are bels.

The decibel is a logarithmic unit. Each time the amount of power of an amplifier, for example, is increased by a factor of 10 we have added 10 decibels (abbreviated db). The table below gives some easily remembered values of decibels and their corresponding power ratios.

N_{db}	Approximate Power Ratio
3	2.0
4	2.5
6	4
7	5
9	8
10	10
20	100
23	200
30	1000

To determine the number of decibels by which two powers differ, we must first determine the ratio of the two powers, then we look up this ratio in a table of logarithms to the base 10 and then we multiply the figure obtained by a factor of 10. Now let us consider the relative sound intensities produced by 100 persons shouting and by 10,000 people shouting. First we determine the ratio, 10,000 ÷ 100 or 100.

The logarithm of any number, A, to the base 10 is merely the number of times 10 must be multiplied by itself to be equal to the number A. In the example above, 100 represents 10 multiplied by itself, and the logarithm of 100 to the base 10, therefore, is 2. The number of decibels expressing the relative loudness of 10,000 people shouting compared to 100 is

$$
\begin{aligned}
N_{db} &= 10 \log_{10} (10{,}000 \div 100) \\
&= 10 \log_{10} 100 \\
&= 10 \times 2 \\
&= 20
\end{aligned}
$$

Now let us see what happens if we double the number of people to 20,000.

$$
\begin{aligned}
N_{db} &= 10 \log_{10} (20{,}000 \div 100) \\
&= 10 \log_{10} 200
\end{aligned}
$$

Now enters an aspect of logarithms which must be kept in mind constantly. It must be remembered that we are interested in the number of times 10 must be multiplied by itself to produce the given number. We know that 10 multiplied by itself produces 100, and when 10 is multiplied by itself three times, 1000 is the result. Clearly 200 is produced by multiplying 10 by itself somewhere between 2 and 3 times. Here is where the logarithm table comes in handy. What do we do? We look up 2 (which is the significant part of 200) in the log table. This gives us 0.3. Therefore we state that the logarithm of 200 to the base 10 is 2.3. Here the 2 indicates that the original number is somewhere between 100 and 1000; and 0.3 gives us the exact place where 200 falls between 100 and 1000.

To arrive at the correct figure, we must use the log table and use our head at the same time. We must determine for ourselves what the number before the decimal point is—in this example, 2—then we use the table to give us the figure after the decimal point.

Therefore

$$\begin{aligned} N_{\text{db}} &= 10 \log_{10} 200 \\ &= 10 \times 2.3 \\ &= 23 \end{aligned}$$

Voltage and current ratios. Strictly speaking, the decibel should be used only when expressing the ratios of powers. Let us suppose that two amplifiers are feeding current into equal resistances. The currents are different. How can we express in decibels the advantage of the one as a current amplifier? We need only find out the ratio of the powers and multiply the logarithm of this ratio by 10. Thus:

$$P_1 = I_1^2 R$$

$$P_2 = I_2^2 R$$

$$\begin{aligned} \text{db} = 10 \log \frac{P_1}{P_2} &= 10 \log \frac{I_1^2 R}{I_2^2 R} \\ &= 10 \log \frac{I_1^2}{I_2^2} \\ &= 20 \log \frac{I_1}{I_2} \qquad [2] \end{aligned}$$

If the resistances are not equal, equation 2 becomes

$$\text{db} = 20 \log \frac{I_1\sqrt{R_1}}{I_2\sqrt{R_2}} = 20 \log \frac{E_1/\sqrt{R_1}}{E_2/\sqrt{R_2}} = 20 \log \frac{E_1\sqrt{R_2}}{E_2\sqrt{R_1}} \quad [3]$$

The factor 20 arises from the fact that when one squares a number the logarithm is doubled. For power ratios, the decibel is 10 times the logarithm; for current or voltage ratios the decibel is 20 times the logarithm of the ratio.

Voltage or current ratios can be translated into decibels only when the impedances into which the current flows, or across which the voltage exists, are taken into account. If these impedances are equal for both currents or both voltages, they cancel out, one being in the numerator and one in the denominator, but. in general, they do not cancel out and must be considered.

The decibel is always an expression for a ratio. We cannot speak of an amplifier that has an output of so many decibels, but if we assign some arbitrary level—say 10 mw—and compare all amplifiers to this amount of power we can say that one has 20 db or 100 db greater output, or less output, or is "up" or "down" 20 or 100 db compared to 10 mw.

Example 2–14. An amplifier has 1 volt applied to its input resistance of 10,000 ohms. Across its output resistance of 4000 ohms appears a voltage of 40. What is the power gain in decibels? The voltage gain? Would it be worth while to increase the output voltage from 40 to 50 volts?

Solution.

$$\text{Power input} \quad P_i = \frac{E_i^2}{R_i} = \frac{1}{10{,}000} = 10^{-4} \text{ watt}$$

$$\text{Power output } P_o = \frac{E_o^2}{R_o} = \frac{40^2}{4000} = \frac{1600}{4000} = 0.4 \text{ watt}$$

$$\frac{P_o}{P_i} = \frac{0.4}{10^{-4}} = 4 \times 10^3 = 4000$$

Power gain $= 10 \log 4000 = 36$ db (because the log of 4 is 0.6 and the log of 1000 is 3 and the log of 4000 is 3.6)

$$\text{Voltage gain} \qquad = 36 \text{ db} = 20 \log \frac{E_o\sqrt{R_i}}{E_i\sqrt{R_o}} = 20 \log \left(40 \frac{\sqrt{R_i}}{\sqrt{R_o}}\right)$$

$$20 \log \left(40 \times \sqrt{\frac{10000}{4000}}\right) = 20 \log 40 \times \sqrt{2.5}$$

or

$$20 \log 40 \times 1.6 = 20 \log 64 = 20 \times 1.8 = 36 \text{ db}$$

If E_o becomes 50 volts,

$$P_o = \frac{50^2}{R_o} = \frac{2500}{4000} = 0.625$$

The gain due to this increased output over P_o (above) is

$$10 \log \frac{0.625}{0.400} = 10 \log 1.56$$
$$= 2.0 \text{ db (approx.)}$$

Thus the gain due to increasing the output from 40 to 50 volts, or from 400 to 625 mw, will be audible to the ear, but the difference is not worth a great deal of effort to attain.

The solution of the above example is characteristic of the solutions of all such problems. Given the power ratio it is only necessary to look up the logarithm of this ratio to get the decibel gain. The student must not forget that all numbers between 100 and 1000 have as the first digit of their logarithms the number 2. Hence all power ratios between 100 and 1000 lie between 20 and 30 db. Multiplying any power by 10 represents a gain of 10 db. Thus, of two amplifiers having outputs of 50 and 500 watts, the 500-watt amplifier is said to be 10 db better than the 50-watt one. A loss of 10 db means that the power in any circuit has been divided by 10. If it is decreased or increased by 100 times the loss or gain in decibels is 20.

Example 3–14. A certain amplifier has a characteristic such that at 100 cycles its amplification in voltage is 8, at 1000 cycles it is 80, and at 6000 cycles, where the amplifier tends to "sing," the voltage amplification is 200. Are these differences appreciable to the ear?

Let us take the amplification at 1000 cycles as a zero level and find out how much above or below this level the other frequencies are. At 100 cycles the voltage ratio is 80/8 or 10. At 6000 cycles the voltage ratio is 200/80 or 2.5. At 100 cycles there is a loss; at 6000 cycles there is a gain. Thus:

$$\text{Loss at 100 cycles} = 20 \log \frac{80}{8} = 20 \log 10 = 20 \text{ db}$$

$$\text{Gain at 6000 cycles} = 20 \log \frac{200}{80} = 20 \log 2.5 = 8 \text{ db}$$

Such a characteristic indicates a poor amplifier. The low notes would be totally lost, and high ones are out of balance.

Example 4-14. In a certain circuit there is a loss of 25 db. What power ratio corresponds to this loss?

Power ratios of 10 = 10 db, 100 = 20 db, and 1000 = 30 db. Therefore the power ratio of 25 db lies somewhere between 100 and 1000. The figure 2 of 25 db tells us that the loss is somewhere between 100 and 1000 times. The figure 5 of 25 db is 10 times the logarithm of 3.1, and so 25 db corresponds to a power ratio of 310.

The solution of such a problem is as follows:

$$25 \text{ db} = 10 \log_{10} \frac{P_1}{P_2}$$

$$2.5 = \log \frac{P_1}{P_2} \quad \text{(dividing both sides by 10)}$$

$$\frac{P_1}{P_2} = \text{antilog } 2.5$$

$$= \text{antilog } 2.0 \text{ times antilog } 0.5$$

$$= 100 \times 3.1 = 310$$

If the loss were a voltage loss of 25 db the solution would be:

$$25 \text{ db} = 20 \log \frac{E_1}{E_2}$$

$$1.25 = \log \frac{E}{E_2} \quad \text{(dividing both sides by 20)}$$

$$\frac{E_1}{E_2} = \text{antilog } 1.25 = \text{antilog } 1.0 \text{ times antilog } 0.25$$

$$= 10 \times 1.78 = 17.8$$

The use of the decibel. The decibel may be used to express any ratio of power, voltage, current, mechanical loss or gain, etc. Thus we may say that a symphony orchestra has a range 60 db in power; that is, when it is playing very loudly, fortissimo, it is 60 db louder than when playing very softly, pianissimo. This corresponds to a power range of one million to one. In the wire circuits which carry the microphone currents from the symphony hall to the broadcast station, the weakest of the desired signals must be 40 db above the noise in the line. The very weak passages of the orchestra are built up by local amplifiers until the currents are greater than the noise currents. The limit to the louder passages is the overloading of the amplifiers either at the hall or in the broadcasting station or by causing "cross talk" from one wire circuit to another. And so the stronger passages are cut down.

Whenever a circuit suffers a loss in power or voltage or current, we may express that loss in decibels. The frequency characteristic of an amplifier, of a loud speaker, or of a telephone line may be expressed in decibels by plotting a curve in which zero level is the amplification or power output at some arbitrarily chosen frequency. Thus, if we choose 1000 cycles as a reference frequency, all other frequencies are either up, down, or flat with respect to the level at 1000 cycles.

Problem 14–14. What in decibels corresponds to a voltage ratio of 100? Power ratio of 100? What voltage ratio corresponds to 100 db? What power ratio?

Problem 15–14. A current of 0.006 amp flows through a resistance of 1000 ohms. A switch reduces this current to 1.0 ma. How much is the power reduced in decibels?

Problem 16–14. An amplifier has a normal output of 1 watt. A switch is provided that its output can be reduced in 5-db steps. What is the output in watts when it is reduced by 5, 10, 20, and 25 db?

Problem 17–14. An amplifier has its power output reduced by 25 per cent. Is such a reduction in power audible to the ear?

Problem 18–14. A radio receiver has a voltage gain in its radio-frequency amplifier of 50 db. Express this in voltage gain, and in power amplification, provided that the same impedance closes the input and output of the amplifier.

Problem 19–14. A radio receiver is so adjusted that a station 20 kc off the frequency at which the receiver is tuned is reduced by 40 db in voltage below the station that is being listened to. What is the ratio in voltage between the desired and undesired station?

Problem 20–14. A broadcasting station increases its power from 500 to 5000 watts. What is this in decibels? What is the increase if the power is increased to 50.000 watts?

Problem 21–14. If an audio amplifier has two stages and each stage has a gain of 25 db, what can the gain of the receiver be reduced to when listening to a 50,000-watt station compared to a 500-watt station provided that they are equidistant? In other words, a given station increases its power from 500 to 50,000 watts. How much audio gain in decibels is this worth to the listener?

Problem 22–14. The noise on a certain telephone line is 40 db down from the broadcasting signals. What is their power ratio? If the broadcasting currents are of the order of milliamperes, what are the noise currents?

Problem 23–14. The maximum power output from a 45 type of tube is 1600 mw. The 2A3 has an output of 3500 mw. How much greater in decibels is the 2A3 power output?

Problem 24–14. A loud speaker is 5 per cent efficient and requires 1.5 watts to give sufficient volume output. If it is made 50 per cent efficient how much can the power input be reduced to give the same output?

Problem 25–14. A tube has a plate resistance of 5000 ohms. Calculate the power into a load which varies from 1000 to 20,000 ohms and convert the ratio between the power at maximum to the power at other values of load resistance in decibels. How great can the difference between the load and the tube resistance be before the ear will note the difference?

Problem 26–14. The sensitivity of a condenser microphone is said to be −60 db where 0 db = 1 volt per dyne of force exerted by an air wave impinging on each square centimeter of the diaphragm. The carbon button microphone has a sensitivity of −40 db. What is the sensitivity, in decibels, of the carbon microphone compared to the condenser transmitter? How many stages of transformer-coupled audio amplication using 2 to 1 transformers and tubes with a μ of 8 will be required to bring the output of the condenser transmitter up to the level of the carbon button microphone? A commercial telephone transmitter, such as is used in ordinary telephones, is tuned to the average speech frequency and therefore amplifies what corresponds to its input about 1000 times. Express in decibels its sensitivity compared to the other two microphones.

Problem 27–14. A radio receiver is tuned to a certain distant station which gives at the receiver input a voltage of 500 μv. A near-by station on a different frequency produces a voltage of 50,000 μv at the same time. How much loss in decibels must be put into the receiver at the frequency of the undesired station to reduce the signals to the same level? How much additional loss must be put into the receiver to reduce the unwanted station to 60 db below the desired signal? The curves in Fig. 19–14 will be interesting in connection with this problem. They were published by Lloyd Espenschied in the *Bell System Technical Journal,* January, 1927. They show the relative selectivity of several types of receiver. The "double detection" receiver is a superheterodyne.

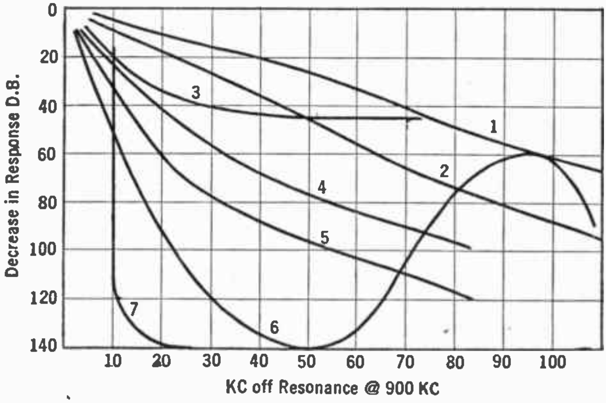


FIG. 19–14. Selectivity of several circuit combinations.

1. Single circuit non-regenerative.
2. Coupled circuit non-regenerative.
3. Single circuit regenerative.
4. Tuned radio frequency.
5. Tuned radio frequency.
6. Double detection (superheterodyne).
7. Ideal characteristic.

Compensating amplifiers at low frequencies. Very often the loss in low frequencies in parts of the circuit may have to be made up in the audio amplifier. For example, if the baffle in which the speaker is placed is too small to radiate the low tones properly, the amplifier may be "tilted" so that these frequencies are over-emphasized. Coupled with the loss in the baffle the over-all response may be fairly flat.

In Fig. 20–14 is a simple "bass-boosting" system which improves the low frequencies. Here a simple network is shunted across the volume-control system in such a manner that at ordinary volume levels the amplifier characteristic may be comparatively flat. At low levels, the bass-boosting system automatically pulls up the bass. This is shown in Fig. 20–14, where a tuned circuit is shunted across part of the volume-control potentiometer.

When the volume is turned down, both low and high frequencies are given a boost, but the low ones are increased relatively

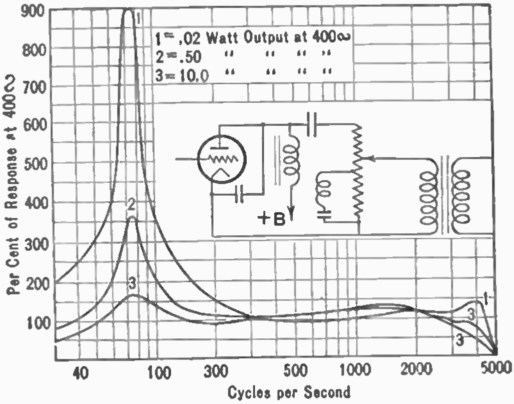


FIG. 20–14. Circuit and response of acoustically compensated volume control.

much more. This is known as the acoustically compensated volume control.

Compensation for high-frequency loss. Just as a tuned circuit may be used to pull up the response at a low frequency, so may high frequencies be increased if desired. For example, in television amplifiers, very wide bands of frequencies should be amplified. Owing to the various capacitances in tubes, sockets, wiring, etc., the high end of the frequency response curve of television amplifiers tends to droop. If a small inductance can be added to the circuit in such a way that it resonates with these small capacitances at a high frequency, the characteristic will tend to rise and the loss in high frequencies may be completely, or at least partially, overcome.

Power levels. Engineers frequently use the decibel system in expressing the output level of a signal, mentally comparing it with some fixed and understood level.

In Fig. 21–14 is indicated the amount the broadcast system compresses the power range encountered in music. A symphony orchestra covering a range in power from minus 40 to plus 20, or 60 db, must be compressed into a range of 27 db when it is broadcast. Here the zero level is the power output of the orchestra when playing at what a musician knows as forte; it is considered to be 1 mw into an impedance of 600 ohms.

Radio engineers have debated whether to rate the sensitivity of receivers in the microvolt or the decibel system. Thus to state that a receiver is a 10-μv set indicates that it requires this input

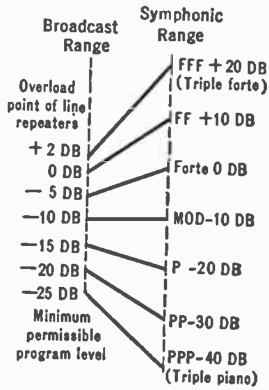


FIG. 21–14. Compression of dynamic range in broadcasting system.

to produce a given output, probably 50 mw. A receiver which requires 100 μv would be a 100-μv. set.

In the decibel system the sensitivity of the set would be rated in decibels below 1 volt. Thus a receiver which required 1 volt to produce a given output would have a rating of zero; one which required only 0.5 volt would have a rating of 6; one requiring only 1 μv would have a rating of 120. Zero level would then be 1 volt, or one million microvolts. Thus a high-sensitivity set would have a high rating number or value.

CHAPTER 15

DETECTION

Before the output of a microphone can be sent through the "air" from a transmitter to a distant receiver, the microphone output voltages must be combined with a radio-frequency (r-f) voltage in some manner. The reason for this necessity is the simple fact that alternating currents will not "radiate" unless the frequencies are much higher than those in the audible range. The lowest frequencies which will be radiated through space are of the order of 15,000 cycles, and transmission at this low frequency is very inefficient.

Modulation. Consider the circuit of Fig. 1–15, in which two sources of alternating voltage are in series. One is a very high-

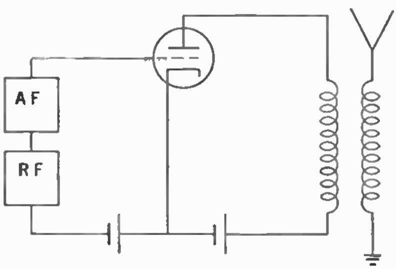


FIG. 1–15. A simple modulating system.

(radio-) frequency voltage, and the other an audible-frequency (a-f) voltage of variable amplitude but constant frequency, or tone. Let us short-circuit the low-frequency generator and turn on the r-f generator. Alternating currents will flow in the antenna, and the amplitude of these currents will depend, among other factors, upon the power output of the tube and the extent to which its grid is excited; in other words, the antenna current will bear some relation to the input voltage.

Now if the a-f generator is turned on, its output voltage is added to the r-f voltage at some instants (when two voltages are in phase) and subtracted from the r-f voltage at other instants (when they are out of phase). The actual grid voltage which excites the amplifier tube varies in amplitude, and of course the tube plate current will vary and so will the antenna current. The a-f generator acts like a constant-frequency, variable-amplitude grid bias. The greater the variations in the a-f voltage, the more will the r-f voltages be changed, and, if at some instants these two voltages are equal, the plate current (and antenna current) may be twice the value attained with the a-f generator turned off. At other instants, when these two voltages are out of phase, there will be no grid excitation and the plate current should be reduced to zero.

The a-f voltages, originally secured from a microphone, perhaps, cause the amplitude of the r-f plate current to rise and fall in unison with the amplitude of the a-f voltages. We have modulated the r-f voltages and of course have modulated the power radiated by the antenna. The high-frequency power, from now on, merely acts as a carrier for the information we wish to transmit, and, once the modulated r-f voltages are picked up at the receiver, we promptly lose interest in anything except the modulations.

A glance at Fig. 2–15 will show the variations in the r-f current produced by the low-frequency modulation. When the amplitudes of the two currents (or voltages) are equal, modulation is said to be complete, or 100 per cent. The high frequency is called the "carrier" frequency, and the low frequencies are called the modulating frequencies. The **percentage modulation** at any instant is 100 times the ratio between A, the peak current of the modulating frequency, and B, the peak current of the non-modulated high frequency.

Thus percentage modulation $M = A/B \times 100$ per cent.

Demodulation. At the receiving station, some process opposite to modulation must take place to secure, from the r-f energy, the modulating low frequencies which carry the desired intelligence. In this process, called demodulation or, more generally, detection, the modulations are separated from the carrier frequency.

There are three basic steps in detection. The first is rectification, by which one-half of the modulated r-f waves are cut off. The amplitude of the remaining r-f waves changes in accordance with the modulations at the transmitter. The second step is to

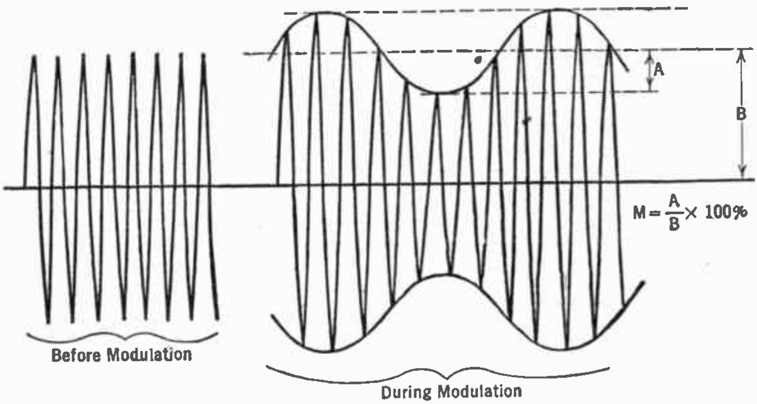


FIG. 2–15. Unmodulated and modulated wave.

filter out the remaining r-f voltages but at the same time to retain the amplitude variations in this voltage. The third step is to remove any direct current produced in the rectification process. Detection, therefore, consists essentially in passing modulated r-f voltages or currents through a rectifier. The output of the rectifier and its accessory apparatus consists in (1) a voltage (or current) whose amplitude varies in accordance with the modulations and (2) some direct current.

The detection process. A device to perform as a detector must have some portion of its input-output characteristic nonlinear. For example, let us look at the characteristic shown in Fig. 3–15, where the slanting line is the *E-I* characteristic of the device. Note that, if voltages more negative than -4 are placed upon the device, there is no output current, and that voltages less negative than −4 produce a current proportional to the voltage. If the voltage is raised and lowered about the −4-volt point, current will flow only when the voltage is raised above −4 volts. In other words, the device acts as a rectifier and the output current will look like one-half of the input.

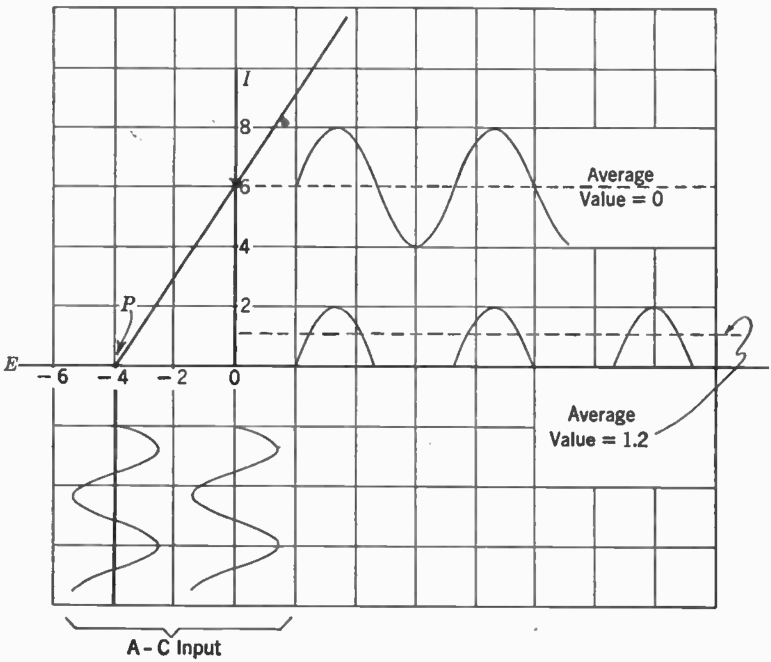


FIG. 3-15. Condition (bias = 0) producing no rectification or detection and condition (bias = −4) in which detection is accomplished.

On the other hand, if a fixed bias of zero volts is placed upon the device, and if this bias is raised and lowered, both halves of the input voltage produce output current. No rectification takes place.

In practice, the bias may be secured from a battery or from any source of constant voltage. The input (modulated r-f voltage) can be applied in series with this fixed bias so that the voltage actually applied to the detector varies up and down from its average value determined by the bias. If the operating point is P, then the output current will consist of waves having the same form (if there is no distortion) as half of the input voltage waves. The current measured by a d-c meter in the output will be some value between the peak of the individual waves and zero; it will reach some average value as indicated.

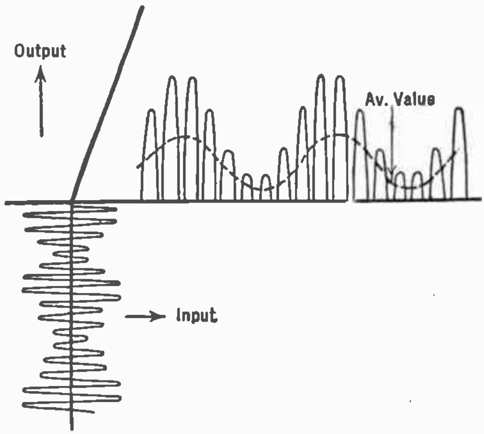


FIG. 4–15. Rectification taking place about a non-linear part of a characteristic.

Now, if, at the transmitter, modulation consists in merely turning on and off the transmitter, then a plate-current meter in the detector will show a flow of current when the transmitter is on but will indicate zero current when the transmitter is off. Instead of a meter, a telegraph sounder could be used to indicate the message which the transmitting operator wished to convey.

If, on the other hand, the operating point is at the zero-bias point, then there will be no rectified current and a d-c meter will not show whether the transmitter is being keyed or not. A detector must be operated about some non-linear point such as P.

In Fig. 4–15 note that the input voltages are modulated by some less abrupt system than simply turning on and off the transmitter. The peak value of the input voltage rises and falls in a rhythmic fashion. The peak values of the rectified current rise and fall in unison with these input variations. A line connecting these peak values is called the **envelope** of the wave, and, if there is no distortion in the detection process, the envelope of the output current will resemble, exactly, the envelope of the input voltage.

The average value of the current also follows the envelope; if the modulating voltage has a frequency low enough for a current meter to follow, then a meter in the output of the detector will indicate a rise and fall in average current in exact accordance with the modulation.

If the carrier frequency is 1000 kc, for example, one million of the individual cycles will be impressed upon the detector input each second; only a few are shown in the figure for simplicity. The peak value of these individual waves varies at the modulation rate. Thus it can be said that each audio or modulating cycle consists of thousands of r-f cycles.

Filtering the detector output. The first step in detection is to rectify the modulated wave. This leaves the positive half cycles of the r-f wave with a modulated envelope. Now that we are through with the carrier voltage, it can be eliminated simply by providing an easy path to ground for the r-f currents without, at the same time, short-circuiting the modulating voltages to ground. Figure 5–15 shows a simple detector made up of some sort of device that passes current in one direction but not in the other. This is the rectifier. The output is a resistance representing any sort of load, say a pair of telephones or the input of an amplifier tube. Across this load is placed a condenser.

Now when the first positive half cycle of current comes along from the rectifier it charges the condenser. When the charging voltage begins to fall the condenser discharges through the load resistance, holding up the voltage across the resistance in spite of the fact that the rectifier output is falling. If the condenser is large enough so that the time required to discharge it is greater than the time interval between positive half cycles, it will completely smooth out the r-f variations across the load; if the con-

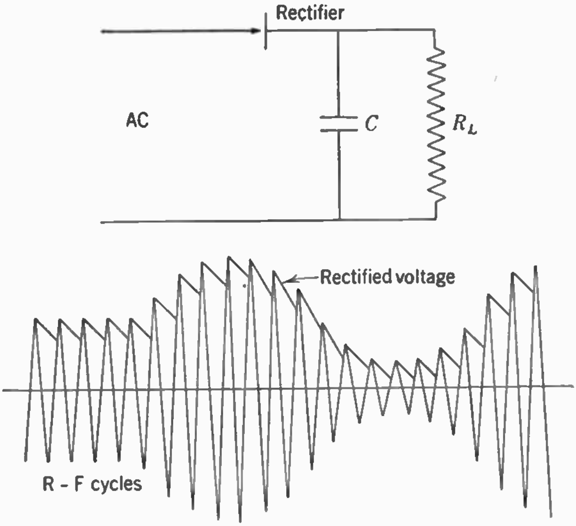


FIG. 5–15. Rectification process in which a condenser is used to maintain the envelope of the input voltage across the load.

denser is too large it will also smooth out the low-frequency modulation. The trick, then, is to choose a value of C that filters out the r-f voltages without at the same time doing away with the modulation.

Across the load resistor, therefore, appear modulating voltages, and through it is a direct current producing a d-c voltage drop. If the following circuit cannot utilize this direct voltage, all that is necessary is to place a blocking condenser between the rectifier load circuit and the following apparatus as shown in Fig. 6–15.

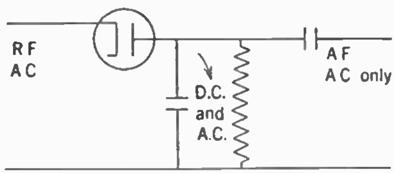


FIG. 6–15. Use of a blocking condenser as a filter to pass the envelope but not the direct current produced by the detector.

Crystal detectors. Very early in the radio art (in the "wireless" days) crystal detectors were utilized. This was before the invention of vacuum tubes. A crystal detector was made up of a piece of galena (lead sulphide) or silicon or molybdenum on which rested a small wire known as a "cat's whisker." Such crystals have unilateral conductivity, passing current better in one direction than in another. In other words, a crystal detector was a rectifier. The crystal was fairly sensitive, but it was very

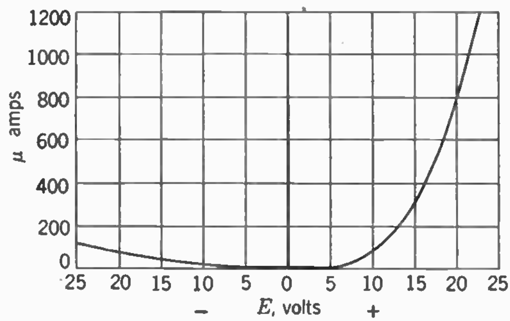


FIG. 7–15. Crystal (Carborundum) detector characteristic.

unstable because of the fine adjustment necessary between the thin wire and the crystal itself.

Now that ultra-high frequencies are coming into wide use, the crystal detector is coming back into vogue. It has one great advantage over a tube detector: being very small, it will not upset circuits in which ultra-high-frequency currents flow, for its capacitance is low compared to the capacitance of a tube, a socket, and attached wiring.

Diode detection. Nowadays the vast majority of detectors are simple two-element vacuum tubes. The diode is about as sensitive as a crystal but is stable. It delivers an output quite free from distortion if properly operated.

The diode may be made especially for the purpose of detection, like the 6H6. Any triode or more complicated tube, however, may be made into a diode by connecting two or more elements together. Thus a triode can operate as a diode if the grid and plate or plate and cathode are connected together. In all diode circuits there is a resistance load shunted by a condenser. The diode has low internal resistance during the time it conducts, and therefore a high load resistance will tend to make the *E-I* characteristic linear and so produce little distortion. The diode draws current from the input circuit and therefore puts a load upon it, drawing power from it. This decreases the selectivity of the input circuit.

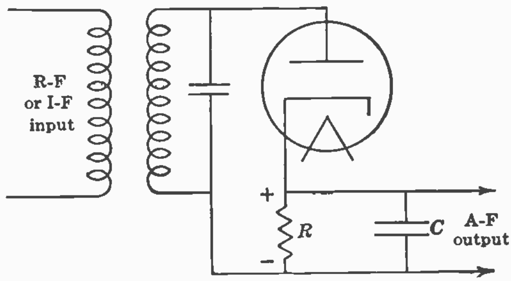


FIG. 8–15. Simple diode detector. Rectified voltages appear across *R*.

The simplest circuit is shown in Fig. 8–15. The load *R* is ordinarily about 0.5 to 1.0 megohm; *C* is about 150 $\mu\mu$f at broadcast frequencies and higher than this value at intermediate frequencies. Except at low input voltages, *R* is large compared with the internal resistance of the tube, and for this reason most of the voltage produced in the rectification process appears across *R*. The actual value of the rectified voltage is almost equal to the peak value of the a-c voltage across the tuned circuit. Values of *C* are chosen so that this condenser offers little reactance to the r-f voltage but has high reactance to the modulating voltages.

If one applies unmodulated alternating voltages across the tuned circuit in Fig. 8–15 and reads the d-c voltage across *R*, a curve like that in Fig. 9–15 will be obtained. This shows, for example, that 10 volts on the input produces a direct voltage across the load resistor, *R*, of approximately 13.5 volts. Now if this input voltage is completely modulated, at some instants it will be zero and therefore no (or very little) rectified voltage will be produced, and at some other instants the input will be 20 volts and approximately 27 rectified volts will appear. In other words, the voltage across R varies from zero to 27 volts and has an average value of approximately one-half of 27 or 13.5 volts.

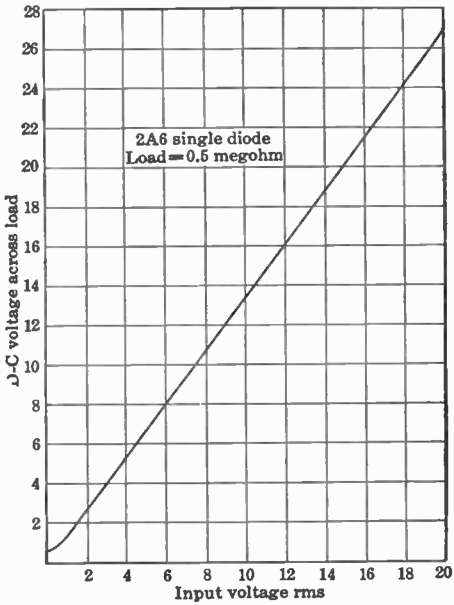


FIG. 9-15. Input-output characteristics of a diode detector.

An a-c voltmeter across R would measure a voltage having a peak value of 13.5.

Since the diode internal resistance is high, it would be very difficult to couple it to the following tube (an audio amplifier) by means of a transformer; therefore resistance-capacitance coupling is used. So long as a diode has a pure resistance load, it will detect without distortion. The envelope of the rectified voltage will be like the envelope of the modulated r-f voltage placed upon its input. But the RC network used to couple the tube to the following tube is not a pure resistance; and, in all receivers in which automatic volume control voltages are secured from the detector, an *RC* circuit is connected in the diode output.

At high values of modulation, the distortion of a diode detector may be appreciable because of these facts. If the diode load is the same for direct current and for the modulating frequencies, then no distortion results. One way to approach this happy state is to couple the following tube into the diode by connecting the tube to the middle of the resistance rather than at the tube end of the resistance. This lowers the voltage picked up and passed on to the following tube but increases the load the diode works into at all frequencies.

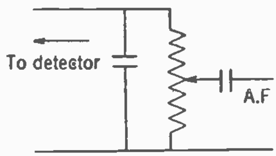


FIG. 10–15. Tapping the output of the detector puts stray capacitances across only part of the detector load, thus increasing the impedance of this load.

Plate circuit detector. In a plate circuit or C bias detector, an amplifier is deliberately overbiased so that distortion of the r-f wave results. In the process some direct current is produced, but in the plate circuit appears a varying voltage with an envelope that is a replica of the envelope of the modulated input.

Advantages of plate circuit detector. Compared to a diode detector, the plate circuit detector is much more sensitive (a given input will produce more output) because the tube actually amplifies the input r-f voltage before rectification takes place. Because the grid is so heavily biased, the impedance of the input is high and puts no load on the circuit out of which the tube works. The diode does not amplify, and, since it draws current from the input, it definitely loads the input down with resistance.

Plate circuit detector characteristics. The magnitude of the a-f voltage one gets from a plate circuit detector when it has certain r-f voltages placed upon the input and the effect of modulation on these input voltages may be determined experimentally as follows. We may fix upon some grid-bias voltage, e.g., for a 6C5 tube about 17 volts with a plate voltage of 250. Then we can put on the input to this tube, operating as an overbiased amplifier, or detector, various alternating voltages which need not be at radio frequencies. These input voltages cause some change in plate current which may be read with a fairly good milliammeter. All that is needed to determine experimentally the detection characteristic of a C bias detector is a source of known alternating voltages and a good milliammeter.

Such a series of curves is shown in Fig. 12–15. For example, with 100 volts on the plate, an input alternating voltage of 12 produces a plate current of 1.5 ma. Across such a family of

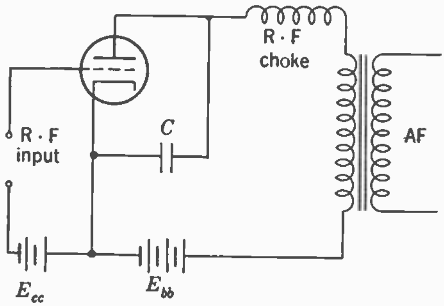


FIG. 11–15. Plate circuit detector. The combination of the choke and condenser C keeps r-f currents from the transformer.

curves a load line can be drawn for any load resistance, in this case 200,000 ohms. This is about the highest resistance load that can be used because of various capacitances which will shunt it and reduce its impedance at the higher audio frequencies.

The rectified output voltage of the detector, e.g., the a-f voltage applied across the 200,000-ohm resistor and hence to the input of the a-f amplifier, can be obtained from such a curve. For example with an input of 12 volts ($E = 12$) the rectified voltage may be found by noting the intersections of the load line with the $E = 0$ line and with the $E = 12$ line. Thus the rectified voltage is the difference between $E_p = 156$ (intersection with $E = 0$) and 66 (intersection with $E = 12$) or 90 volts. Taking several of such voltages a curve like that in Fig. 13–15 can be plotted.

Not only does this curve give the rectified voltage due to various values of *carrier* voltage, but also, by knowing how strongly this carrier is modulated, the actual a-f voltages applied across the load resistance may be ascertained. For example, if the

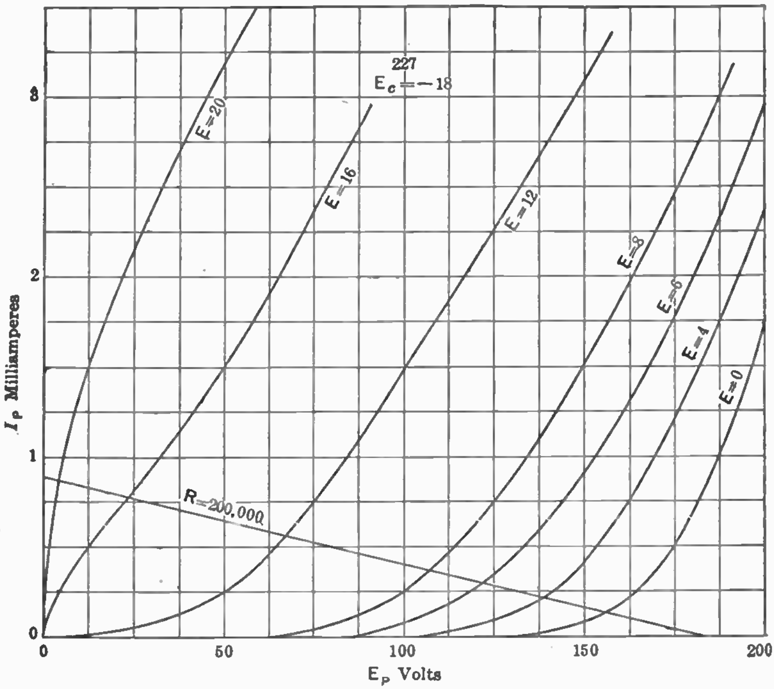


FIG. 12–15. Plate current curves as controlled by plate voltage and input carrier voltage E applied to grid.

plate voltage is 180 and $E_{cc} = -18$ volts, suppose that a carrier voltage of 10 is modulated 33 per cent. The carrier voltage will then vary between $10 - (10 \times 33$ per cent) and $10 + (10 \times 33$ per cent), or between 6.7 and 13.3 volts. These values of carrier voltage represent rectified voltages of 50 and 100, and, because the carrier swings as far up as it goes down from its unmodulated value of 12, the audio voltage produced by these variations is $(100 - 50) \div 2$ or 25 volts.

Detection in a r-f amplifier. Because the r-f stages of a receiver are biased, it often happens that some detection takes place in one or more of them, probably in the first. What happens is something like the following: the first tube is so biased, and may have such a low load impedance in its plate circuit at the frequency under consideration, that the operating point is

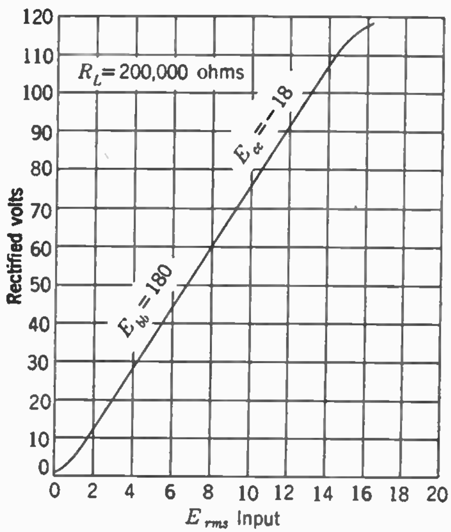


Fig. 13–15. Typical characteristic curve of plate circuit detector.

on a part of considerable curvature. Now a strong signal comes in, a large a-c voltage is impressed on the r-f grid, and detection is the inevitable result. A pair of headphones in the plate circuit of this tube would have an a-f voltage across them, due to the rectified voltage, and would give an audible response.

For example, suppose that the receiver is tuned to a frequency of 600 kc. This means that the load impedance in the plate circuit of the r-f amplifiers will be high at 600 kc but low to all other frequencies. The tube will have a curved characteristic to any signal of frequency other than this. A powerful local station on some other frequency puts a strong signal on the grid of the r-f tube, and the rectified voltages modulate the r-f voltage of the 600-kc wave so that what gets into the second r-f stage is a 600-kc signal modulated with what is going on at the studio of the other station. The modulation of the distant 600-kc sta-

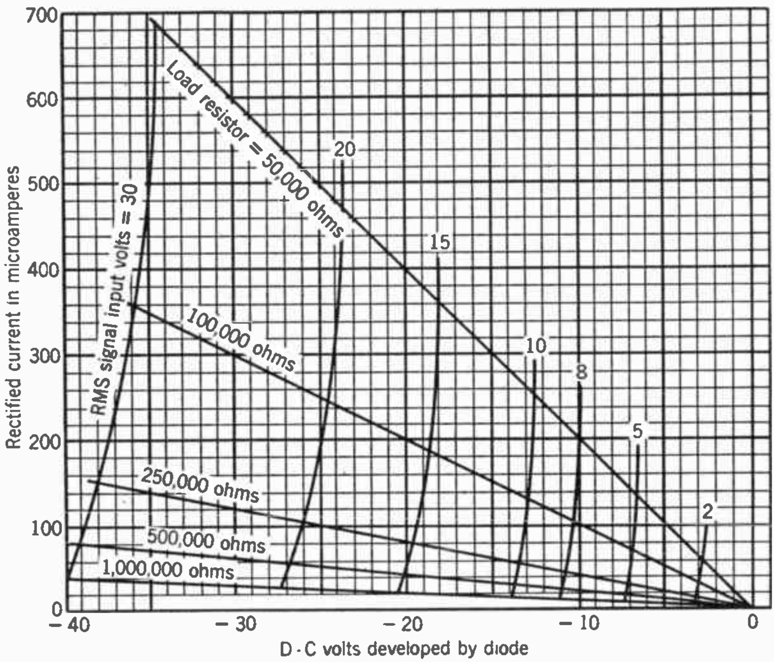


FIG. 14–15. Rectification diagram showing d-c output as a function of a-c input for various load resistances.

tion may be inaudible. A wave trap tuned to the offending local station is a good remedy for such trouble; supercontrol tubes are another remedy.

Grid-leak and condenser detector. In the plate circuit detector, we may look upon the signal as having first been amplified by the tube and then as going through the detection process when it reaches the plate circuit. There is little amplification in such a tube at radio frequencies because of the low plate load impedance at radio frequencies. This low r-f load impedance is deliberate, being part of the process of getting rid of the r-f variations in plate current. It is the result of *C* and the choke in Fig. 11–15.

The grid-leak and condenser detector shown in Fig. 15–15 is more sensitive and more complex in theory but has limited power-handling ability. In this case we think of the r-f signal as going through the demodulation or detection process in the grid circuit of the tube and then having the resulting audio tones amplified in the plate circuit just as in an ordinary amplifier.

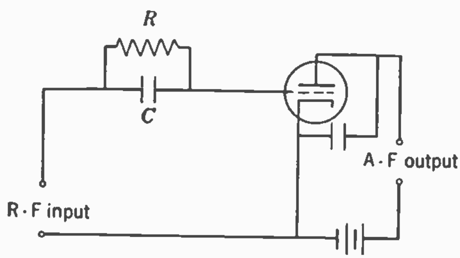


FIG. 15–15. Grid-leak and condenser detector.

Because of this amplification, this type of detector is more sensitive than the C bias (plate circuit) type.

Since such a tube detects on its grid-current curve, the grid voltage must be such that the operating point is on a curved part of the grid-voltage grid-current curve; and because it amplifies on its plate-current curve the plate voltage must be fixed so that the operating point for the modulating voltage is on a straight part of the plate-current curve.

The grid current plotted against grid voltage of a typical tube is shown in Fig. 16–15. It will be noted that, even though the grid is negative, a certain amount of grid current flows, owing to the few electrons which strike the grid on their way toward the plate. These few electrons bias the grid slightly negative, and the input a-c signals are applied about this point as the operating point. Because the characteristic is not linear but curved at this point, the positive half cycles of the input voltage are amplified more than the negative half cycles and rectification takes place in the grid circuit. The direct current which results flows through the grid-leak resistance back to the cathode. The d-c voltage drop along this resistance, due to the rectified current, changes the bias of the tube in accordance with the modulation

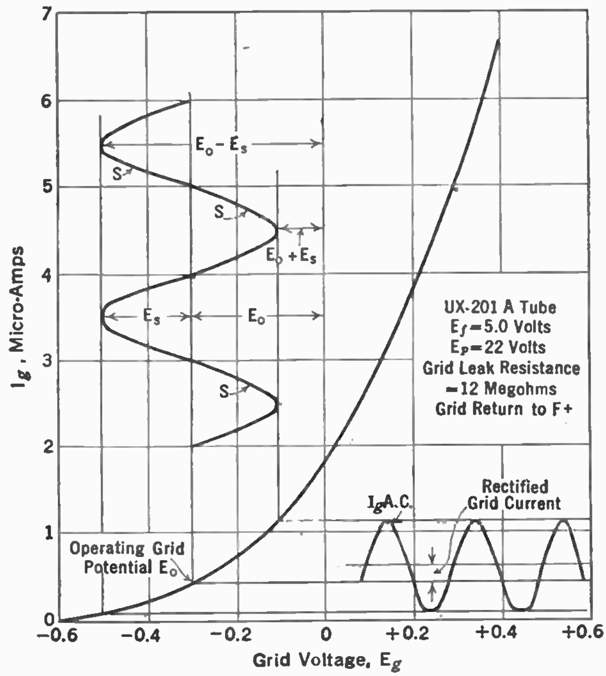


FIG. 16–15. Rectification on the grid voltage-grid current curve.

envelope. This change in bias changes the plate current accordingly, an *increase* in signal producing a *decrease* in plate current.

The plate voltage of the grid-leak detector is adjusted so that the maximum change in grid voltage takes place (maximum curvature of the grid characteristic curve), and this calls for a low plate voltage. Low plate voltages cause a curvature of the plate current curve so that distortion of the modulation envelope is produced, which increases as the percentage of modulation increases. The distortion is the result, not only of distortion products created in the plate circuit by non-linearity, but also of the fact that rectification in the plate circuit opposes rectification in the grid circuit, and so, at higher modulation, the envelope does not show the proper amplitude.

If the grid resistance is greatly decreased and the plate voltage increased, and if the input to the detector is greatly increased, much of the distortion of the grid-leak detector is eliminated.

Effect of grid-leak and condenser values. Changes in grid-leak value produce no other change in the detector action than is produced by changing the operating point. This is the entire purpose of the grid leak. The purpose of the grid condenser is to by-pass the high grid-leak resistance as far as r-f currents are concerned so that the greatest possible r-f voltage may be built up across the grid-filament input of the tube. If the condenser is too small, there will be an appreciable r-f voltage loss in it, and the tube will not get all possible of the input signal. If the grid condenser is too large, the a-f voltages built up across the grid leak will be by-passed. A good value to use is 0.00025 μf; a value of 0.0001 may be satisfactory. Smaller condensers than this produce some loss in r-f voltage and result in decreased sensitivity. Grid-leak detectors may be thought of as a diode detector plus a stage of audio amplification.

Automatic volume control. A notable advance in radio receiver design took place when automatic-volume-control (a-v-c) circuits were developed and placed in active use. By means of automatic volume control the effects of wide variations of voltage at the antenna terminals (fading) are overcome to a great extent. Furthermore, blasts of loud signals are prevented when tuning the receiver from a weak station for which the volume control might be turned up high to a powerful station for which the volume control must be turned down.

As the name implies, a-v-c circuits control the gain automatically. This is accomplished by varying the amplification of the r-f and i-f circuits inversely as the strength of the incoming signal. When this signal is high the gain is low, and when the incoming signal is weak the gain of the receiver is high. The circuits attempt to maintain a constant voltage at the detector terminals and are able to do this over a very wide range of incoming signals.

Although a separate tube may be employed to produce the voltages necessary for automatic volume control, it is common practice to utilize the d-c voltages developed in the diode detector. Thus in Fig. 8–15 the lower end of the diode load resistor has a negative potential with respect to the cathode of the diode tube. This negative potential may be utilized as automatic volume control by impressing it upon the grids of r-f and i-f tubes as bias. The stronger the signal impressed upon the input of the receiver, the stronger will be the signal impressed upon the detector input (in the absence of automatic volume control). The stronger the incoming signal, the more direct current is developed in the diode by the process of rectification, and this current, flowing through the diode load resistor, produces a voltage drop along it. A greater voltage is developed when a greater current flows through the resistor, and this greater voltage applied to the previous tubes as bias will reduce their gain.

What actually occurs is that the transconductance of the amplifier tubes is controlled by varying the bias on the grids of the tubes; the bias is caused to vary automatically with the strength of the signals applied to the receiver but in an inverse sense, that is, the stronger the signal the more negative bias that is applied.

In Fig. 17–15 are two a-v-c circuits. Note that a filter made up of series resistance (R_2) and shunt capacity (C_2) is placed in the a-v-c lead as it is taken from the diode. This procedure is necessary because the voltage across the diode load, R_1, varies with the modulation of the incoming signals. If the voltage were applied directly to the grids of the amplifier tubes, the bias of the tubes would go up and down with the modulation, and again in the inverse sense, so that all that came out of the detector would be a variation of tones all of the same strength. There could be no variations of audio volume.

In practice, then, the voltage from the diode charges the condenser C_2 through the resistance R_2. Condenser C_2 can discharge at only a slow rate, and therefore the voltage applied to the amplifier grids as bias cannot follow the modulations of the incoming signals although it may follow the variations in the signals due to fading or tuning from one station to another. The design of this filter is important. The values of R_2 and C_2 must be chosen so that the time taken by the condenser to discharge (its **time constant**) is such that the lowest modulation frequencies will not cause any variation in the grid bias of the amplifier tubes, but the discharge time must not be so large that a

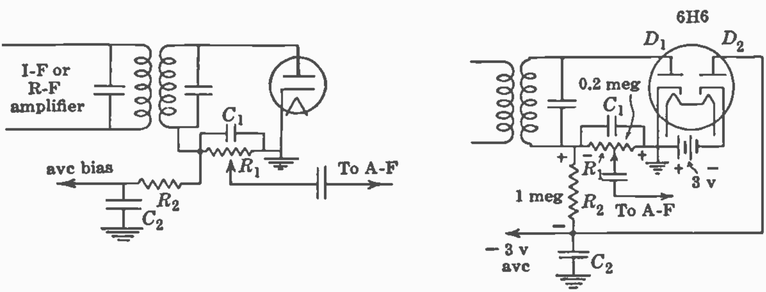


FIG. 17–15. Typical a-v-c circuits. In the second a permanent bias of −3 volts is applied to i-f tubes as "delayed" a-v-c voltage.

delay occurs when the system is recovering from a crash of static. A time constant of 1/10 to 1/5 sec is usually employed.

There are many variations of this simple a-v-c circuit. Sometimes the a-v-c voltages are amplified before they are applied; sometimes more automatic volume control is applied to one tube than to another; sometimes none is applied until the incoming signal reaches a certain value. This is known as "delayed" automatic volume control since the action is delayed until signals of a certain strength are attained. Sometimes the a-v-c tube is connected to a point in the circuit at which the selectivity is less than that at the detector terminals. In this system, noise encountered when the receiver is tuned off resonance is reduced. Sometimes a parallel chain of amplifier tubes is used to produce the a-v-c voltages.

Delayed automatic volume control. One of the circuits in Fig. 17–15 represents a system for delayed automatic volume control. The tube is a 6H6 which has two diodes in it. Diode D_1 acts as detector and a-v-c tube. R_1 and C_1 are the load of this diode. R_2 and C_2 are the filter for the automatic volume control. A 3-volt potential exists between the two diodes, and a direct current flows through R_1, R_2, and D_2. The a-v-c tap, therefore, is approximately -3 volts potential with respect to ground. This is a permanent negative bias of -3 volts which is applied to the grids of the previous amplifier tubes. So long as not more than 3 volts is developed by the diode across R_1, this bias remains at -3 volts. When, however, stronger signals are placed upon the input to the diode, the plate of diode D_2 becomes more negative than the cathode of D_2 and so current ceases to flow in the circuit. The a-v-c lead, therefore, takes the voltage that is produced by the drop across R_1. This negative voltage is applied as bias to the amplifier tubes, in accordance with the signal strengths applied to them.

Thus the amplification of the receiver is at a maximum for weak signals but varies inversely as the strength of strong signals.

Manual volume control. The a-v-c system supplies constant voltage to the detector over a very wide range of incoming signals. Variations in loud-speaker volume, therefore, must be controlled by a manual adjustment, by taking more or less of the audio output of the detector and applying it to the a-f amplifier.

Where automatic control is not available, volume may be controlled in several ways. One is to decrease the gain of the r-f or i-f amplifiers manually by adjusting the bias. Since more bias adjustment is required for the amplifiers near the detector than for those near the antenna, some means is usually provided for proportioning the volume-control potential inversely with the signal applied to each tube.

Sometimes the bias on the oscillator tube is adjusted at the same time the amplifiers are controlled, either manually or by a-v-c circuits.

Direct plate current as a function of a-c grid voltage. The vacuum-tube voltmeter is really a C bias or plate circuit detector, and the change in plate current is a function of the a-c voltage on the grid. It is only necessary to calibrate the detector, using any source of alternating current and any standard a-c voltmeter as a standardizing voltage, or a known current can be passed through a known resistance and the voltage drop used for calibration.

The detector in a radio receiver is also a vacuum-tube voltmeter although it is not so calibrated. If a sensitive meter, reading, say, up to 0.5 ma, is placed in the detector plate circuit of a radio receiver, changes in the reading will be noted when a strong signal is tuned in. If greatest sensitivity is desired, the steady no-signal current may be balanced out and then a sensitive microammeter may be used. The changes in this meter reading may serve as a measurement of fading, signal strength, etc. No change will occur unless the r-f voltage on the input to the tube changes. Modulation will not cause any change unless the transmitter is over-modulated.

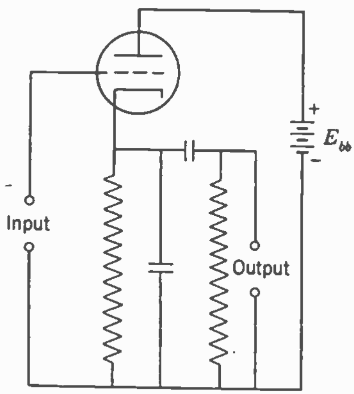


FIG. 18–15. Infinite-impedance detector, a form of plate circuit detector in which the load and the plate battery are interchanged.

Infinite-impedance detector. A form of plate-circuit detection is shown in Fig. 18–15, where the plate battery and the plate load resistor are interchanged. The load resistance is quite high, sufficient to cut off the plate current, and for this reason the input impedance of the tube is infinite. Thus it puts no load on the circuit from which it derives its voltage. Rectified plate current flowing through the cathode resistor builds up a voltage across it which is only slightly less than the envelope amplitude. The peak voltage at modulation frequency is large, approaching half the plate supply voltage. The sensitivity is about the same as that of a diode.

CHAPTER 16

RECEIVER SYSTEMS

Basic qualities of a radio receiver. Three basic qualities of a radio receiver must be weighed not only by the designer but also by the ultimate user: **sensitivity, selectivity,** and tone **fidelity.** The sensitivity of a receiver is an indication of the over-all amplification from antenna-ground binding posts to loud speaker. A receiver that is very sensitive requires but a small input voltage to deliver considerable output power. The selectivity of a receiver is an indication of its ability to discriminate between wanted and unwanted signals. An infinitely selective receiver would be one that would respond only to a given station and not at all to another no matter how powerful this undesired signal was, nor how close in frequency it was to the desired signal. That there is no such receiver goes without saying. The fidelity of a receiver tells how well it reproduces what comes from the microphone at the transmitter. A receiver that delivers a high-fidelity signal is one which has a flat a-f response curve over a wide frequency band and is free from noise and distortion.

A receiver that is perfectly selective and infinitely sensitive and that delivers a perfectly faithful signal over a wide audio band is impossible to obtain. A receiver which is selective and sensitive enough for practical purposes, and which has a frequency curve so good that the ear would not detect any distortion, is not difficult to design, to build, or to operate.

Field strength. The voltage that is set up across a receiving antenna is called the **field strength** of the transmitter at that particular point on the earth's surface. It is of the order of microvolts or millivolts, and, because a higher antenna will pick up a greater signal, that is, the voltage across it will be greater, it is the practice to rate field strength as so many microvolts or millivolts per meter. Thus an antenna that has an effective height of 1 meter and has 4 μv across it is situated in a field strength of 4 μv per meter. The effective height of the antenna is somewhat less than its actual physical height above ground; in most receiver measurements it is assumed as 4 meters (13 ft). An antenna that has an effective height of 4 meters and a voltage of 10 μv across it is immersed in an electric field due to some transmitting station whose strength is 2.5 μv per meter.

The greater the field strength at a given point the more volume one can get out of a receiver with a fixed amount of amplification. Similarly, the greater the field strength, the less receiver amplification is necessary to give out a certain amount of power.

Desirable signal strengths. Experience has indicated that signal strengths of three general classifications are necessary to provide good service to listeners. In city business areas 10 to 25 mv per meter is required to override high interfering electrical noise and the shadows cast by large buildings; in the residential areas of large cities, field strengths of 2 to 5 mv per meter are required; and in rural areas where man-made noise is low, fields as low as 0.1 to 0.5 mv per meter will provide satisfactory service. It must be remembered that the absolute value of the signal is seldom the important quantity; it is how much louder the signal is compared to the noise; and that in a locality where noise is low (as in the country) the signals need not be so strong.

The signal strength at a remote receiving point is proportional to the square root of the power at the transmitter; it depends also upon the frequency of transmission, the heights of the receiving and transmitting antennas, and the kind of soil or terrain between transmitter and receiver. Also, of course, the signal strength decreases as the distance increases. At a point far from the transmitter a much greater area is permeated by the same energy as at a point near the transmitter, and therefore a given area, such as that covered by a receiving antenna, picks up less total energy at the distant point. Some further loss is due to absorption of energy in foliage and other material of high resistance.

The purpose of the transmitting station is to provide a good lusty signal that will override static and other disturbances; the purpose of the r-f amplifier is to provide the listener with good loud signals from the field strengths which the stations produce.

Advantage of high power at transmitting station. Whatever voltage exists across the antenna, whether noise or desired signal, is amplified by the r-f amplifier; there is therefore a distinct advantage, as far as the receiver is concerned, in using large amounts of power at the transmitting station. The greater the ratio of signal to noise the better will reception be. No matter how great the voltage gain of a r-f amplifier, it cannot bring a weak signal out of the noise and give satisfactory reception. The signal must always be about 40 db above the noise level (in an amplitude-modulated system) to provide entertainment free from a noisy background. Whenever the noise increases, as on a warm summer day, and the transmitter station power remains constant, reception suffers, and it suffers more the farther the receiver is removed from the station. The noise about a given receiver is more or less constant under a given set of conditions, whereas the field strength due to a transmitter decreases with distance from the transmitter.

If a receiver is situated in a quiet locality, where the noise level made up of stray voltages from street cars, elevators, arc lamps, power leakages from high-tension wires to trees, sputtering flatirons, x-ray machines, electric razors, etc., is weak, the greater the amount of over-all amplification, the greater the distance away a transmitter of a given power can be and still provide an adequate signal; and, of course, with the same amount of amplification, the weaker a station can be at a given distance to provide the desired receiver signal. The purpose of the r-f amplifier is the same as that of a telescope—to decrease the effective distance between transmitter and receiver. Like a telescope, a receiver can be aimed at a given transmitter by means of directive antennas with the result that the signal-to-noise ratio is increased.

One of the virtues of the high frequencies is the fact that there is less natural static on them; but man-made static and electric noise increase with frequency until very short wavelengths are reached.

If the transmitter power is increased by 10 db (10 times the power), unwanted signals are automatically decreased 10 db compared to the desired signals. This desirable effect may be achieved by means of a directive antenna at transmitter or receiver or both.

Amount of amplification necessary. Theoretically it would be possible to add as many stages of amplification as desired. They could operate at the frequency of the incoming signals, or at the modulation (audio) frequencies secured by the detection process. By means of the superheterodyne principle, these stages could operate at some frequency intermediate between the incoming frequencies and those which are the audible final result. But definite difficulties stand in the way. In the first place, amplifiers tend to become unstable if the gain is too high, owing to the difficulty, especially at high frequencies, of preventing some of the output from getting back into the input. In the second place, some noise is always created in the amplifying and detecting process. This noise will be amplified with the signal and may seriously interfere with or completely mask the desired signals.

In practical receivers, some of the amplification takes place at the frequency of the incoming signals and some at audio frequencies. In the superheterodyne receiver the incoming signals are changed to a lower frequency (still above audibility) and are amplified at this frequency. In such a receiver, there may be amplifiers operating at radio frequency, at intermediate frequency, and at audio frequency.

If a diode detector is employed, not less than 10 volts (rms) will be necessary on the input if distortion is to be kept low. Between the detector input and the antenna-ground binding posts must be sufficient amplification so that 10 volts is produced from the incoming signal, whatever this voltage may be. Under ideal conditions, inherent noise at the antenna-ground binding posts is no lower than 1 μv and is nearly always higher. To enable the receiver owner to "reach the noise level" will require that a total amplification of $10 \div (1 \times 10^{-6})$ or 10 million times up to the detector be available. A receiver is usually designed to have less amplification than this, for the simple reason that it is seldom that a gain of 10 million times can be used.

There is usually a gain of 2 to 5 times in the input system to the first tube. If a stage of r-f amplification precedes the frequency changer, it may have a gain of 30 times. Another gain of 20,000 will occur from the grid of the frequency changer to the grid of the detector, if a two-stage amplifier is employed. Some amplification occurs between the detector and the loud speaker. The total amplification amounts to 3 million times, which is about the limit to be found in a good receiver.

Amount of selectivity necessary. The selectivity depends entirely upon the use to which the receiver is to be put. If it is to be a high-class receiver for broadcast reception the selectivity problem is much greater than if the receiver is to be used on general communication circuits. In such service high selectivity is possible because the band width to be passed is limited, a frequency range from 250 to 2500 cycles being all that is needed for intelligibility. A high-class broadcast receiver, on the other hand, should pass all frequencies from approximately 60 cycles to 10,000.

Up to the time of the second World War, manufacturers of broadcast radio receivers had made only feeble attempts to provide equipment with sufficient selectivity without impairing the tone fidelity at the same time. In endeavoring to compete with each other on a price basis, many short cuts were made in engineering, with the result that the average receiver reproduced very little beyond 4000 or 5000 cycles and still did not have proper selectivity.

Code receivers employing a quartz crystal filter are able to separate two signals only a few hundred cycles apart. Of course such a receiver would be unsatisfactory for voice or music reception.

Tone fidelity required. If the receiver is to be used on a code circuit, only a very narrow band is to be received. If a general communications receiver is used for both code and voice, better tone fidelity is required. If the receiver is to be a high-fidelity broadcast set, then very good tone fidelity is absolutely necessary.

Tone fidelity means not only wide-band response but also freedom from noise and distortion in that band. The expense of a receiver mounts rapidly when the band width that is to be reproduced accurately is increased. A loud speaker which is relatively free from distortion and which transmits frequencies from 60 to 10,000 cycles may cost as much as a small radio receiver of the pre-war variety. Thus the cost of a high-fidelity receiver may be several times that of the average radio.

Types of receiver circuits. Several types of circuits are used for radio receivers:

1. Simple crystal detector or a single tube receiver.
2. A regenerative detector plus audio amplifier.
3. Tuned-radio-frequency amplifier (t-r-f receiver).
4. Superheterodyne.
5. Superregenerator.

Crystal detector receiver. In Fig. 1–16 is a simple receiver circuit which satisfied wireless operators for many years before vacuum-tube equipment was available. In those days, problems due to the multiplicity of radio stations were not so great; signals were desired only over fairly short distances (from ship to shore); and headphone reception was all that was required. Nowadays, a crystal detector would be used only where sufficiently strong signals were available, far enough removed from interfering signals in distance or in frequency to cause no trouble. In place of the crystal, a grid-leak detector tube could be utilized. This would have about 10 times the sensitivity of the crystal, but the problem of selectivity would be unsolved.

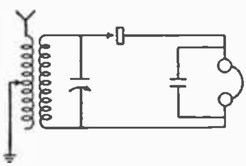


Fig. 1–16. Early wireless detector using a crystal as rectifier.

Regenerative detector. If, as in the circuit of Fig. 2–16, part of the output of the tube is fed back into the input in a regenerative phase, output signals will be increased and the sensitivity and selectivity of the circuit will be improved. If the regeneration is increased sufficiently, the tube will "oscillate," that is, it will generate alternating currents of its own. If the receiver is made to oscillate at a frequency slightly different from that of the incoming carrier, say 1000 cycles off, the combination of the incoming signal and the locally generated signal creates a third signal whose frequency is the difference between

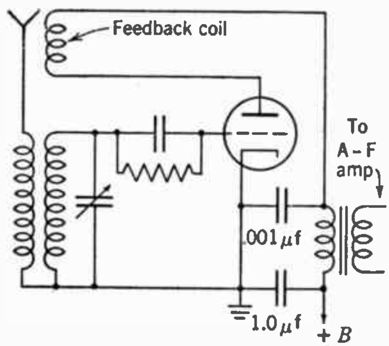


Fig. 2–16. Voltage fed back from plate to grid circuit increases sensitivity and selectivity of simple receiver.

the two combining frequencies. In this example it is 1000 cycles, and this 1000-cycle note will go on or off if the distant transmitter is turned on or off (keyed), and thus the circuit may be employed for code reception.

Up to the point where the receiver actually breaks into oscillation, voice or music can be heard, but after oscillations begin, only unintelligible squeals are heard unless both the frequency of the incoming signal and that generated by the tube itself are very stable. Then the operator may tune the receiver to exact resonance with the incoming signals and can hear music or voice.

As it is possible to generate the difference frequency by means of a separate oscillator, the detector may be tuned exactly to the incoming signal. The difference or "beat" frequency may be made lower than that of the incoming signal, but still inaudible.

This "intermediate" frequency carries the modulation of the incoming signal. It can be amplified and then applied to a detector from which the audio tones are secured. The superheterodyne is based on this principle.

A "communications" receiver is arranged so that signals can be heard from voice-modulated stations or from code stations. When voice or music is to be received, the receiver acts as a simple superheterodyne or a simple t-r-f receiver. For the reception of code, a beat oscillator is turned on, producing an audible tone which is modulated by the code of the distant transmitter. The tone of the beat note can be varied by variations in the tuning of the receiver proper, the beat oscillator producing a fixed frequency. A beat oscillator is provided in communications receivers; it is not the same thing as the "converter" oscillator, which is part of the conventional superheterodyne.

A regenerative receiver made up of a single regenerative detector and one stage of audio amplification is remarkably sensitive, considering the simplicity of the circuit and paucity of apparatus required. It is most sensitive at the point where the detector is on the verge of breaking into oscillations. Here the pass band accepted by the tuned circuit of the receiver is quite narrow, since the effect of regeneration is to reduce the resistance of the tuned circuit so that its response curve becomes very sharp. At this point, the tone fidelity is not at all good, but with a pair of headphones the signals are remarkably acceptable from the tonal standpoint.

At the point of greatest sensitivity, the detector is unstable and a change in d-c tube voltages or a mechanical jar may set the circuit into oscillation. Now the tuned circuit and its accompanying tube become a generator of signals, and, if the coupling between the tuned circuit and the antenna is close, the signals produced by the receiver may radiate from the antenna, causing a disturbance which may be heard in other people's receivers over a wide area. In wartime, such radiating receivers are not tolerated, and great care is taken to prevent any signals generated in a receiver from radiating.

The oscillator in a superheterodyne is also a generator, and in cheap receivers using this circuit the oscillator r-f voltages may produce disturbances in near-by receivers. To prevent such unwanted radiation, it is common practice to interpose a stage of r-f amplification between the antenna and the regenerative detector or the converter of a superheterodyne, and to shield all oscillator circuits well. The r-f stage not only acts as a buffer

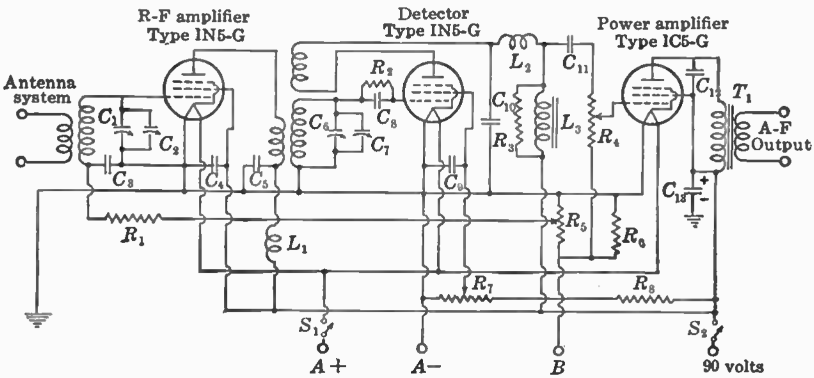


FIG. 3–16. Typical short-wave receiver using one stage of r-f amplification and a regenerative detector.

between the source of oscillations and the antenna but also produces a certain amount of amplification of the incoming signals and a certain amount of selectivity. The straight regenerative receiver which utilizes no preliminary amplification is very broad in its tuning unless the regeneration applied is nearly at the point of instability. The addition of the r-f stage helps the selectivity problem materially.

Feedback to produce regeneration may be secured in several ways, as is indicated in Fig. 4–16. Such receivers have two controls—one for tuning the receiver to respond to the incoming signals, and one to adjust the amount of regeneration. A receiver of this type can be extremely sensitive since the application of regeneration may increase the response of the tuned circuit several hundred times compared to the same circuit without regeneration.

Superregenerative receiver. It will be found that a simple regenerative circuit is most sensitive at the instant just before it breaks into open oscillation. In the superregenerator, invented by Major E. H. Armstrong, means are provided for allowing the circuit to oscillate for a small fraction of a second, and then to break off the oscillations. In other words, the tube and circuit are allowed periodically to approach the oscillation point and then are forced to recede from it. This function is per-

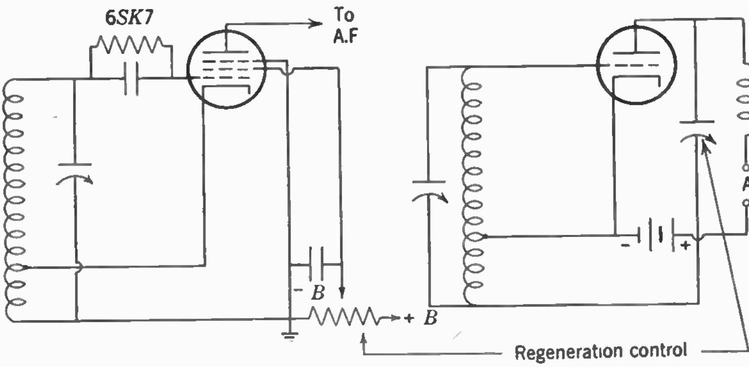


FIG. 4–16. Two feed-back circuits for improving sensitivity of a detector.

formed by an additional "quench" frequency (50,000 cycles, approximately) impressed on the tube so that some characteristic, which tends to make the tube oscillate, is periodically changed into and out of the oscillating region.

A one- or two-tube receiver of this type is extremely sensitive. Molecular motion in the input circuits produces a characteristic "rush" noise until a carrier signal stronger than the thermal voltages appears. Then the noise disappears and the desired signal comes in. The "quench" frequency is filtered out of the circuit so that the listener is not bothered by it.

This circuit is most useful in the very high-frequency bands (5 to 10 meters) where it is difficult to get high amplification in any other way. The circuit is not selective, but the lack can be supplied by the use of one or more stages of r-f amplification ahead of the superregenerative detector.

Tuned-radio-frequency receivers. If the detector is preceded by several stages of r-f amplification, the receiver is called a tuned-radio-frequency or simply a t-r-f set. The tuned circuits, of which there may be as many as six, are all tuned simultaneously by means of a "gang" condenser. This is a variable capacitor with as many sets of plates as there are circuits to tune. Pentode tubes are generally used, coupled together by means of transformers. The primary of these transformers is usually of high inductance, 4 mh (for the broadcast band), and the secondary size depends upon the tuning condenser size. For the band 550–1600 kc a secondary with an inductance of 270 μh is satisfactory when employed with a condenser having a capacitance range of 15 to 300 $\mu\mu$f. If the coil has a Q of 100 the voltage gain will be approximately 30 per stage. Much higher amplification is possible, but not with stability.

The selectivity of a t-r-f amplifier depends entirely upon the characteristics of the transformers employed. Since the resistance of the coil and the other apparatus to which it is attached governs the selectivity and since the resistance is different at different frequencies, the selectivity of a t-r-f receiver is not constant, but is high at the lower carrier frequencies (600 kc) and low at higher frequencies (1600 kc).

Furthermore, the t-r-f receiver is difficult to align and to keep lined up so that each of the individual stages tunes to resonance at the same setting of the variable condenser. One more difficulty is undesired feedback from output to input which leads to instability and often to outright oscillation.

Before the advent of screen-grid and pentode tubes, much trouble was experienced with triodes in t-r-f receivers because of the grid-plate capacitance of the tubes. By the addition of a small condenser to the circuit, approximately equal to the grid-plate capacitance of the tube, feedback to the input could be added to the circuit in such a manner that the inherent feedback through the grid-plate capacitance could be balanced out. A voltage opposite in polarity to that fed back through the tube was added to the input by this means, and thus the circuit and tube were stabilized. Several "neutralizing" circuits of this sort

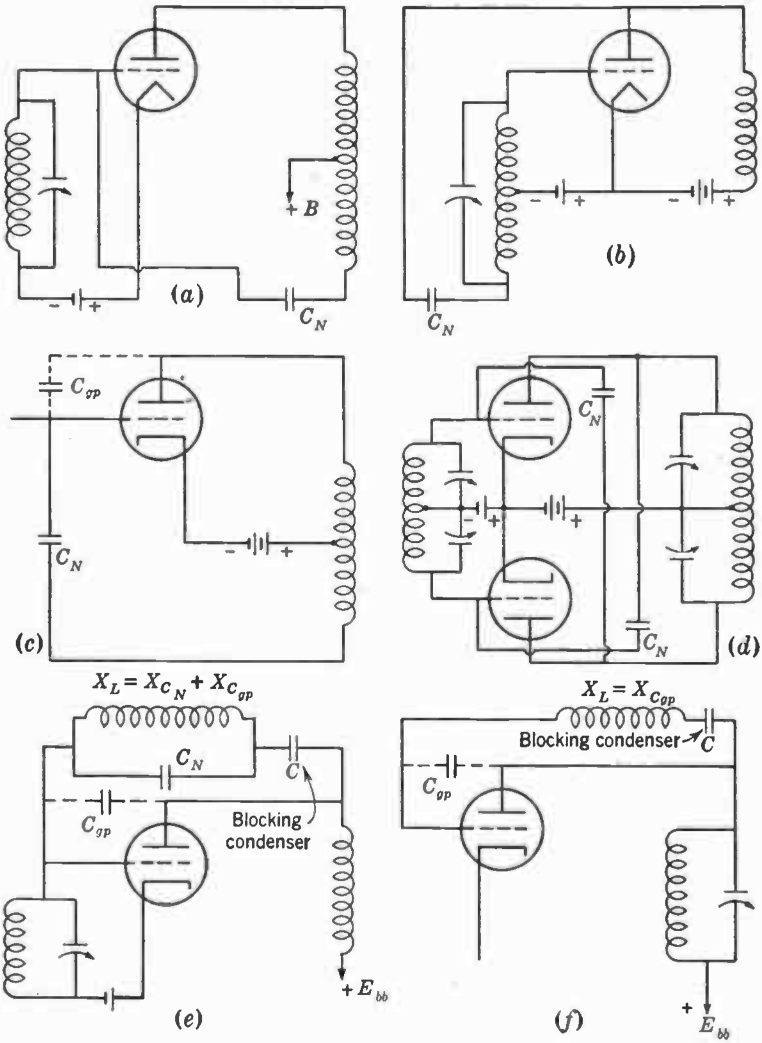


FIG. 5–16. Neutralizing circuits.

(a) Use of split plate transformer.
(b) Split transformer in grid circuit.
(c) Equivalent bridge circuit of (a).
(d) Neutralized push-pull amplifier.
(e) Network connected from plate to grid.
(f) Resonant circuit for neutralization.

were developed and are still used in r-f amplifiers for transmitters. The use of triodes as r-f amplifiers in receiving circuits, however, has been discontinued.

Superheterodyne. By far the vast majority of present-day receivers are built on the superheterodyne principle. In such a receiver, the incoming signals are changed in frequency immediately, or after some preliminary selection and amplification. The modulation of the incoming signals is transferred to the new, and usually lower, frequency. The new frequency, called the

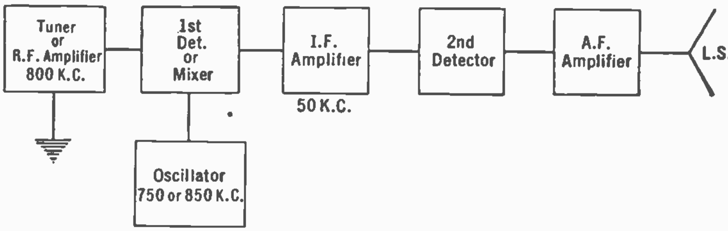


FIG. 6-16. Symbolic diagram of superheterodyne.

intermediate frequency, is amplified as much as is desired in a r-f amplifier whose response characteristics can be carefully adjusted at the factory and thereafter maintained constant. The i-f amplifier can be as selective or as broad as is desired for the service to be performed. After this amplification, demodulation takes place in a conventional detector; the audio signal is amplified further in an a-f amplifier and is thereafter employed for whatever use the receiver is designed.

The advantages of the superheterodyne compared to a t-r-f receiver are as follows: (1) All incoming signals are changed to a single fixed frequency and are amplified at this new frequency. In the t-r-f receiver, all amplification takes place at the frequency of the transmitting station. This brings up the difficulty of securing uniform amplification and uniform selectivity over a wide frequency band. (2) There is less danger of feedback troubles at the lower frequency for which the i-f amplifier is designed. Therefore, somewhat greater amplification can be secured with the same degree of stability, compared to a t-r-f receiver. (3) The selectivity of a stage of i-f amplification will be about twice as great as a stage of r-f amplification if the intermediate frequency is half that of the radio frequency.

This comes about in the following way. Suppose that the desired signal has a frequency of 1000 kc and that an interfering signal is on 1010 kc. The unwanted signal is 10 kc away from the 1000-kc signal, or off resonance by 1 per cent. If both signals are converted to a lower frequency, say 456 for the desired 1000-kc signal, the undesired signal will become 466 kc or 10 kc in 456 or about 2.2 per cent off resonance. If both the 1000-kc and the 456-kc resonance curves are equally sharp the 466-kc signal will be farther down on the curve and will, therefore, produce less interference.

Heterodyne action. The frequency is changed by mixing the incoming signal with a locally generated signal that may be higher or lower than the incoming signal by the amount of the intermediate frequency. Consider an incoming signal of 1000 kc. If a locally generated signal of 1456 kc is mixed with the 1000-kc signal, new frequencies representing the sum and difference of the two originating frequencies will be produced. In this example the difference frequency, 456 kc, will be utilized. This is the intermediate frequency. The local oscillator, which produces the locally generated signal, is tuned to a higher frequency than the incoming signal, and of course it must be tuned to a new frequency each time an incoming signal of different frequency is to be received. The problem, then, becomes one of maintaining a constant frequency difference of 456 kc between the incoming signal and the signal generated in the local oscillator. This is done, usually, by a system of series and shunt condensers placed across the main tuning condenser. Actually, the exact difference frequency of 456 kc is produced only at three points in the broadcast band, but the variation over the entire band is not great enough to destroy the virtue of the system. The sum frequency is filtered out by the tuned interstage transformers.

Preselection. A stage or two of r-f amplification ahead of the frequency conversion has several advantages, but the disadvantage of any t-r-f system is the difficulty in maintaining constant amplification and selectivity over the frequency band to be received. The advantages are: (1) Coupling between the oscillator and the antenna is minimized, with a reduction in the danger of interference in near-by receivers due to the radiation of the signal generated by the receiver. (2) Interference created by signals coming into the input of the receiver from a transmitter operating on a frequency equal to the intermediate frequency is reduced. (3) "Image" frequency troubles which come about for the following reason are reduced. If the intermediate frequency is IF, and the oscillator frequency is f, then incoming signals having frequencies of either $f - IF$ or $f + IF$ will be heterodyned and will pass through the i-f system. If the system is designed to operate on the $f - IF$ frequency, some means must be provided for keeping $f + IF$ signals out. For example, if the local oscillator is tuned to 1455 kc and if the IF is 455, then incoming signals of 1000 kc (1455 – 455 kc) or 1910 kc (1455 – 455) will produce the desired IF and will be sent through the system. Preselection reduces the voltage of the 1910-kc signal because the tuned circuits of the preselector, being tuned to 1000 kc, will have little response to voltages of a frequency of 1910 kc. (4) A higher signal voltage will be introduced into the frequency converter system and will increase the signal-to-noise ratio of the converter output. The receiver, therefore will be quieter in operation.

A tube used in the i-f system, however, will provide greater amplification and greater selectivity to signals on channels adjacent to the desired signals than if used in a r-f amplifier, and it will cost less. The temptation, therefore, is to use the tube in an i-f amplifier and to neglect the preliminary amplification wherever possible.

Frequency conversion. Before special tubes were available for frequency conversion, a triode or pentode tube was used for a **mixer** tube. In this tube two frequencies were impressed on the grid, one corresponding to that of the desired incoming signal, and the other that generated by the local oscillator. These two signals were introduced into the mixer tube in various ways, as by inductances or capacitances, and, as a matter of fact, could be introduced in any of the methods by which one frequency is modulated by another.

In the mixer tube, the two frequencies beat together, producing the sum and the difference frequencies whose values are determined by adding or subtracting the individual component frequencies. The output of the mixer was tuned, usually, to the

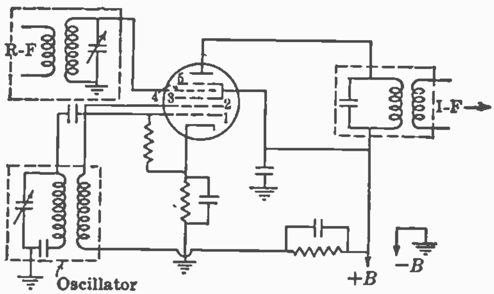


FIG. 7–16. Pentagrid converter circuit for frequency changing.

difference frequency, and this was the intermediate frequency. It bore the audio modulations of the incoming signal.

In a more elegant method, the two frequencies were coupled by means of an electron stream. In the pentagrid-converter tubes (like 6A8) there are five grids, a cathode, and an anode. With these seven electrodes, all functions of oscillation and frequency mixing take place. A typical circuit is shown in Fig. 7–16. Grids 1 and 2 and the cathode form a triode oscillator with external connections to inductances, etc. Grid 2 (actually a pair of side rods) acts as the anode of the oscillator. This anode plus the cathode of the tube can be looked upon as a "virtual" cathode supplying for the remainder of the tube an electron stream which varies at oscillator frequency.

Grid 3 is a positive screen like that in any screen-grid tube. Grid 4 is connected to the signal input, and its voltage varies at the frequency of this input signal. As the electron stream flows past grid 4 (which is already modulated at oscillator frequency), it gets its second modulation from grid 4. The final grid is a continuation of grid 3 or the screen grid and is followed by the anode which collects the doubly modulated electron stream. The screen grid hurries on the electrons by its positive potential and at the same time acts as a screen to prevent unwanted electrons from getting to grid 4.

A good example of the tube designer's ingenuity is evidenced by a modification of this type of tube. In the 6A8 tube and circuit, good frequency conversion takes place up to moderately high frequencies, but, as the frequency rises, the performance decreases for two principal reasons. In the first place, the output of the oscillator falls off at the higher frequencies; in the second place, there are certain unavoidable couplings between the oscillator and the signal frequency sections. These couplings increase as the frequency increases.

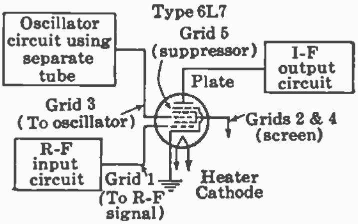


FIG. 8–16. Mixer tube and circuit useful at high frequencies.

In tubes like the 1R5 and the 6SA7 these faults are overcome. Here grid 1 is the oscillator grid. The second grid is connected inside the tube to grid 4 and acts as anode for the oscillator but at the same time as shield or screen to protect the signal grid, 3. Grid 5 acts as suppressor. In these tubes the space charge around the cathode is unaffected by electrons from the signal grid or by the electrostatic field of this grid. Thus r-f voltage on the signal grid has little or no effect upon the cathode current, and, therefore, changes in a-v-c bias have little effect upon the oscillator transconductance or the capacitance of grid 1. The effect is an increased freedom from troubles due to oscillator detuning by variations in a-v-c voltage.

In the 6K8, still another tube structure is employed in which the oscillator grid and grid 1 are the same. The functioning of the tube is not critical with respect to changes in oscillator plate voltage or signal grid bias. It is used in all-wave receivers where it is important to reduce frequency shift at the higher signal frequencies.

The 6L7 tube, which requires a separate oscillator, is a mixer tube especially adapted for high frequencies. It has two control grids, 1 and 3. The first is a remote cut-off type and the second a sharp cut-off. Incoming signal voltages are applied to grid 1

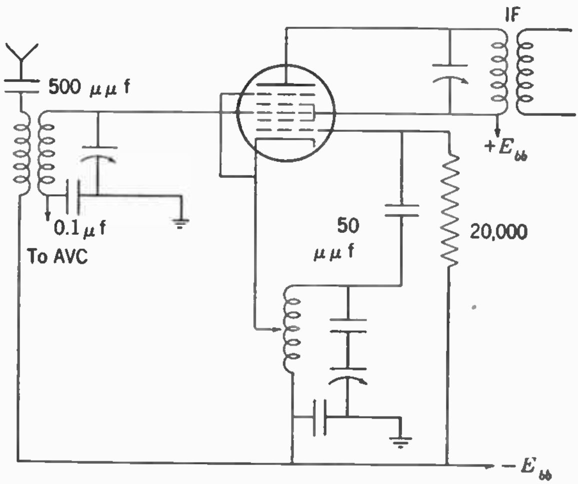


FIG. 9–16. Use of 6SA7-type of tube as frequency converter.

and oscillator voltages to grid 3. Grids 2 and 4 are connected and act as shield for 3 and at the same time accelerate the electrons. The final grid 5 is a suppressor. The 6L7 is known as a pentagrid-mixer tube.

In the 6J8G, the oscillator grid is connected internally to the injector grid. This tube is actually a 6L7 and a triode oscillator assembled over one another and placed in the same envelope. The 6K8 is more like a 6A8 with the anode grid rods replaced by an anode which is on one side of the cathode and with the mixer anode on the other side.

The **translation gain** of circuits using the pentagrid converter, like the 1A6, is of the order of 40. The term translation gain describes the gain in converting radio frequency into intermediate frequency. Thus a microvolt of r-f energy put into the frequency changer becomes 40 μv of intermediate frequency in the output.

The frequency converter tube is often called the "first" detector, and the entire receiver is often called a "double detection" system since there are two detectors. The signal which produces the intermediate frequency by being added to the local oscillator frequency is called an "image" frequency; the i-f signal is called the "beat" frequency, and the process of frequency conversion by heterodyne action is known as "beating" one frequency with another.

Choice of intermediate frequency. When preceded by some preselection, the chance of image-frequency trouble decreases as the intermediate frequency increases. At the same time trouble from stations separated from the desired station by the intermediate frequency is decreased. For these reasons it is desirable to have a high intermediate frequency.

With a low intermediate frequency the chances of harmonic troubles due to energy from the second detector getting back to the input of the set are lessened because only the higher and weaker harmonics would fall in the broadcast band. At the same time, lower frequencies are amplified better in the i-f amplifier. And, finally, more selectivity against adjacent channels can be obtained if the intermediate frequency is low.

All manner of beat frequencies are used, from 175 kc to several hundred kilocycles, depending upon the manufacturer, the type of set, the input frequencies to be received, etc. It has become almost standard practice to use intermediate frequencies in the vicinity of 456 kc for home radio receivers. In other types of receivers the intermediate frequency may be as low as 60 kc. The image-frequency ratio of a high-class receiver is about 20,000 to 1. That is, the voltage of a signal differing from the desired signal by twice the intermediate frequency must be 20,000 times as strong as the desired signal to produce the same output.

Problem 1-16. A broadcast receiver is to cover the band from 1600 to 550 kc. If the oscillator is tuned higher than the input circuit so that a beat frequency of 456 kc is produced, what must be the frequency band of the oscillator? Suppose that identical inductances are used in the tuned circuits and that the maximum capacity of the input circuit tuning condenser is 250 $\mu\mu$f. What will be the approximate maximum value of the oscillator tuning condenser?

Problem 2–16. The intermediate frequency is 456 kc in an all-wave receiver. What frequency band must the oscillator cover if one of the short-wave bands tunes from 4.5 to 12 megacycles?

Frequency converters. Any tuned r-f receiver can be converted into a superheterodyne by the addition of an oscillator and a mixing tube. In such a system the amplifier of the t-r-f set is used as the amplifier and is adjusted to give high amplification at any desired frequency within the range over which it will tune. Thereafter, all tuning is accomplished by varying the frequency of the external oscillator. The beat frequency between the incoming signals and the locally generated signal passes through and is amplified by the t-r-f receiver.

Band-pass amplifiers. The ideal response characteristic of a receiver would be a flat top with very steep sides. The flat top should be approximately 10,000 cycles wide, or more, and the sharper the sides the more selective would be the receiver without cutting off the higher modulation frequencies. An amplifier with such a characteristic may be said to be a "band-pass" amplifier, meaning that it passes and amplifies a band of frequencies and rejects all falling either above or below that band. Numerous attempts have been made to accomplish this effect.

By means of coupled circuits it is possible to approach the ideal. Thus in Fig. 10–16 if both coils are individually tuned to the same frequency, and then coupled together, the resultant response characteristics may be a single narrow-topped curve, like that of a single tuned circuit, or a flat-topped curve, or a curve of two more or less widely separated peaks with a hollow between, depending upon the degree to which the circuits are coupled.

In Fig. 10–16 is given the result of coupling two such circuits together with various degrees of coupling. It will be seen that too close coupling gives the widely separated peaks, proper coupling gives a comparatively flat-topped characteristic, and too loose coupling gives a sharply tuned circuit. In a multistage circuit, one might use a combination of two coupled circuits, one like 3 and the other like 1 or 2 in Fig. 10–16. The over-all effect would be better than one alone.

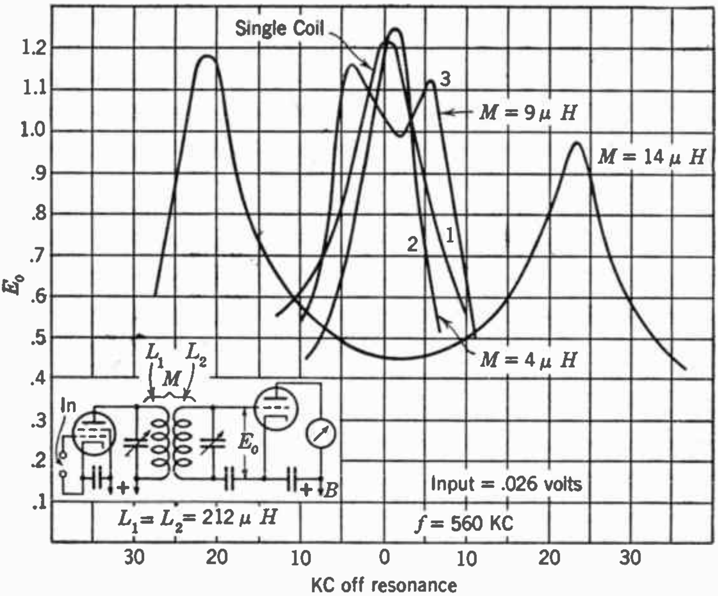


Fig. 10–16. Experimental determination of effect of close coupling in band-pass filter at high frequencies.

Superheterodynes can use such a band-pass circuit in the i-f amplifier. Here the frequency is fixed, and one adjustment will do for all signals put into it. When the band-pass arrangement is used at broadcast frequencies, the width of band passed may differ at each frequency to be received. If the coupling is by inductance the band will be broad at the high frequencies; if the coupling is capacitive, the curve will be broad at low frequencies. Some combination may be arranged so that a more or less uniform band is passed at all broadcast frequencies.

Modern practice is to approach the effect of band-pass tuning by using an intertube transformer with both primary and secondary tuned. The coupling is adjusted to get the desired broadness at the top of the response curve.

Second detector. After sufficient amplification at the intermediate frequency is obtained, any type of detector may be employed to remove the audio modulation from the intermediate frequency. Modern practice is to use a diode. This is usually followed by a single stage of audio voltage amplification and then by the power amplifier. The direct currents produced by the rectification or detection process and flowing in the diode output circuit are used to maintain the voltage input (automatic volume control) to the detector more or less constant in spite of variations in the input voltages secured from the antenna.

Tone control. Manual control of the tone emitted by the loud speaker is incorporated in many broadcast receivers and in some receivers made for other purposes. Frequencies at either end of the audio range may be accentuated or decreased with respect to each other or to the middle region. In general, a tone control is a device for reducing the high-frequency response of the receiver to get rid of noise either generated inside the receiver or coming in with the signal. It consists of a resistance in series with a capacitance, the combination being shunted across the input to the audio amplifier. The capacity shunts out the higher audio tones; the resistance, which is the variable feature, increases or decreases this shunting effect. In other words, the usual tone control is nothing but a device to eliminate high notes.

Automatic tone control. When static is bad, as when receiving a weak station, cutting out the higher audio notes is advantageous. Often a program that is hopelessly lost in noise can be received with a certain amount of pleasure if the tone control is advanced to the point where little beyond 2000 cycles is received.

Circuits have been devised which give some measure of automatic tone control. Thus when the receiver r-f gain is low, as it would be if receiving a strong local station, the tone-control function is not in operation. Then as weaker and weaker stations are tuned in, this tone-control function, which is connected with the a-v-c system, cuts off more and more of the high frequencies. In certain circuits this function is inherent; in others, an additional tube acting as a variable resistance or reactance is employed.

Such circuits have not come into general use. Automatic selectivity control (a-s-c) circuits which make the set less selective when receiving strong signals and more selective on weak signals will probably be incorporated in receivers ultimately. The circuit in Fig. 11–16 uses overcoupling in an interstage transformer for the purpose of broadening the tuning so that higher audio frequencies can be heard. A simple switch connects or removes the additional coupling inductance.

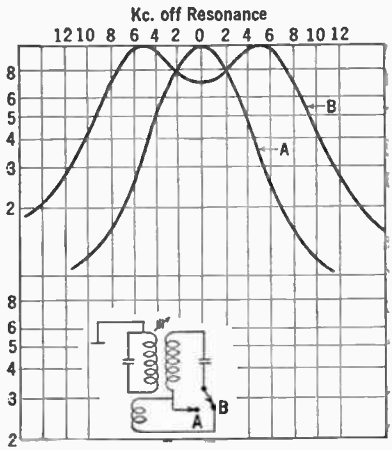


Fig. 11–16. Selectivity curve for two-position i-f transformer.

Noise suppressor systems. In receivers having automatic volume control the sensitivity is at maximum when no signal is being received, for example, between carrier signals. Therefore, any static or other radio noise will be amplified to the limit of the circuit and be passed to the loud speaker. Circuits have been developed which shut off this noise. They operate in various ways. For example, a tube may be so connected that its plate current will be high in the absence of signal. This plate current can be used to supply bias voltage to the a-f systems. Absence of signal thus will over-bias the audio amplifier, preventing any sound coming from it.

These are known as "quiet a-v-c" or "squelch" systems. They are not found in home receivers to any extent, especially since push-button sets, with which no interchannel noise is heard, came into vogue.

Receivers used on fixed frequencies, such as police receivers, almost always have squelch systems so that nothing is heard by the operator until the transmitter comes on the air.

Automatic frequency control. Another automatic feature which has been incorporated into modern receivers is the maintaining of the oscillator frequency accurately at the desired value. The frequency at which an oscillator functions is dependent upon several variables: one of them, the voltage on the plate; another, the capacity of the input (grid to cathode). If, after a receiver is tuned to the desired station, these values should change because of temperature or line voltage variations, or for any other cause, the oscillator would not be tuned to the frequency which would provide the required intermediate frequency.

Furthermore, the difficulty of accurately tuning a highly selective superheterodyne with several tuned stages is very real. When it is tuned off resonance with the desired signal, bad distortion results.

How automatic frequency control works. A fundamental circuit for an a-f-c system is shown in Fig. 12–16. An i-f amplifier feeds two tuned secondary windings of a transformer by means of two parallel primaries. The upper secondary is tuned slightly higher and the lower slightly lower in frequency than the exact intermediate frequency. If the frequency from the i-f amplifier is exactly between the two frequencies to which the secondaries are tuned, the secondaries have the same voltage across them and the two diodes have equal voltages impresed upon them.

If, however, the intermediate frequency is slightly higher than the correct value, then the upper tube has somewhat greater voltage impressed upon it than the lower tube, and the a-f-c bias voltage will be negative with respect to its original value. If, now, the intermediate frequency is below the correct value, the lower tube will produce a higher current and the a-f-c bias will be positive. Proper design will make the a-f-c voltage linear with respect to the amount the intermediate frequency is off the correct value.

These variable voltages are applied to the grid of a tube which acts as a variable reactance. The variable reactance is connected into the tuned circuit, which is part of the oscillator, with the result that positive reactance (inductance) is added to the tuned circuit when the oscillator tends to tune off frequency in one direction, or negative reactance (capacitance) is

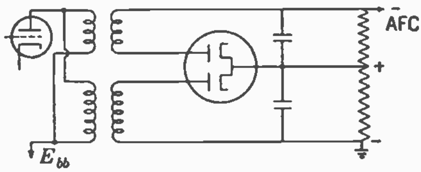


FIG. 12–16. Basic automatic-frequency-control circuit.

added when the oscillator tends to drift off frequency in the other direction.

These circuits will be described in greater detail in connection with frequency modulation.

Not only do a-f-c systems keep a receiver tuned to the correct frequency, but also they aid the operator in tuning his receiver. When the circuits are tuned to the vicinity of resonance with the incoming signal, the a-f-c system takes over the tuning process and forces the oscillator to generate the correct frequency to heterodyne the signal to the proper intermediate frequency.

Although automatic frequency control was a splendid development and worked successfully, it required more tubes and more apparatus, and receiver manufacturers shied away from adopting the system as standard. The virtues of receivers which do not go out of tune were made evident, however, and soon components manufacturers began to improve their condensers and coils so that shifting of frequency with time or temperature variations was greatly reduced.

Pushbutton receivers. Three types of the pushbutton tuning system have been applied to receivers, one or the other of which is now found on most of the automobile sets and on many home receivers as well. The advantages are simplicity, quietness, and speed of tuning from one station to another.

In one system, individually tuned circuits are provided for each pushbutton. These circuits are tuned to the desired stations by variations in capacitance or inductance. The tuning may be effected by a service man or the owner by simple screwdriver adjustments. The pushbuttons are mechanically connected to switches which select the required tuned circuits. The number of stations that can be tuned in is limited by the number of pushbuttons provided.

In a second system, the gang-tuning condenser may be rotated throughout its range by means of the conventional dial. In addition, cams are attached to the condenser shaft so that when a pushbutton is depressed the shaft is rotated to the proper position to get the desired station. The position of these cams on the shaft may be changed easily.

In a third system an electric motor is utilized to turn the condenser shaft. This method is more expensive and more likely to get out of order, but it permits remote control and has the advantage that any desired channel may be selected.

None of these systems would work well without the development of stable components or without automatic frequency control. The trend is toward the use of components, i.e., condensers and inductances, which do not vary in their electrical characteristics with time, temperature, or humidity.

Tuning indicators. Highly selective receivers must be correctly tuned; otherwise severe distortion results. Early examples of receivers of this type utilized d-c milliammeters with various kinds of indicating needles to advise the operator when the receiver was tuned to the exact center of the pass band. The current for operating these indicators was obtained from the a-v-c system, the greatest amount of a-v-c voltage being obtained when the receiver was tuned to the center of the i-f band.

Introduction of electron-ray tubes simplified the tuning indication problem. These tubes indicate visually the condition of correct tuning by means of a fluorescent target which is bombarded with electrons and which glows when so bombarded. There are several types, but each depends upon the same fundamental principle. Between the cathode (source of electrons) and the fluorescent target is a ray-control electrode. Variations in voltage on this electrode control the flow of electrons to the target and therefore control the amount of the target that glows.

In one tube, such as the 6N5, there is a triode in addition to the ray tube. The triode acts as a d-c amplifier. The grid of this tube is supplied with voltage from the a-v-c system. In Fig. 13–16 the flow of plate current, under control of the voltages supplied to the grid of the triode through the resistor R, fixes the voltage of the control electrode of the indicator tube. When the voltage of this electrode decreases, less of the target is illuminated. When more plate current flows from the triode through R the voltage drop in R increases and there is less voltage on the control electrode of the indicator tube. The shadow on the target increases. When the triode grid is more negative, less plate current flows, more voltage is across the electron-ray tube, and the target becomes brighter.

In another tube of this general type there is no triode amplifier but there are two control electrodes, one on either side of the cathode, and both are brought out to base terminals. Thus, two symmetrically opposed patterns or two unlike patterns may be obtained, depending upon how the ray-control electrodes are connected. This tube requires an external d-c amplifier.

Still other tubes have sharp cut-off or remote cut-off ray-control electrodes.

In operation, the electron-ray tubes get their controlling voltage from the a-v-c system. Since the maximum of negative voltage is produced in the a-v-c circuit when the receiver is tuned to resonance with the desired station, a maximum of negative voltage is supplied the control grid of the triode acting as d-c amplifier for the indicator tube. Under these conditions the least plate current will flow through the resistor R, the greatest voltage will appear across the ray tube, and there will be the smallest shadow angle on the target.

Electron-ray indicator tubes are found in many measuring instruments, taking the place of indicators of the milliammeter type. The sensitivity of the indication may be increased by using a separate d-c amplifier to control the action of the ray-control electrode as shown in Fig. 13–16.

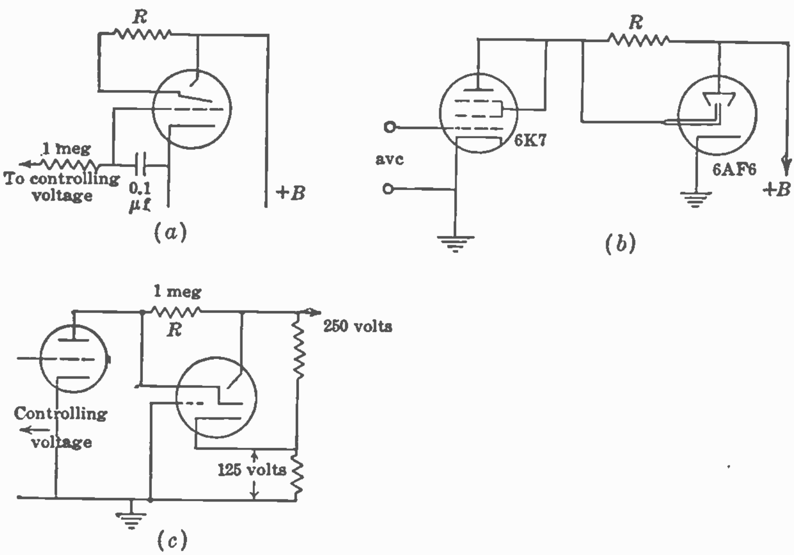


FIG. 13–16. Electron-ray tuning indicator tube circuits. In (b) and (c) separate d-c amplifiers are used, whereas in (a) the triode amplifier is in the same envelope with the indicator element.

Loud speakers. The loud speaker is usually the final link in a radio system, and because of its position with respect to the listener it is frequently blamed for much of the bad reproduction that really originates somewhere else.

The task of the loud speaker is to translate into sound energy the electrical energy in the power tube. It must do this as effectively and faithfully as is necessary for the particular job for which the radio is used. It is useless to design and operate a high-class amplifier with a poor loud speaker. The wide range of tones coming from the amplifier is lost in the loud speaker and does not get to the listener. Likewise, it is absurd to install a perfect loud speaker in the hope that the fidelity of reproduction from a poorly engineered receiver will be bettered. The full benefit of a wide-range loud speaker cannot be attained until the complete chain of apparatus is perfect—amplifier, power tubes, plate-voltage supply, etc.

Elements of the loud speaker. Every speaker has three essential elements. First, a **motor** converts electrical into mechanical energy which is imparted to a **diaphragm.** The diaphragm, the second element, vibrates and sets air in motion to form sound waves. Third, a **horn** couples the diaphragm to the impedance of the air load. Thus the horn is a sort of transformer or transmission line to couple the impedance of the diaphragm to the load, which is the air that is to be set in motion. Some loud speakers have no horn, proper, but the diaphragm is extended in size and accomplishes the needed coupling by itself.

Each of these elements may take one of several forms, and one assembly may be considerably more efficient than another, that is, more of the electrical energy may be converted into acoustic energy in one type of speaker than in another. Some speakers will handle quite a lot of electrical energy with relative freedom from distortion. Other types are limited in the volume they can handle. Some types will reproduce a wide range of frequencies. and others will convert electrical energy into acoustical energy only over narrow frequency bands.

The engineering of loud speakers is quite as complex as the engineering of an audio system or the designing of a radio receiver. The terms and quantities used are different, but there are certain analogies between the electrical and mechanical-acoustical systems. Each involves coupling a generator to a load at as great efficiency as possible over as wide a frequency band as necessary.

The horn speaker. This type of speaker is highly efficient, perhaps 30 to 50 per cent. A source of sound of small physical volume but high intensity is coupled to a large volume of air where the intensity of vibration is low. The horn must permit the air vibrations (the sound wave emitted by the source) to expand in such a manner that they will leave the horn without reflections back into the horn. The small end of the horn is called its **throat;** the large end, its **mouth.** The rate and manner in which the horn changes its shape are known as the **taper.** The relative dimensions of the mouth and throat determine the load on the diaphragm which produces the sound at the throat end.

The impedance seen by the diaphragm, looking into the horn, should equal the impedance of the diaphragm for best transfer of power. The lowest frequency that can be radiated into space by the horn is controlled by the mouth of the horn. The area of the mouth should be the area of a circle having a diameter one-fourth wavelength at the lowest desired frequency.

If the impedance seen by the diaphragm (source of sound) is not high enough, an "air" transformer is used. This is a chamber in which air pressure may be built up. Its dimensions are critical; its volume should be small, and its radius should not be greater than one-fourth wavelength at the highest frequency to be reproduced.

The taper usually follows an exponential law; that is, the cross-sectional area increases at a rate which is proportional to the area. The taper controls the efficiency of sound transmission along the length of the horn. High frequencies are transmitted quite well, but at the lower frequencies the propagation falls off. Until 1941, the exponential horn was most used because it gave the best transmission at low frequencies of any taper utilized up to that date. In 1941, however, a new horn design (known as the Hypex) was worked out mathematically and put into production. At the low frequencies it performs appreciably better than the horns using exponential tapers.

Diaphragm speakers. The horn is large and, therefore, unfitted for use in home radio receivers. For this service the cone speaker has been developed to a high state of efficiency. To transmit from a localized source of sound to an acoustic load spread over a large area, the cone diaphragm must be light (to avoid loading up the motor with a resistance load) and at the same time it must be rigid so that it does not break up into all sorts of vibrations when driven by the motor. The highest frequency that will be transmitted satisfactorily will depend upon the size and weight of the cone, a small cone transmitting high frequencies better than a large one. On the other hand, to couple energy at low frequencies to the air load, the cone must be large, in order to move a large mass of air. Thus the loud-speaker designer is in a dilemma. If a wide range of frequencies is to be reproduced, he compromises by using two speakers, one large to reproduce the low frequencies, and one small, often called a "tweeter," to reproduce the high frequencies. They may be on the same axis or may even be part of the same diaphragm, or they may be distinct speakers, physically displaced from each other.

Cone speakers can be made which will "sound loud" although really little acoustic energy is being radiated. They distort badly, but the listener is happy if the sound seems loud. Speakers can be made which will sound as though they reproduced low frequencies, although the cone is small. Again, the distortion is high; the low frequencies are really synthetic, and a comparison with a good low-frequency speaker will instantly show the difference.

The moving-coil speaker. The construction of the moving coil or dynamic type of speaker is shown in Fig. 14–16. A strong permanent magnet, or an electromagnet energized by a direct current from a storage battery or from part of the plate-current supply system or from the 110-volt a-c line by means of a rectifier, furnishes a steady field. The voice-frequency currents coming from the final power tubes in the amplifier are passed through a few turns of wire around a small movable coil to which is attached the diaphragm or cone. When alternating currents flow through the coil, it tends to move at right angles to the lines of force across the air gap. These motions of the coil are imparted to the cone and thence to the air.

The impedance of this movable coil is very low, of the order of 5 to 10 ohms, and is almost constant at audio frequencies. The tube feeds power into this impedance through a step-down transformer. Because the coil can move through a considerable distance, good low-frequency response is possible. Considerable sound energy can be got from such a speaker. The resonant frequency of the moving part is usually lower than the lowest audio tone to be reproduced.

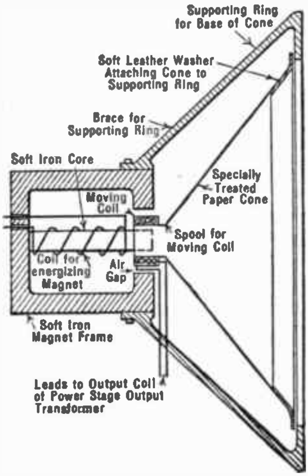


FIG. 14–16. Modern "dynamic" or moving-coil loud speaker.

The impedance of the speaker tends to rise at high audio frequencies, owing to the inductance of the moving coil, so that distortion is caused when the speaker is operated from a high-impedance source. Therefore condensers are shunted from plate to cathode of pentodes and beam power tubes to lower the high-frequency impedance when these tubes are used.

Baffles for dynamic speakers. It is necessary to install a dynamic speaker in the center of a rather large and heavy "baffle" if the low notes are to be properly reproduced. Otherwise the wave set up from the back of the cone can interfere with the wave set up by the front, and little or no sound will get to the listener. The baffle increases the air path between front and back and should be great enough so

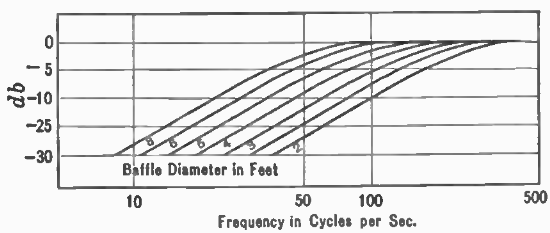


FIG. 15–16. Dynamic speaker response as controlled by baffle size.

that the shortest mechanical path between front and back edges is at least one-quarter wavelength for the lowest note to be received. Since the wavelength of sound, like that of radio waves, is equal to the velocity at which it travels (1100 ft per sec in air) divided by the frequency, it is not difficult to prove that a baffle at least 32 in. square is necessary for notes as low as 100 cycles, and 110 in. for notes as low as 30 cycles. When the unit is mounted in a box, peculiar resonances are set up which spoil the good qualities of the moving-coil speaker. These resonances are sometimes smoothed out by the introduction of resonating chambers or diaphragms which absorb energy at the offending frequency.

Improvement in loud speakers. Several methods have been developed for increasing the bass response of loud speakers (for

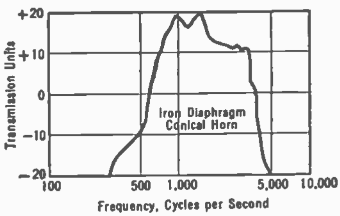


Fig. 16–16. Limited response of horn speaker

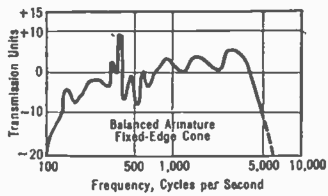


Fig. 17–16. Response characteristic of good cone.

it is at the low frequencies that it is difficult to get good coupling between motor and load) and for eliminating troublesome resonances. One method of reducing resonance in the cabinet was to place cones (not driven by motors) in the front of the cabinet. These cones were of such dimensions that they were resonant to the frequencies at which the cabinet "boomed," they absorbed energy and acted as dissipators of the unwanted resonance energy.

To eliminate the loss of low frequencies because the wave from the back of the diaphragm, at low frequencies, could cancel some of the energy radiated from the front of the diaphragm, several schemes have been applied. One makes use of a total enclosure like a box closed on all sides. Ordinarily with such a cabinet the system will boom at some frequencies and at low frequencies may be quite inefficient. If, however, vents or holes of the right size are placed in the front of the enclosed cabinet correctly with respect to the speaker diaphragm, the low-frequency response will be increased appreciably and at the same time the distortion at low frequencies will be much less.

The "bass reflex" principle just described makes it possible to improve the low frequencies as much as 10 db compared with

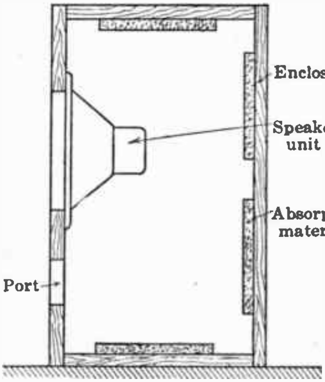


FIG. 18–16. Bass reflex housing for improving low-frequency response of dynamic speaker.

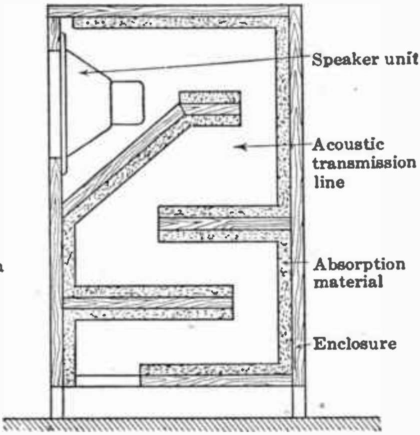
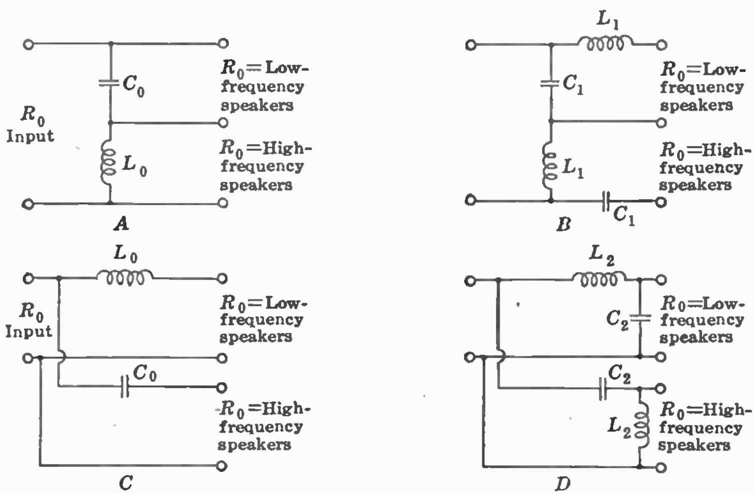


FIG. 19–16. Construction of the labyrinth speaker enclosure.

a total enclosure without the vent. The distortion may be reduced to as little as one-third. The vent acts like a second loud speaker which furnishes acoustic power only at frequencies near the cut-off frequency of the speaker. Other frequencies are absorbed within the cabinet by lining it with material which is highly absorptive at the intermediate and higher audio frequencies. The speaker diaphragm looks into a high anti-resonant impedance at the low frequencies, and for this reason the mechanical movement of the diaphragm is small. At high excursions of the diaphragm, distortion is created. The vent supplies most of the low-frequency energy, and since there is nothing in it to move (no mechanical parts) there is no non-linear flux or non-linear edge stiffness of the paper cone. No distortion is created.

Another method of improving low-frequency response is to connect a transmission line (acoustic) to the diaphragm in such



$L_0 = \frac{R_0}{2\pi f_a}$ $\qquad L_1 = \frac{L_0}{\sqrt{2}}$ $\qquad L_2 = \sqrt{2}\, L_0$

$C_0 = \frac{1}{2\pi f_a R_0}$ $\qquad C_1 = \sqrt{2}\, C_0$ $\qquad C_2 = \frac{C_0}{\sqrt{2}}$

f_a= Cross-over frequency of network
Inductances in henries
Capacitances in farads

FIG. 20–16. Dividing networks to keep low frequencies out of high-frequency speaker, and vice versa.

a manner that it brings the back-side radiation of the diaphragm in phase with the front-side radiation and thereby improves the response. The acoustic transmission line is called a "labyrinth"; it is shown in Fig. 19–16. Distortion at the low frequencies is reduced compared with a cone speaker without the labyrinth.

Other methods involve a series of resonant pipes placed in the cabinet, and, of course, the two-speaker system is widely used when a wide range of frequencies must be reproduced. In this system a tweeter is used for the high frequencies and a large speaker called a "woofer" for the lows. By means of a "dividing network" the low frequencies are kept out of the small speaker and the high frequencies out of the large speaker. Thus each is required to handle only that power which will produce useful acoustic output. Each speaker can be engineered with its own restricted range in mind, rather than expecting a single speaker to cover the whole audio range equally well.

Loud-speaker measurements. One method of measuring the performance of a loud speaker is to hang a calibrated microphone in front of the speaker which is actuated by various tones of known amplitudes from an oscillator. The output of the microphone is amplified and measured, and thus a curve of output versus frequency may be obtained. Another method is to rotate a microphone through a rather large arc in front of a loud speaker. This tends to average the effects due to resonances in the room in which measurements are made.

CHAPTER 17

OSCILLATORS

Not only will an amplifier tube increase the amplitude of an alternating voltage placed upon its input but it will also generate alternating currents in the absence of any exciting voltages placed upon the input from any external source. The frequency of these currents may be anything from a few cycles per hour to many hundreds of millions of cycles per second; the efficiency of generating a-c power from d-c sources (batteries, etc.) may be as great as 70 per cent or more; the stability and constancy of the frequency and amplitude may be as great as is desired practically; the power output may be as great as thousands of kilowatts.

How does an amplifier tube produce alternating current from a d-c source?

Consider first a simple amplifier. Suppose that 1 volt placed upon the grid-cathode circuit appears as 10 volts across a load resistance. The circuit and tube amplification, therefore, is 10. Now suppose that by some sort of network we can take 1 volt from the output and insert it in the input in series with the original input voltage. The grid-cathode input circuit now has 2 volts across it, and the output voltage will be 20 instead of 10 provided that the amplifier works as well as before. We can now remove the original 1 volt input exciting voltage and the output will drop to 10 volts, but the tube now furnishes its own excitation. It has become a self-excited oscillator. It is producing a-c power from d-c power furnished by the batteries or by a voltage supply system of any sort attached to the cathode, grid, and plate terminals.

There is only one catch—the voltage secured from the output and inserted into the input must have the proper magnitude and the proper phase. If the phase is reversed, the feedback voltage will be subtracted from the original input voltage and the output voltage will be reduced and not increased. If the external excitation is removed under these conditions, the output will soon fall to zero and no a-c power will be generated.

If the power fed back to the input from the output is insufficient to supply the losses of power in the input circuit, even if this feedback of power is in the correct phase, then there will be no generation of a-c power because all the feedback energy is used up in heating the resistance of the input. None is left over to excite the tube to produce its own power.

When an amplifier of this general sort is producing its own a-c power, it is said to be an "oscillator" and to be producing oscillations.

Practical oscillators. Many circuits have been devised for feeding back energy from output to input in correct amplitude and phase. In none of them are there any external exciting voltages. When the source of electrons is available and when plate and grid voltages are supplied, any slight disturbance, such as the change in plate current during the warming-up period, will start an originating oscillation through the system. Part of the output of this oscillation is fed back into the input automatically, and soon the production of power is proceeding at the amplitude and frequency determined by the tube and circuit constants. The tube and circuit are generating power.

The tuned plate or tuned grid circuits are easy to understand since the functions of the individual components of the circuit are separated. The feedback coil is merely coupled to the tuned circuit, and the extent of the coupling (the mutual inductance between the two inductances) determines the amplitude of the feedback voltage while the direction of the coupling determines the phase. If the connections to the feedback coil (tickler) are reversed, the circuit will not oscillate.

Figure 1–17 represents a straight "tickler" feedback or tuned grid oscillator. In Fig. 2–17 the tuned circuit is in the plate circuit and feedback is provided by a grid tickler. In Fig. 3–17 both grid and plate circuits have tuned circuits in them, feedback being provided by the grid-plate capacitance within the tube and the capacitance of the leads.

In all these circuits, note that the plate battery occupies its normal position, between the lower end of the plate-circuit ap-

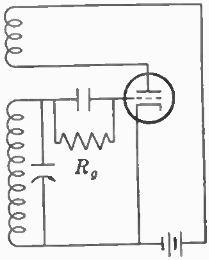


FIG. 1-17. Tickler feedback circuit.

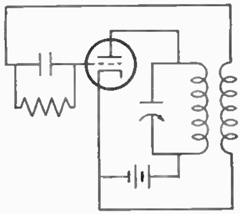


FIG. 2-17. Feedback coil in grid circuit.

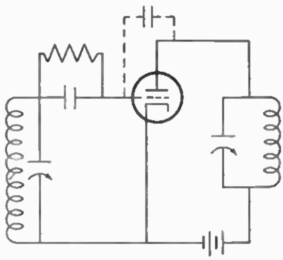


FIG. 3-17. Feedback through grid-plate capacitance.

paratus and the cathode. Note, too, that each circuit has a grid condenser and a resistance connected across it or connected from grid to cathode. This *RC* circuit is to provide bias to the grid in a manner to be described later.

In Fig. 4-17 is shown the Hartley circuit in which a single tuned circuit has a tap to which the cathode is connected. A portion of the alternating voltage developed across the inductance is applied to the grid as excitation. The tap is a means of adjusting the grid alternating voltage to the value best for efficiency or any other criterion the operator desires. Ordinarily the tap will be near the grid end of the coil since any voltage applied to the grid is multiplied by the amplification factor of the tube. A small grid voltage, therefore, multiplied by the

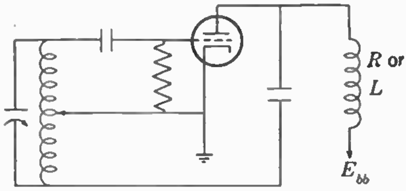


FIG. 4–17. Hartley oscillator.

amplification factor will produce the desired sustained alternating current in the tank circuit.

In Fig. 5–17 the correct voltage division is secured by a split condenser. This circuit is often used in superheterodynes as the oscillator. The upper condenser is the tuning capacitor; the lower condenser acts as the padding capacitor to make the circuit tune to the correct intermediate frequency over the whole frequency band over which the receiver will be used.

Note in Figs. 4–17 and 5–17 that the plate voltage is supplied through a shunt circuit composed of a resistance or a choke coil.

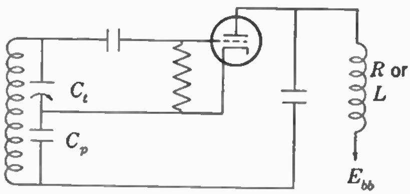


FIG. 5–17. Colpitts oscillator.

The virtue of this method lies in the fact that no d-c voltages exist on the tuning coils or condensers. "Shunt feed" may be employed on any of the circuits shown here. In the Hartley circuit, neither terminal of the tuning condenser is at ground potential, and the circuit, therefore, is somewhat subject to "hand capacitance" effects. If the condenser is properly adjusted with a long insulated shaft, and then if the hand is brought near the plates of the condenser, the frequency will change because the condenser capacitance has been effectively increased by the capacitance of the operator's body with respect to ground.

In any of these circuits, a combination of more than one kind of feedback may be employed. For example, if in Fig 5–17 a plate tickler is employed in addition to the capacitive feedback through C_p, the oscillator output can be made quite uniform over a range of frequencies. Since some feedback is provided by C_p, the number of turns on the tickler coil can be quite small so there is little danger that this coil will become resonant at a high frequency and cause trouble.

If the tuning condenser across the grid coil in Fig. 3–17 is removed and the inductance increased accordingly, the Q of the grid circuit will be low (because of the large number of turns on the grid coil) and the circuit will be rather broad in tuning, that is, not so critical with regard to tuning. The tuned plate circuit, therefore, is the frequency-determining element.

If the tuning condenser in Fig. 4–17 is connected across the plate part of the circuit only, the Hartley circuit becomes the reversed feedback circuit of Fig. 2–17.

A resistance feedback that is employed in laboratory oscillators is shown in Fig. 6–17. As in all such circuits, the grid excitation is greatest when the largest coupling between grid and plate coils is used. Then the plate-current variations are considerable in magnitude, they use curved parts of the characteristic, and harmonics are generated. In a laboratory oscillator where harmonic production is to be kept to a minimum, the coupling between input and output circuits should be adjusted

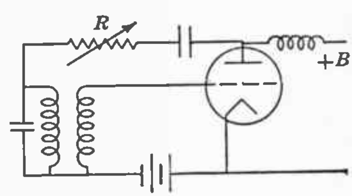


FIG. 6–17. Circuit often used in low- (audio-) frequency oscillators. Feedback is provided by resistance, which also loads up the plate circuit and straightens out the E_g-I_p characteristic.

at each frequency to the least possible amount that will insure stable oscillations. The resistance feedback method is useful in such oscillators because of the mechanical ease of adjusting the feedback voltage. The feedback resistance increases the total resistance in the plate circuit, and variations in tube resistance due to voltage variations are not so important. The circuit is, therefore, more stable. The dynamic characteristic of the tube becomes straighter because of the high resistance load, and harmonic generation decreases.

This circuit is used to generate audio frequencies for laboratory purposes. In all audio oscillators it is worth noting that direct plate current should not be allowed to flow through iron-cored inductances since core saturation is likely to occur with resultant generation of harmonics and wave-form distortion.

Adjusting the oscillator. When the oscillator is used to deliver power it is desirable to attain the adjustment which will either deliver maximum output or secure maximum efficiency so that greater inputs may be used. The tuned plate-tuned grid circuit has no adjustments. The operator tunes either the plate or grid circuits until the tube oscillates and there is nothing else he can do. In fact, if he does not tune the circuit to an oscillating condition the plate current may be very high unless the grid is biased with a battery instead of through a grid leak and condenser.

In the Hartley circuit, however, it is possible to move the center filament tap and so to get some control over the strength of oscillations, the feedback voltage, etc. In Fig. 7–17 is illustrated the result of varying the filament tap on a simple 40-meter oscillator of the Hartley type. The curves give the plate current, I_b, the current, I_a, into an antenna coupled to the plate coil, and the ratio of antenna current to plate current as a measure of the efficiency of the circuit. What is wanted is much antenna current and little plate current. The greater this ratio, the greater is the efficiency of the circuit.

The relation between plate voltage and oscillating current is shown by Fig. 8–17 to be linear. The grid bias measured across a 5000-ohm resistor is plotted too. The tube was a UX-210 oscillating at 1225 meters in a tuned plate circuit. The effect of changing the C bias resistor of a tube in the tuned grid-tuned plate circuit is shown by Fig. 9–17.

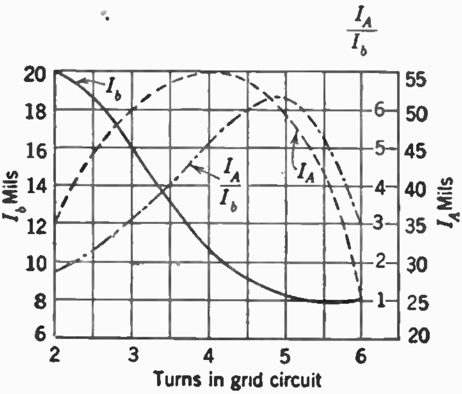


FIG. 7–17. Effect of varying excitation by changing position of center tap in Hartley circuit.

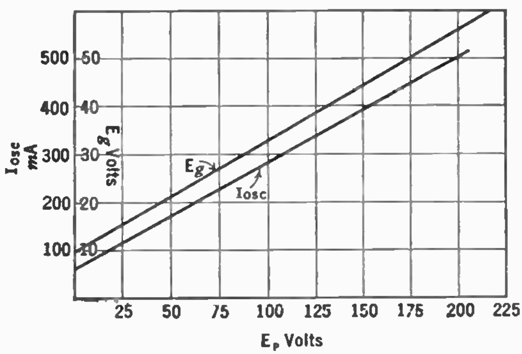


FIG. 8–17. Relation between plate voltage and oscillatory current and grid bias.

Electron-coupled oscillator. An ingenious circuit employing a screen-grid tube or its equivalent has been devised in which the frequency of the voltages produced is practically independent of the plate voltage and load variations. The simple circuit shown in Fig. 10–17 is called an electron-coupled oscillator be-

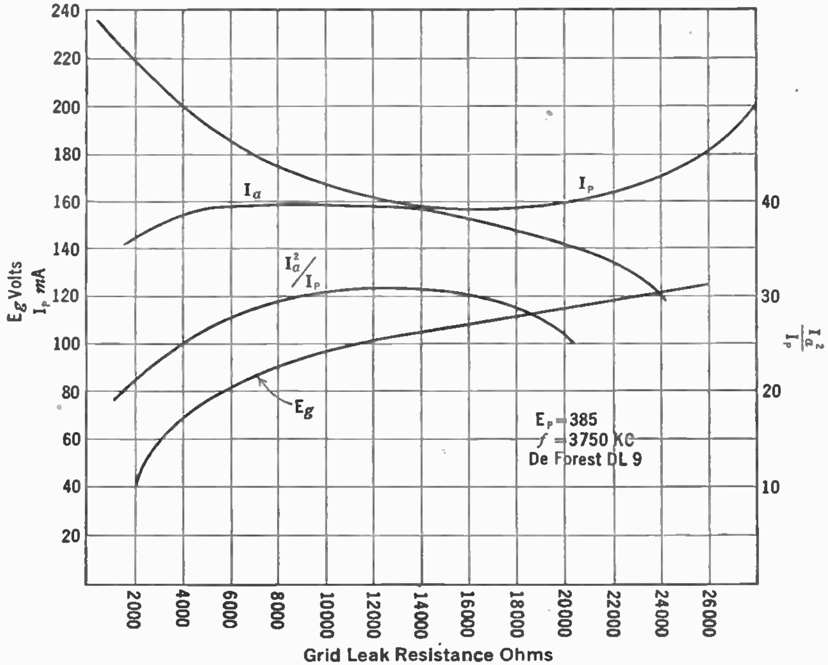


FIG. 9–17. Effect of varying grid-leak resistance.

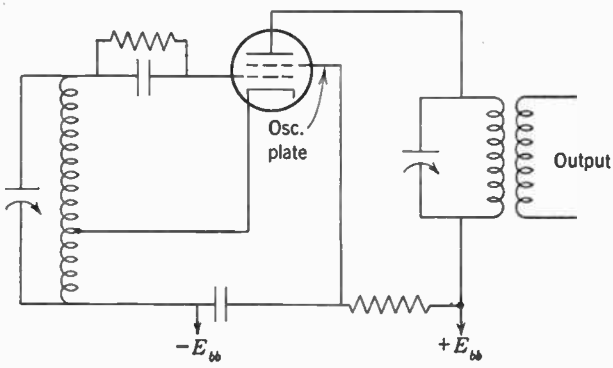


FIG. 10–17. Electron-coupled oscillator.

cause the output is coupled to the oscillator only through an electron stream. The actual oscillating circuit is made up of the apparatus external to the tube with the cathode and two grids acting as a triode oscillator. The screen grid acts as anode of the oscillator. Plate current will be modulated by the alternating voltages produced in the triode. At the same time the oscillator is effectively shielded from what occurs in the output or load circuit because the screen grid is at ground potential as far as alternating currents are concerned. As long as the ratio between screen and plate voltages remains constant (as is assured by taking both from the same source through a potential divider) the frequency will be independent of the actual voltage on the plate.

Electron-coupled oscillators are widely used in the frequency-converter stages of superheterodyne receivers.

Grid bias for oscillators. The steady plate voltage may be supplied either in series with the oscillatory circuits or in shunt with them. The steady grid voltage may similarly be supplied by a shunt method or by a battery in series with the grid oscillatory elements. The grid may be made to supply its own bias by forcing it to draw current during some portion of the oscillatory cycle, making this current flow through a resistance, and utilizing the voltage drop along the resistance as bias. Thus, in Fig. 1–17, grid current flowing through R_g produces the desired bias voltage.

If the output current increases, through some cause or other, the feedback excitation on the grid increases, more grid current flows through the grid resistance, and more negative bias is placed upon the tube. The increased bias decreases the output current. For this reason a self-biased oscillator tends to have a more stable output, as far as amplitude of oscillation is concerned, than if battery bias is employed. There is one disadvantage: if the tube stops oscillating for any reason, there is no bias on the grid, and the plate current may rise high enough to damage the tube. Power oscillators usually have some, if not all, of the bias obtained from a cathode resistor.

The power represented by the flow of grid current through the grid resistor must be supplied by the plate battery, and it lowers the over-all efficiency of the system.

Note in Fig. 11–17 how the steady plate current of a grid-leak bias oscillator decreases when oscillations begin. This is be-

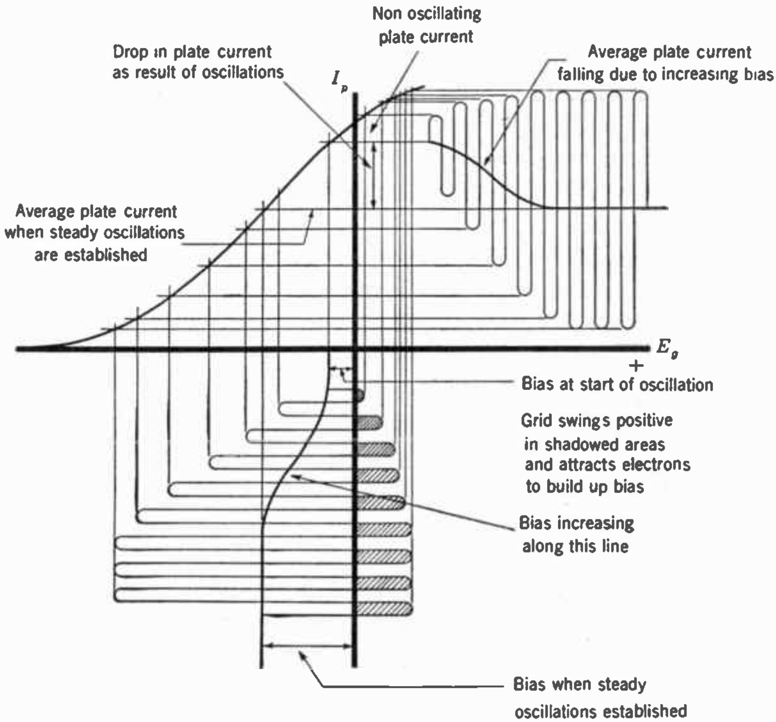


FIG. 11–17. Gradual build-up of oscillations with resultant increase in bias and decrease in steady plate current.

cause more electrons are attracted to the grid as oscillations build up, and the greater grid current flowing through the grid resistor produces a greater voltage drop.

Other types of oscillating circuits. In any *LC* oscillator a very great change in the wave form of the alternating currents produced will be obtained if the grid condenser and leak are made very large and if the coupling between grid and plate is made very tight. The grid bias now builds up very rapidly, and, since the time constant of the RC circuit is high, a considerable period must elapse before the voltage across the condenser decreases to the point where the tube can conduct current again. In other words, the bias builds up to the cut-off point, no current flows in the plate circuit for a period, and then the cycle starts over again. Short pulses of current are produced in this manner, the periods of conduction being separated by an interval long compared to the time of conduction.

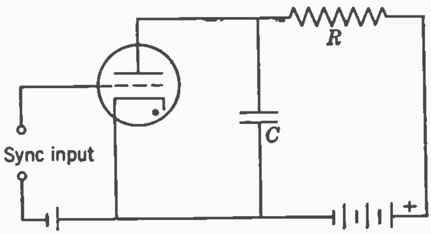


FIG. 12–17. *RC* oscillator using gaseous triode and with synchronizing voltages placed on grid to control time of condenser discharge.

The timing of the pulses can be controlled by an external timing voltage (synchronizing or "sync" voltage). This type of oscillator is often called a "blocking" oscillator since the grid blocks and no current flows. Such a circuit is employed in television and other applications of cathode-ray tubes.

Some circuits employ no L or C at all, relying on the blocking action described for proper operation. The wave form of the plate-current pulses may be sawtooth, square, or sometimes sinusoidal.

A very simple type of oscillator operates as follows. A gaseous triode is shunted across a condenser which is charged through a resistor from a source of potential. When the voltage across the condenser reaches the value at which the gaseous tube begins to conduct, the condenser discharges suddenly through the gas tube and then begins to charge again. The wave form of the voltage across the condenser has two distinct portions, a long slowly rising portion representing the time during which the condenser is charging through the resistor and then a steep drop representing the sudden discharge period through the tube. At this point the voltage across the condenser is too low to maintain current flow through the tube and it ceases to conduct, thus removing the effective short circuit across the condenser, which again recharges.

By proper design, the wave form of the output of the non-feedback oscillators may be made sinusoidal, and, by using two tubes in a symmetrical circuit, both halves of a sine wave may be produced.

Negative resistance oscillator. If, in any circuit, a *decrease* in current is produced by an *increase* in voltage, the dynamic resistance of that circuit is negative and oscillations will be sustained. Thus in Fig 13–17 a tuned circuit in the plate circuit

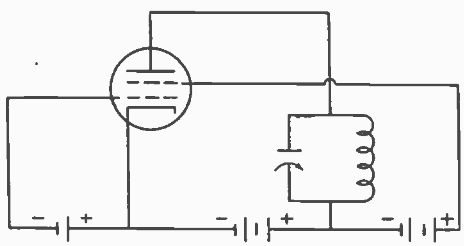


FIG. 13–17. Negative resistance produced by operating screen at higher potential than plate causing secondary emission.

of a screen-grid tube will have oscillations in it if the plate voltage is lower than the screen-grid voltage. The slope of the e_p-i_p characteristic in this region is negative, and, since the slope of this curve is a measure of r_p, r_p must be negative.

Oscillator efficiency. By adjusting the grid bias so that plate current flows only during a small part of the cycle, power $(I \times E)$ dissipated on the plate may be kept very low. Under these conditions the plate voltage may be increased beyond the value which would be safe if plate current flowed all the time. The circuit efficiency is high, but the alternating plate current will be non-sinusoidal, that is, it will contain many harmonics. Since, however, the tank circuit and the load coupled to the tank are tuned to the fundamental frequency the harmonics will not create much voltage in the load because the load will have a low impedance to frequencies higher than the fundamental. Such high-efficiency circuits may, therefore, be employed at radio frequencies, whereas they could not be used at audio frequencies. They perform like self-excited class C amplifiers.

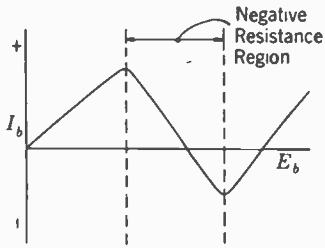


FIG. 14–17. Tube characteristic employed in circuit of Fig. 13–17.

Frequency stability. Complications set in when oscillators are used for transmission or for laboratory measurements where a constant frequency output is desired. The frequency of such circuits is determined chiefly by the inductance and the capacitance in them. The tuning condenser is shunted by the tube capacitances. The grid-cathode capacitance, for example, is a function of the plate load and the grid-plate capacitance of the tube. In fact the input capacitance C_i is a function of the load, and this input capacitance is

$$C_i = C_{gk} + C_{gp}\left(\frac{\mu R_L}{R_L + r_p} + 1\right)$$

which shows that any change in the plate resistance r_p of the tube, or in the grid-plate capacitance of the tube, or the output load R_L produces a change in the grid-cathode capacitance which may have a share in determining the frequency to which the system oscillates. Changes in filament temperature, in C bias, or in plate voltage will affect r_p and change its relation to the load resistance. Such changes produce a change in frequency of the oscillator output.

One way to lessen this difficulty is to shunt a fairly large capacitance directly across the grid and cathode to increase the total effective input capacitance so that small changes in the internal capacitance of the tube will have little effect upon the tuning. This is accomplished by using a small coil and a large condenser (low L/C ratio). High-capacitance circuits, however, have large circulating currents in them and may be quite inefficient.

Crystal oscillators. For many reasons it is desirable to keep an oscillator on the desired frequency. For example, the oscillator of a broadcast station determines the carrier frequency of

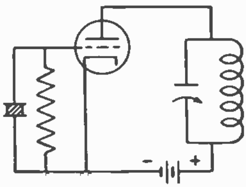


FIG. 15–17. Quartz crystal acting as a tuned circuit of very high Q and high stability.

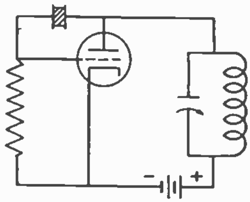


FIG. 16–17. Circuit in which crystal is connected from grid to plate.

that station, which should not vary. Various means are available for maintaining the frequency of an oscillator constant. By far the greatest frequency stability is obtained by means of a slab of quartz crystal, of correct dimensions, cut properly with reference to its optical and electrical axes.

Such a slab of quartz has the peculiar and highly important property called **piezoelectricity.** If the two surfaces of the plate are compressed mechanically, an electrical voltage appears across these faces. Conversely, when the voltage across the faces is changed, the crystal tends to change its size. It is, in effect, a mechanical resonator or a mechanical tuned circuit, and when inserted into a vacuum-tube oscillator it will serve the same purpose as a tank circuit made up of an inductance and a condenser.

Such a tuned circuit has an extremely high value of Q, acting in fact like a very large inductance of low resistance shunted by a very small capacitance. In Fig. 15–17 such a crystal is represented as being in the grid circuit of a tuned-grid tuned-plate oscillator. When the plate circuit is tuned a bit higher in frequency than the actual resonant frequency of the crystal, the entire circuit oscillates at the crystal frequency. For several degrees of variation of the tuning condenser, as long as the resonant frequency of the plate tank is higher in frequency than the crystal, the oscillator will still produce power at a frequency controlled by the crystal. This means that minor variations in the capacitance of the circuit, such as are occasioned by changing tubes, variations in plate voltage, etc., have no controlling effect upon the frequency of output.

Crystal cuts. There are many ways in which the crystal plates can be sawed with reference to the axes of the mother crystal. These are known as "cuts," and some of them are good for one purpose and some for other purposes. Since the frequency at which the quartz crystal resonates depends to some extent upon the temperature of the crystal, much research has gone into determining the cuts which will produce the smallest temperature coefficient, that is, cuts which will oscillate at the desired frequency regardless of temperature.

The frequency of some cuts is determined primarily by the thickness and that of other cuts by the breadth or length of the plate. In other cuts the temperature coefficient is controlled by one dimension while the frequency is determined by another; it is thus possible to control these two important factors independently.

In those cuts in which the thickness determines the frequency, thinner crystals oscillate at higher frequencies, the relation being 1960 kc per mm, for example, in the AT cut. For this reason the higher frequencies require very thin crystals, and, since these plates vibrate when in service with an amplitude dependent upon the amount of power they are controlling, oscillators working at high frequencies are not controlled directly by the crystal because of danger of crystal breakage. Instead, a thicker crystal is utilized to control an oscillator of lower frequency than that desired. This frequency is doubled or tripled or multiplied still higher in subsequent circuits. If the power output of the station is high, not all this power is directly controlled by the crystal. Instead, a low-powered oscillator is controlled by the quartz, the output of the oscillator being raised in frequency and amplified as much as desired in subsequent stages.

It is generally considered wise to keep the r-f current flowing through the crystal circuit lower than 100 ma, not to attempt to control more than 10 to 15 watts directly with the crystal, and not to control directly a circuit whose output is higher in frequency than about 15 megacycles. With care, a crystal-controlled oscillator should not vary in frequency more than a few parts per million over an appreciable time. That is, an oscillator having an output frequency of 1,000,000 cycles (1000 kc) should not vary more than 10 to 15 cycles from this frequency. With great care stabilities as high as 1 part in 10^8 may be obtained.

CHAPTER 18

TRANSMITTERS

A radio transmitter is a combination of an oscillator, voltage and power amplifiers, modulating equipment, and power supply apparatus all designed to put r-f power into an antenna from which it is radiated in the form of electromagnetic energy. The oscillator stage is of low power, perhaps 5 watts or so, is crystal controlled, and in high-frequency stations operates at some submultiple of the final radiated power frequency. It is followed by amplifiers which build up the voltage output of the crystal stage to a point high enough to drive the final power output stage. These intermediate amplifiers also act as "buffers" to isolate the crystal from the power stage so that the crystal can operate without any reaction back upon it by the power stages. These amplifiers may also act as frequency amplifiers, increasing the frequency by some integral multiple of the oscillator frequency. The frequency multiplication may take place in several stages. Connected to one of the stages is a modulating system which gets its excitation from a microphone or a telegraph key.

Types of r-f power amplifiers. When one generates and transmits r-f power, efficiency becomes important since the quantity of power that is purchased from a public utility, or generated locally by the station, must be paid for according to the amount used and not according to the amount usefully radiated. A 50-kw broadcast station that consumes 240 kw from the power line is certainly less efficient and will have a higher annual bill for electrical energy than one which delivers the same power to the antenna but which needs to draw but 105 kw from the power lines.

Since the output of a r-f power amplifier feeds a tuned antenna, it is easy to provide filtering in the output stage so that harmonics of the transmitter frequency do not get on the air. For this reason it is not necessary to use class A amplifiers in a radio transmitter. Much higher efficiency can be secured from class C amplifiers, which, it is true, produce many harmonics of the fundamental radio frequency, but these harmonics can be prevented from reaching the antenna.

If the modulation takes place in a low-power stage, then all subsequent stages must be linear in order that the modulation (which is the important message-bearing portion of the transmitted wave) will not be distorted. If modulation takes place in the final stage, the modulated stage can be operated as class C. If the transmitter is for transmitting code rather than voice or music, all stages may be operated as class C.

The principal difference between class B and C amplifiers is in the amount of C bias used and the portion of the a-c cycle during which plate current is permitted to flow.

In a class B amplifier the bias is so adjusted that the plate current is almost zero when the grid excitation is removed. If sine-wave voltage is applied to the grid, plate current consists of a series of half sine waves similar to the output of a half-wave rectifier. The load impedance is adjusted so that the relation between plate current and grid voltage is linear. The grid swings positive on excitation peaks so that grid current flows. During peaks of plate current the plate voltage reaches a minimum value because of the high voltage drop across the load. The tube power, then, is the product of a high current and a *low* voltage; but the load power is the product of a high current and a *high* voltage. This means simply that, of the total power taken from the plate battery, most is supplied to the load, and least is used up in heating the plate of the tube. Power does not flow continuously from the plate supply battery. Efficiency increases so that a maximum value of 78.54 per cent is attained.

Two tubes may be operated in a class B r-f amplifier if desired, and twice the power output of a single tube will be secured. Tubes in class B r-f amplifiers can be operated to produce full power output, whereas in a class A amplifier (audio) the maximum power is never attained.

A class C amplifier has its bias adjusted so that plate current is definitely zero for a certain portion of the excitation cycle. The grid excitation is high and may produce saturation current from the cathode of the tube on the positive portions of the excitation cycle. Other conditions are quite similar to those of a class B amplifier. The plate efficiency of the class C amplifier is high, approaching 90 per cent in large tubes.

In code transmitters, all stages may be operated class C. In a telephone transmitter only the stage in which modulation takes place and preceding stages are operated class C.

Neutralization. Buffer or intermediate power amplifiers may be triodes or, better, screen-grid tubes and pentodes. Triodes

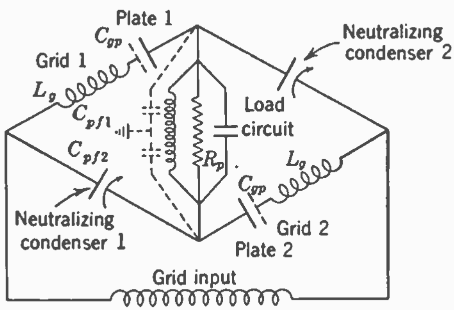


FIG. 1–18. Bridge circuit made up of neutralizing and tube grid-plate capacitances in a neutralized push-pull amplifier.

must be neutralized to keep them from oscillating or to keep variations in their output from feeding back to the crystal stage. Neutralization consists in making the tube into a bridge circuit in which an output voltage, equal to that fed back to the input by the grid-plate capacitance, is fed to the input. This additional voltage is of such a direction that it opposes (neutralizes) the effect of grid-plate capacitance voltage.

The several types of neutralizing circuits may be roughly broken down into three general schemes: (1) plate neutralizing, (2) plate-to-grid neutralization (as used in RCA and Western Electric transmitters), and (3) push-pull cross neutralization. Diagrams illustrating these methods are shown in Fig. 5–16.

The question naturally arises why triodes are employed nowadays when screen-grid tubes are available. The reasons are as follows: (1) plate modulated stages must use triodes since varying the plate voltage of a screen-grid tube has little effect upon plate current; (2) high-power screen-grid tubes have not become available; (3) at the high frequencies (1–2 meters wavelength) the input and output capacitances of screen-grid tubes are so high that simpler tubes of the triode type must be used.

Power amplifier design. Since the grids of class B and C amplifiers are driven positive, the grid circuits pass current and therefore require power. This power must be provided by the stage that precedes and drives the class B or C stage.

Actual design of a class C stage is a matter both of calculation and of cut and try. Since the plate current does not flow continuously, the internal resistance of the tube varies over the cycle and a numerical value which can be employed in calculating the load resistance, power output, etc., cannot be supplied by the manufacturer (or measured by the user). Several steps must be taken in the preliminary calculations.

1. The maximum grid excitation voltage and minimum plate voltage must be decided upon. Ordinarily it is desired that the excitation power be kept small. This means that excessive grid current should not flow and, therefore, that the maximum excitation voltage should not be too high. The maximum or peak value of the grid exciting voltage may be taken as 80 per cent of the minimum instantaneous plate voltage.

2. The proportion of the excitation cycle in which plate current is to flow must be decided upon. If current flows for only a small portion of the cycle, high efficiency will be secured but the power output will be low. If the current flows for a longer period, efficiency will be low but the output will be high. The actual time during which current flows is expressed as an angle as shown in Fig. 2–18. If current flows for the entire half cycle in which the grid voltage is increasing in a positive direction, an entire half cycle (180°) of current will flow. Ordinarily the angle of flow will be of the order of 120° in a voice-modulated stage.

3. Determine the cut-off bias. This will be equal to

$$E_c = E_{bb} \div \mu$$

4. Determine the required excitation voltage. For a class C amplifier this will be somewhere between 1.5 and 4 times the cut-off bias.

5. Calculate plate current, grid current, plate dissipation, efficiency, etc. If the values as calculated are not satisfactory,

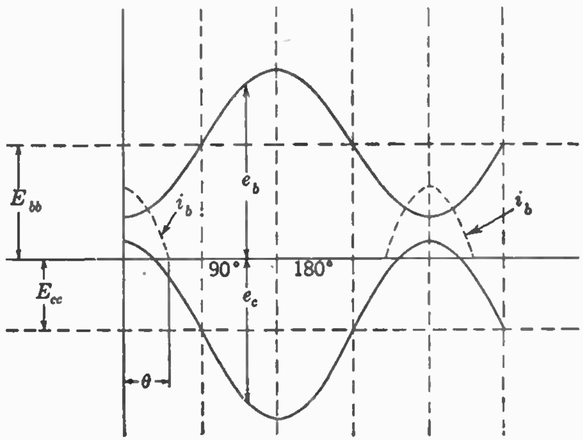


Fig. 2–18. Relations existing in a class C amplifier. θ is the angle during which plate current flows.

other assumed values of angle of flow, maximum grid voltage, and minimum plate voltage must be used and new calculations made.

Power relations in class C amplifier. The power supplied by the plate battery E_{bb} is

$$P_{\text{input}} = E_{bb} I_b$$

The output to the tank circuit is

$$P_{\text{tank}} = \frac{E_0 I_p}{2}$$

where E_0 is the peak voltage across the tank.
I_p is the peak a-c plate current.

The peak voltage across the tank is

$$E_0 = E_{bb} - E_{p\ \text{min.}}$$

Thus the peak tank voltage is almost equal to the plate battery voltage. The power lost in the tube itself is the difference between the total power supplied by the battery and that supplied to the tank. The efficiency is the ratio of the power supplied to the tank to the power drawn from the power supply.

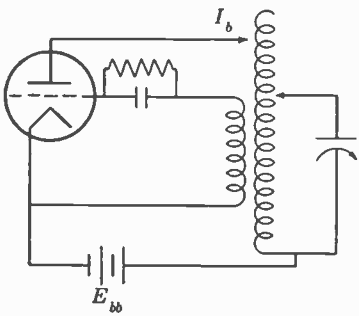


FIG. 3–18. Use of taps on plate tank inductance for impedance adjustments between tube and load.

Since the load of the amplifier is tuned to resonance at the frequency which is to be transmitted, this load represents almost a pure resistance to the tube with a value equal to L/CR, where L, C, and R are the constants of the tank circuit. This value of load should be equal to the effective impedance of the tube for greatest power transfer. The load resistance is also equal to E_0/I_p by Ohm's law, and so the value of the load resistance represented by the anti-resonant circuit is $R_b = L/CR$, where R is the effective a-c resistance of the tank coil including the resistance coupled from the load.

If the value obtained from the above expression does not give maximum power into the tank circuit, the tank inductance may be used as an auto-transformer to adjust the circuit for maximum power output to the tank. This really changes the ratio of L to C without changing their product (LC), which determines the resonant frequency. If the tank circuit is tuned to a different frequency, L, C, and R will change, but the ratio remains essentially constant.

Use of tank for flywheel effect. The student may wonder why tank current can flow continuously if the tube passes current only at periodic intervals, and why the tank current may be sinusoidal if the plate current is in the form of pulses.

It must be remembered that the only power dissipated in the tank is in the inherent resistance of that circuit, or the resistance reflected into it by the process of supplying radiation power to an antenna. Aside from this loss of power, energy put into the tuned circuit is transferred back and forth from condenser to inductance, just as, in a swinging pendulum, energy is transformed from energy of position (potential) to energy of motion (kinetic). To maintain current flowing in the tuned circuit, all one needs to do is to supply the power losses, not necessarily continuously but at intervals. The pendulum of a clock keeps swinging although it gets an impulse from the source of energy, the spring, only at the proper points in its cycle, by the escapement mechanism, and similarly the tube acts as an escapement mechanism in the class C amplifier, supplying energy to the circuit at the proper points in the cycle and of the proper amounts to overcome the losses inherent in the circuit plus energy transferred to the antenna for radiation. If the tuned circuit has a high Q, 10 or more, the currents in it will be sine-wave.

Frequency multipliers. Since the amount of power that can be controlled by a quartz plate is limited because the plate vibrates when it is in an electrical circuit tuned to the frequency of the plate, and since the thickness of the plate and its mechanical strength decrease as the frequency of resonance increases, it follows that high-power high-frequency stations cannot be controlled directly by the quartz crystal. Instead the r-f energy is generated at low power and low frequency, and both power and frequency are increased in subsequent stages.

A frequency multiplier is an amplifier so biased that it generates high-amplitude harmonic voltages in its output; it is a class C amplifier. Thus a crystal may control the frequency of a 5- to 50-watt oscillator working at, say, 1 megacycle. If the voltage output of this oscillator is used to drive a voltage amplifier which has a high second harmonic, 2-megacycle voltages will be available in the amplifier output. If the third or higher harmonics are high, 3-megacycle or even higher-frequency voltages can be obtained in the amplifier output. This multiplication of frequency may take place in several stages. For example, if the antenna is to radiate energy at 45 megacycles, a 5-megacycle crystal may be employed. The first multiplier may increase this to 15 megacycles, and the next "tripler" may increase it to 45 megacycles.

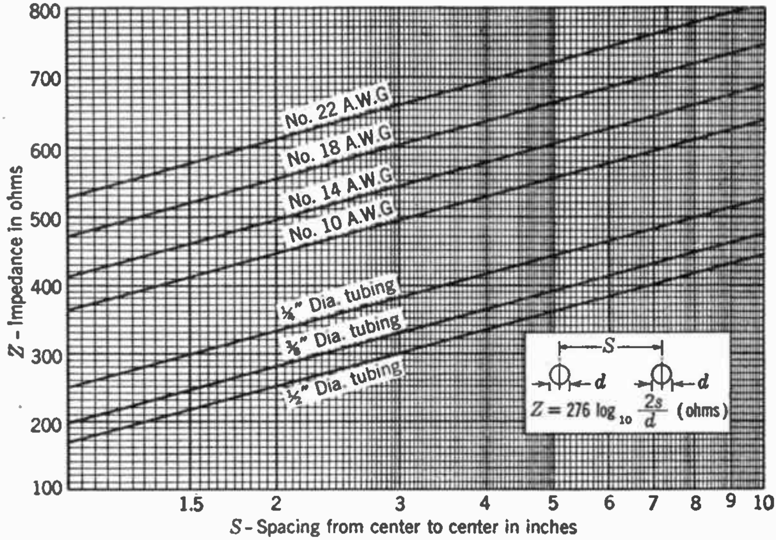
The plate tank circuit is tuned to the harmonic. This makes a frequency multiplier less efficient than a class C amplifier because the fundamental component of the plate current is wasted at the plate, since the impedance of the tank circuit is low at the fundamental and is reactive rather than resistive.

Any tube or circuit which has a non-linear relation between grid voltage and plate current will have harmonics in its output. Overbias of the grid is one method of securing a high harmonic content. Special circuits have been devised for the purpose of frequency multiplication.

Coupling the transmitter to the antenna. As an antenna in free space operates much more efficiently than one surrounded by buildings, vegetation, etc., it is common practice to locate the antenna at some distance from the transmitter. This necessitates provision for transferring the power in the final stage of the transmitter to the antenna. Some sort of transmission line is the usual means employed.

A very simple line is made up of two parallel conductors, each about ⅛ in. in diameter, the spacing being 12 to 15 in., arranged horizontally equidistant from the ground. Another type of transmission line consists of two conductors, one inside the other and separated from it by spacing insulators. This concentric line has the great advantage that the outer conductor can be grounded or actually placed underground. It is common practice to maintain a slight gas pressure in the concentric line to keep out moisture. Each of these lines has a characteristic impedance, more or less independent of its length, depending upon the diameters of the individual conductors and their distances from each other. In Fig. 4–18 will be seen the relation between these factors.

Since the greatest amount of power will be transferred to the line from the transmitter when the impedances of the line and transmitter are properly matched, and maximum power will in



Impedance of two-wire transmission line.

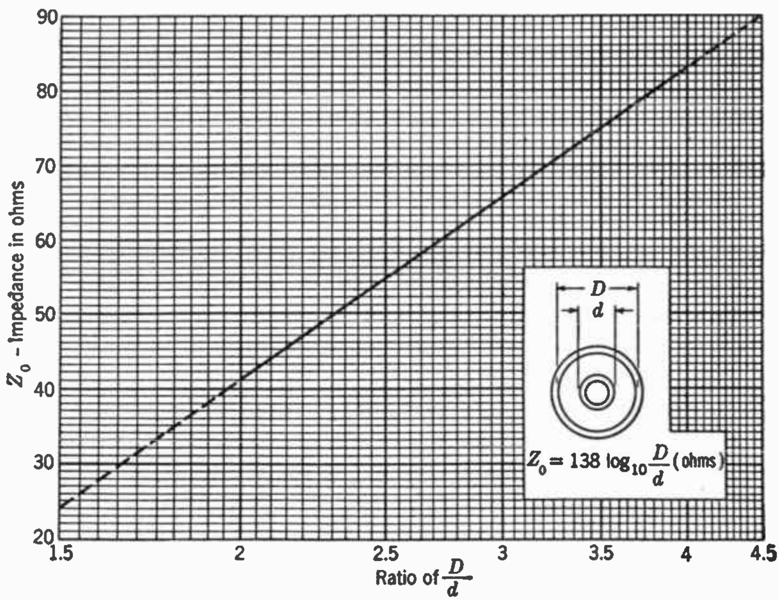


FIG. 4-18. Impedance of concentric line.

turn be transferred to the antenna from the line when these two impedances are matched, it is a matter of great moment to the transmitter engineer to design properly the transmission lines, the antenna, and the coupling systems for connecting them together. From the standpoint of the line, it is highly desirable that it operate between two pure resistances whose values are

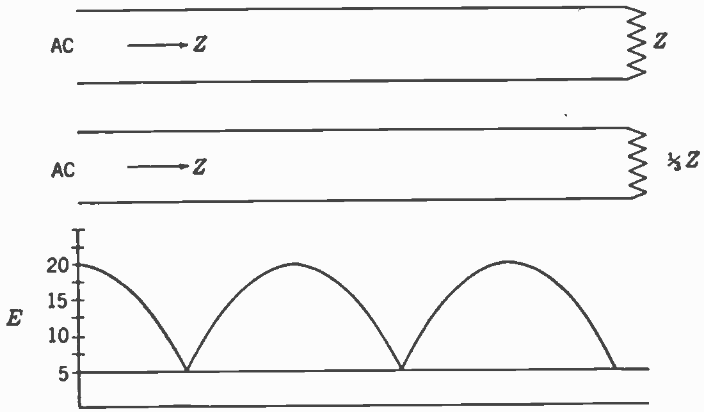


FIG. 5–18. If the line is terminated in a load equal to the line impedance, maximum power will be absorbed by the load. But if the load has a lower or higher impedance, standing waves of current and voltage will occur along the line caused by power flowing back from the load toward the source.

equal to the impedance of the line. If there is some reactance at either end, "standing waves" will be set up along the conductors and not only will they tend to radiate energy into space, acting like antennas, but also the current flow will increase, giving rise to greater power loss (I^2R) in the line's inherent resistance. Standing waves are periodic positions of maximum and minimum currents (and voltages) and are evidence that the line is resonant and therefore acting like a tuned circuit.

If the line is properly terminated, it will transfer r-f power from transmitter to antenna with very slight loss, which is made up entirely of the I^2R power loss in heat. Naturally, the higher the voltage that can be put on the line, the lower the current.

and the lower the loss in power. Further information on transmission lines will be found in the chapter on antennas.

Keying a code transmitter. If the radio station is to transmit code only, its modulation system can be very simple. One merely inserts a key in some portion of the power supply system so that when the key is up the oscillator does not generate r-f

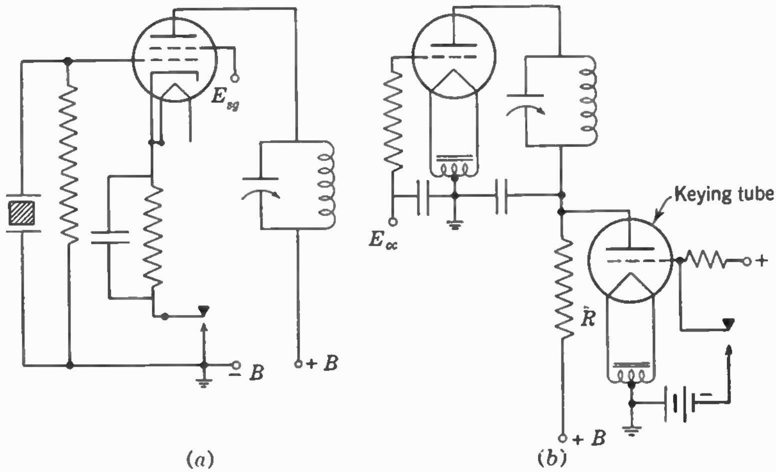


FIG. 6–18. Methods of keying an oscillator or an amplifier. At (a) the plate current is keyed directly; at (b) a keying tube is employed. Voltages across the key in (b) are much lower than in (a).

power, or the amplifier does not operate, thus preventing any of the oscillator output from reaching the antenna. Keying is usually accomplished in a buffer stage between the crystal oscillator and the final power amplifier stage. The grid of this keying stage may be overbiased so that no plate current is drawn until the key is pressed when the bias on the tube is reduced. If the power taken by this tube is not great, keying may be accomplished in the power supply lead to the cathode. The tube side of the key is "hot," and it is dangerous to operate a hand key in this position. It is better to use a mechanical relay here, which, in turn, is operated by the hand key. In Fig. 6–18 (b) a keying tube is shown. When the key is down, the keying tube draws no current since it is biased to cut-off. When the keying tube is up, the grid is somewhat positive, drawing a high current which produces a high voltage drop across R and reduces the voltage to the power tube.

Audio portion of broadcast transmitter. Any station which transmits speech or music must necessarily be more complex than one which transmits code. Between the microphone and the final power stage feeding the antenna must be high-quality audio amplifiers having sufficient voltage amplification and power output to raise the output of the microphone to the level required to modulate the r-f portion of the transmitter. All these amplifiers and the modulating equipment must be electrically quiet in operation. The modulator must not introduce noise, hum, or distortion into the signals that come from the microphone.

Audio-frequency range. The wider the a-f range to be transmitted, the more difficult is the problem of keeping noise and distortion out of the system. A voice-frequency station serving a police or fire department need not have as wide a frequency range as a high-fidelity broadcast station. Speech can be transmitted with good intelligibility if a frequency range of 250 to 2500 cycles is transmitted. If the higher frequencies are transmitted, the intelligibility will be better. If the lower frequencies are transmitted, the sounds will be louder and the sender will not have to talk so loud. Such a frequency range would be entirely inadequate if an accurate reproduction of a musical program were to be transmitted.

It is recognized that perfect reproduction of speech and music requires the transmission, without volume or frequency distortion, of a frequency band from 30 to 15,000 cycles. The broadcast audience has never had reproduction of this quality. Transmitters are placed in the standard broadcast band so close together, in order to accommodate all of them, that frequencies higher than 5000 cycles conflict with stations on adjacent channels. Manufacturers of receivers have spent the energies and capabilities of their engineering departments in reducing the cost of receivers for the sake of mass production of low-priced units instead of in improving the tone fidelity of reproduction. So much distortion and noise are inherent in the average broadcast receiver that, if the frequency response band were widened out to reproduce what is available from the average broadcast transmitter, the receiver would be unsalable. The average receiver has very little response at 5000 cycles or at 100 cycles.

Nevertheless broadcast stations transmit wide bands of audio tones, free from distortion and noise. Their output is vastly beyond the capabilities of the average broadcast receiver to reproduce with fidelity.

The advent of frequency-modulation broadcasting, however, can provide the public with wide-band noise-free reproduction, if the receivers are properly designed, so that the standard broadcast receivers will either be improved to compete with frequency-modulation receivers or will come to be recognized for what they are—devices of limited tone response full of distortion products which make the reproduced sounds quite unlike those picked up by the studio microphones.

Microphones. The input to the whole broadcast station is the microphone. If it has limited tone response, or if it is noisy, the entire transmission system will be faulty. All microphones have but one purpose—to translate a mechanical motion into an electrical voltage. Sound waves are variations of air pressure; they produce mechanical changes in the microphone which produce electrical changes that are representative of the sounds affecting the microphone.

The earliest microphones, used at present in the common telephone, contained carbon grains between two metallic plates across which was placed a voltage. Sounds impinging upon the diaphragm of the microphone caused the carbon grains to pack closer together or to be released from pressure, and these variations in pressure produced variations in resistance of the carbon. The current through the microphone changed in response to sounds in the vicinity of the microphone. The electric current through the device was "modulated" by the sound waves, and the variations of current were characteristic of the air vibrations or pressure variations.

The next microphone was made up of two plates of a condenser, one plate being movable with respect to the other. When sound waves hit the variable plate, acting as the diaphragm, the spacing between the two plates changed and the voltage across the condenser changed in response to the changes in capacitance produced by the varying spacing between the two electrodes.

A more modern microphone is a dynamic loud speaker in reverse. When a coil placed in a steady magnetic field moves in response to sound waves impinging upon the diaphragm attached to it, potentials are produced across the coil. These varying voltages may be amplified and finally used to modulate the broadcast transmitter. The ribbon microphone has a light metallic ribbon suspended in a magnetic field. The crystal microphone has a piezoelectric crystal of Rochelle salt. When sound waves strike the diaphragm of this device, mechanical pressures are exerted across the two faces of the crystal, and corresponding voltages are produced across them.

No matter what kind of microphone is employed, the voltages produced may be amplified and used to modulate the radio transmitter. The voltages are quite low, in general being about 60 to 80 db below 1 volt when the microphone is picking up sound of average intensity, such as a normal speaking voice. This is a very small fraction of a volt. Considerable amplification is needed to bring the microphone output up to the level at which it must be placed upon studio or telephone lines to be sufficiently greater than the line noise not to be masked by it.

Pre-amplifiers. The first amplifier which the microphone output goes through is called a pre-amplifier or preliminary amplifier; it generally has a voltage gain of approximately 50 db.

Additional amplifiers are needed after the signals come in from the line if the pickup is at a distant point, and often amplifiers are used for the sole purpose of isolating the line from whatever apparatus the isolating amplifier feeds into. Since these amplifiers produce voltage gain, losses are introduced deliberately so that the level out of the amplifier is not greater than desired.

Volume range and compression. A violin playing very softly has an output of about 4 mw, and a full orchestra at its peak has one of about 70 watts. This is a volume range of 43 db. Before the weakest sounds picked up by the microphone can be placed upon lines or upon the transmitter they must be raised in level above any possible line or transmitter noise. Before the loudest sounds picked up are utilized they must be reduced in level to the point where they will not cause "cross talk" from one line to another or will not overload the transmitter. Thus the original 43-db volume range must be compressed, usually to a volume range of about 25 to 30 db, either manually, by operators who control the gain of some amplifier so that the output falls within these limits set by the system, or automatically.

If the average level of modulation is raised, listeners will get louder signals from a given broadcast station. If peak sounds are not to overload the transmitter, however, the average modulation must be kept quite low, perhaps 30 per cent. Some amplifiers now employed act as volume compressors, limiting the peak voltages automatically by reducing the amplifier gain and then allowing it slowly to come back to normal after the peaks have passed. In this manner overloading of the transmitter is avoided and at the same time the average modulation level can be raised.

Modulation. If the r-f power radiated from the antenna is to convey intelligence to the distant receiver, some means must be provided for modulating the power with the intelligence. Simply keying the power on and off in accordance with a prearranged code is a means of so modulating r-f power. In this system the amplitude of the transmitting antenna current is varied completely, from zero to maximum value.

If voice or music is to be used to modulate the transmitter, some means must be provided for varying the r-f power in the transmitter antenna according to both the strength and frequency (tone) of the voice or music. The power is varied in two general ways at the present time. In the older system used in the standard broadcast band, the **amplitude** of the transmitter antenna current is varied in accordance with the strength of the microphone voltages. Thus a loud audible sound produces a correspondingly great change in the antenna current. The rate at which the antenna current (not its amplitude) is varied depends upon the frequency of the audio sound. This is **amplitude modulation.**

In newer systems of police radio and in the frequency-modulation broadcasting band, the **frequency** of the transmitter is varied in accordance with the audio tones. When the microphone is quiet, no modulation, the station sends out a fixed-frequency carrier wave just like an amplitude-modulated station. When the microphone picks up a sound, its voltage output is used to vary the frequency of the transmitter, the rate (that is, the number of times per second) at which it is varied depending upon the frequency of the audio tone and the extent of the variation being a measure of the microphone input.

Low-level vs. high-level modulation. Amplitude modulation can occur practically anywhere along the line-up of transmitter tubes from the oscillator to the final output stage. If it is performed early, say just after the crystal, modulation is said to take place at a low level. All succeeding stages must amplify this modulated r-f voltage. If modulation takes place in the final stage where the power level is already high, it is said to take place at a high level. Each method has its advantages.

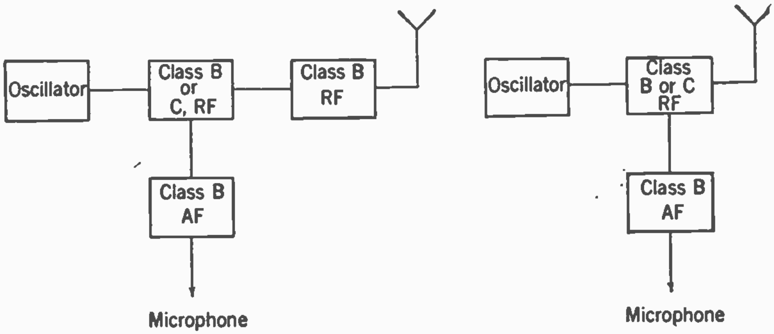


FIG. 7–18. Modulation can take place anywhere between oscillator and antenna. At left is a low-level system; at right, a high-level arrangement.

Power is required to modulate power, and power is represented in the audio variations which modulate a r-f carrier. If modulation takes place at a low level, each of the succeeding amplifiers must handle both the carrier power and the modulating power and must be linear, low-efficiency amplifiers. The amount of power required to modulate a low-level stage, however, is small; in fact, all that is required is a good audio power amplifier. If modulation is produced in the final stage, the amount of modulation power must be greatly increased, but this disadvantage is partly overcome by the fact that the modulated and preceding stages will be more efficient (class C) than any stage following modulation (class B).

Modulation systems. In general, modulation can take place in the plate circuit, control grid, suppressor grid, or cathode circuit of an oscillator or r-f amplifier. What is required is that the r-f oscillatory current in the output of the amplifier or oscillator be proportional to the voltage of the tube electrode being modulated. For example, if the output current of a r-f amplifier is proportional to the plate voltage, then, when the plate voltage is varied in accordance with microphone voltages, the output current will be so modulated in amplitude, and the antenna current will vary accordingly. Voltages picked up at the distant receiving station will bear these modulations, which can be utilized by the detector.

Heising modulation. An early system (named for its inventor), widely used until more efficient methods were developed, involved varying the plate voltage of an oscillator in accordance with the modulating voltage. The output of the oscillator was then amplified and fed to an antenna.

Consider Fig. 8–18, in which the reactance of the choke L is high at all audio frequencies compared with the impedances of the two resistances. The resistance R_{mod} is the resistance of the modulator tube which is simply an a-f power amplifier. The

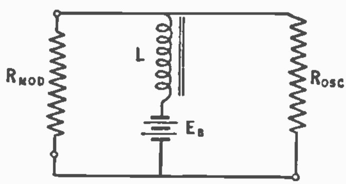


FIG. 8–18. Equivalent of Heising modulation system.

resistance R_{osc} is the resistance of the oscillator tube. Suppose that R_{mod} is caused to vary at some audio rate. The current taken from the plate battery will not vary at this rate because of the large choking effect of the inductance L. The total current, then, from the plate battery is constant. If R_{mod} increases, less current will be taken by it and more can be taken by the oscillator. If R_{mod} decreases, more current will be taken by it and less will be available for R_{osc}.

If the variations in R_{mod} vary at 1000 cycles per second, the current taken by the oscillator will vary at 1000 cycles per sec-

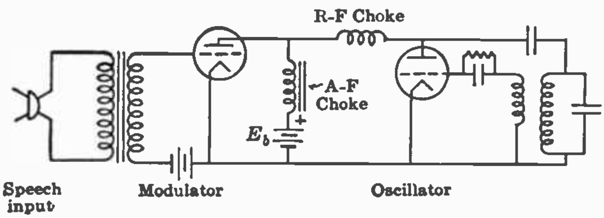


FIG. 9–18. Complete Heising modulation system.

ond. This is another way of saying that the plate current of the oscillator is modulated at a 1000-cycle rate. The modulator acts as a class A amplifier, and the theory of such amplifiers is adapted to this type of modulation. To modulate the oscillator completely, the modulator plate current must be reduced to zero at one instant in the audio cycle and at the next half cycle the plate current must be doubled. Now in a class A amplifier it is impossible to reduce the plate current to zero without causing severe distortion. Therefore a tube having the same power rating as the oscillator will not completely modulate the oscillator in this circuit. If, however, the modulator is supplied with a higher voltage than the oscillator, so that it operates with a higher zero-signal plate current, or if the choke, L, is replaced by a step-up transformer, complete modulation may be accomplished.

One trouble with modulating the oscillator is that the frequency generated by the oscillator is not independent of the plate voltage. Since the plate voltage to the oscillator varies from zero to twice the average plate voltage, the frequency of the oscillator varies over the audio cycle. This produces frequency modulation. For this reason it is no longer the practice to modulate the oscillator, except in the laboratory, or for other

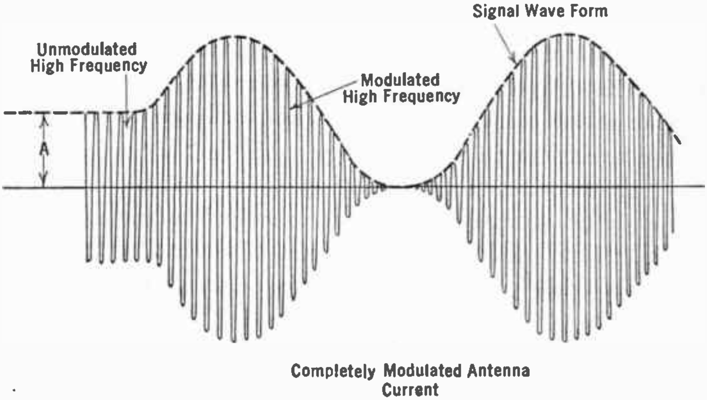


FIG. 10–18. A r-f wave before and after complete modulation.

uses where the variation in carrier frequency is not a disadvantage.

Class C modulator. It is standard practice now to modulate an amplifier stage rather than to modulate the oscillator. Since the oscillator is crystal controlled and since it is highly desirable that it be permitted to operate as isolated from the rest of the circuit as possible, one or more stages of isolation amplifiers (buffer amplifiers) are placed between the crystal and the final power stage. Modulation can take place in any of these amplifiers, for example in a highly efficient class C stage, so biased and otherwise arranged that its control grid, plate, cathode, or suppressor grid may have its voltage varied in accordance with the audio modulations with the certainty that the output will be a good replica of these modulating voltages.

In the plate-modulated class C amplifier, r-f voltages of constant frequency and amplitude are fed to the grid circuit, producing in the plate circuit a-c power of the frequency of the exciting voltage. In the absence of modulation this tube acts merely as a powerful r-f amplifier, feeding its output to an antenna where the r-f power is radiated at a constant frequency and unvarying amplitude. When the microphone in the studio picks up sound waves and converts them into voltages, and when they are amplified sufficiently, they may be added to the plate voltage (E_{bb}) of the class C amplifier, at some instants adding to the voltage, and at other instants subtracting from it. Current and power transferred to the antenna vary in accordance with the variations in plate voltage. All that is necessary is to arrange the circuit and voltages properly so that the envelope of the modulated antenna currents is exactly like the microphone currents. This condition will be met if the r-f output of the tube varies linearly with respect to the plate supply voltage.

If the class C tube has 1000 volts supplied to it from the plate voltage supply system, then 1000 peak volts of audio must be supplied to modulate it completely. At these instants of 100 per cent modulation, the peak power output of the class C tube is 4 times its unmodulated output. Part of this additional power is supplied by the modulator, and the remainder comes from the plate voltage supply system. To modulate the amplifier completely, the modulator must supply 50 per cent ($E_bI_b/2$) of the d-c plate power (E_bI_b) that is required under non-modulated conditions. The modulator, therefore, must be capable of considerable power output. It is usually a highly efficient class B audio amplifier, or one of the newer forms of amplifiers developed within recent years.

If the class C tube is modulated by varying its grid bias (grid modulation), much less modulation power is required, but there will be less power output from a tube of given power dissipation.

Analysis of modulation. In the antenna of a radio station flow alternating currents of the frequency of the carrier of that station. As long as the transmitter is not modulated, the amplitude of these alternating currents is constant. When, however, the microphone is spoken into or the station is otherwise modulated, the plate voltage on the modulated amplifier changes, the currents in the plate circuit and in the tank circuit change, and the currents in the antenna change. This is the essence of modulation—a variation of the carrier current by the microphone voltages. The antenna current increases and decreases in accordance with the frequency of the modulating tones.

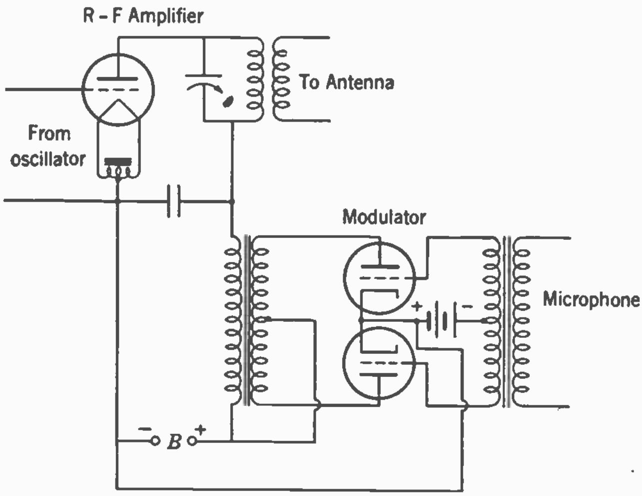


Fig. 11–18. Plate-modulated r-f amplifier.

The **modulation factor** is the ratio between the audio peak current and the radio peak current E_{AF}/E_{RF}. When they are equal the transmitter is modulated 100 per cent; the peak currents flowing in the tank and antenna circuits will have doubled; the peak power developed by the transmitter will have quadrupled; the average power in the antenna will have increased by 50 per cent. More power is actually radiated at these instants.

A modulated wave form like that in Fig. 10–18 is the equivalent of a single carrier frequency plus two "side-band" frequencies displaced from the carrier by the audio frequency. Thus if the carrier is 1000 kc, and if a pure 1000-cycle tone is impressed upon the modulator grid, the antenna will radiate the three frequencies, 1000 kc minus 1 kc, 1000 kc, and 1000 kc plus 1 kc, and a sensitive and selective detector will pick up these radiated frequencies.

If the station is a broadcast station accepting all frequencies up to 10,000 cycles, for example, the side bands of the station occupy the region between the carrier minus 10,000 cycles and the carrier plus 10,000 cycles. This is why the carriers of broadcast stations cannot be placed closer together than the highest

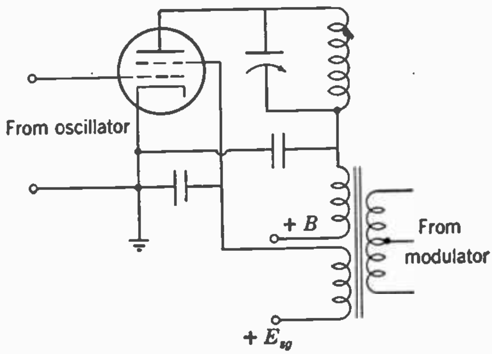


FIG. 12–18. Circuit in which both plate and screen voltages are modulated.

modulating audio frequency. Otherwise their side bands will overlap and create interference and distortion in the listeners' receivers.

Increase of antenna current with modulation. The antenna current of a transmitter will increase when modulation occurs because the power into the antenna increases by the amount of the modulation. Therefore the increase in antenna current can be taken as a measure of the amount of modulation in percentage. This additional power is represented by the side bands of modulation.

The average power in the carrier is $I^2R/2$. The amplitude of the current in each side band is the current in the carrier multiplied by $m/2$, where m is the modulation factor. Therefore the power in each side band is $m^2I^2R/8$ or $\frac{I^2R}{2} \times \left(\frac{m}{2}\right)^2$, and the total power for the carrier and side bands, on the basis that each side band is a replica of the other, is

$$\text{Total antenna power} = \frac{I^2R}{2}\left(1 + \frac{m^2}{2}\right)$$

Since the current is proportional to the square root of the power the current at any degree of modulation is proportional to

$$\sqrt{1 + \frac{m^2}{2}}$$

Thus the following table can be calculated.

Percentage Modulation	Percentage Increase in Antenna Current	Percentage Modulation	Percentage Increase in Antenna Current
0	0.0	50	6.0
10	0.22	60	8.7
20	1.00	70	11.7
25	1.6	80	15.0
30	2.2	90	18.7
40	3.9	100	22.6

Modulator design. The relation between modulator and modulated amplifier in a Heising system may be understood from the following analysis. When the amplifier is unmodulated, its apparent resistance is the plate voltage divided by the plate current, or the plate voltage squared divided by the watts taken by the tube. Thus

Total power from battery

$$= \text{Power to modulator} + \text{Power to amplifier}$$

$$= P_M + P_A$$

But

$$P_M = E_{bb}^2 \div R_M$$

and

$$P_A = E_{bb}^2 \div R_A$$

where E_{bb} = battery voltage.
R_M = apparent resistance of modulator.
R_A = apparent resistance of amplifier.

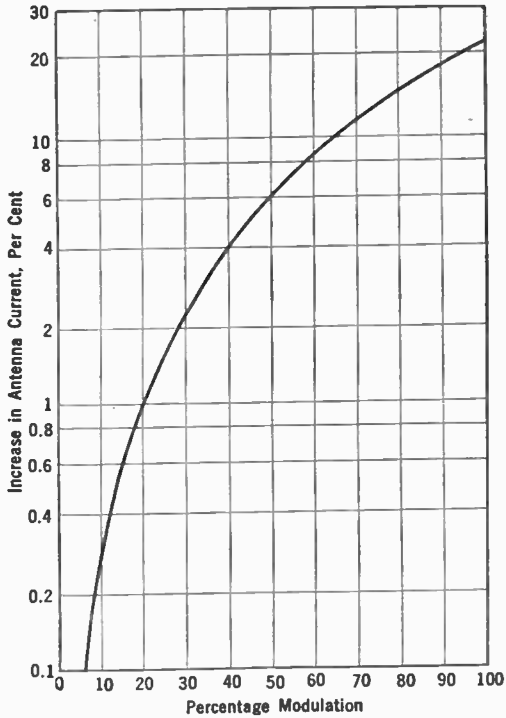


FIG. 13–18. Increase in antenna current as a function of the percentage modulation.

Now, if $P_M = P_A$,

$$P_M + P_A = \frac{E_{bb}^2}{R_M} + \frac{E_{bb}^2}{R_A} = \frac{2E_{bb}^2}{R_M} = \frac{2E_{bb}}{R_M} \times E_{bb}$$

Since m is the ratio between the peak a-c voltage delivered to the modulated amplifier and the battery voltage

$$m = \frac{\sqrt{2}e_p}{E_{bb}}$$

and

$$E_{bb} = \frac{\sqrt{2}e_p}{m}$$

e_p = a-c voltage developed by the modulator

$= \frac{1}{2}\mu e_g$

since $R_M = R_A$.

Therefore,

$$E_{bb} = \frac{\sqrt{2}\mu e_g}{2m}$$

$$P_M + P_A = \frac{2\sqrt{2}\mu e_g}{2mR_M} \times E = \frac{\sqrt{2}\mu e_g}{mR_M} \times E_{bb}$$

$$P_A = P_M + P_A - P_M = \frac{\sqrt{2}\mu e_g}{mR_M} E_{bb} - \frac{E_{bb}{}^2}{R_M}$$

where μ = amplification factor of modulator.
R_M = plate resistance of modulator.
m = desired percentage modulation.
e_g = rms signal to modulator grid.
E_{bb} = plate voltage applied to modulator and amplifier.

This, then, is the power that can be modulated to a given percentage by a modulator tube.

Example. Assume the following data on an 842 which is to modulate a tube from a 350-volt source of supply.

$E_{cc} = -88$ $\qquad e_g = 88/\sqrt{2} = 48$

$E_{bb} = 350$ $\qquad I_p = 14$ ma

$r_p = 2400$ $\qquad \mu = 3$

$$P_A = \frac{\sqrt{2} \times 3 \times 48 \times 350}{0.6 \times 2400} - \frac{350^2}{2400}$$

$$= 44.4 \text{ watts}$$

when $m = 0.6$ or 60 per cent.

The load of the modulator, which is the apparent resistance of the amplifier, is given by

$$R_A = \frac{E_{bb}{}^2}{P_A} = \frac{350^2}{44.4} = 2260 \text{ ohms}$$

The peak voltage produced by the modulator across the amplifier load is

$$\sqrt{2}e_p = mE_{bb} = 0.6 \times 350 = 210 \text{ volts}$$

CHAPTER 19

ANTENNAS AND ELECTROMAGNETIC RADIATION

Now that we have produced high-frequency energy, how are we going to make it convey intelligence to a distant radio receiver? What is the nature of the process called radiation? What is fading? How can the greatest amount of energy be radiated at the transmitter and the greatest amount picked up at the receiver?

Radiation. We are already familiar with the fact that a wire or a coil carrying a current will set up a magnetic field in the vicinity. This field can be detected and explored by means of a compass. If two coils are coupled together, and if the current in one coil is changed, a voltage appearing across the terminals of the other coil will change in unison with the changes in current in the first coil. Energy is stored in this electromagnetic field, and this energy can be put to work in various ways.

We are also familiar with the fact that a charged condenser has energy stored in it. This energy is said to reside in the dielectric of the condenser in the form of an electrostatic strain, like a stretched rubber band. If the ends of two wires connected to the two terminals are brought near each other, a spark will jump across, showing the presence of stored energy.

If a resonant circuit is excited by voltage of the resonance frequency, energy can be stored in the circuit, at one instant all of it residing in the magnetic field produced by the coil and at another instant being shifted to the electrostatic strain existing in the condenser. If the coil and condenser have no resistance, there will be no loss of energy in the circuit and no power will be consumed in exciting it.

Coils and condensers are concentrated pieces of equipment. They occupy small space. When we come to consider similar phenomena associated with antennas we are no longer dealing with small compact objects; we are dealing with structures having inductance and capacitance and resistance but spread out over large space. The electrical quantities L, C, and R are no longer concentrated; they are distributed more or less evenly throughout the whole length of the conductor. Nevertheless, energy can be stored in the fields of an antenna, and this energy can be detected at vastly greater distance than energy associated with a coil of wire—a concentrated inductance.

Radiation field. Let us again consider a compact coil of wire. We know that, when we connect it to a battery, lines of force move out from the wire and produce a field of force, and that when we disconnect the battery they fall back upon the coil. In a coil (ignoring all resistance) all the energy taken from the battery to set up the field is given back when the field collapses.

Now suppose that, after the field is established, we suddenly reverse the direction of current flow in the coil. Let us do this so fast that the lines of force do not have time to collapse back upon the coil and reverse their direction. What happens? The energy in the field cannot be given back to the battery; it cannot remain in the coil, for there is an oncoming new field set up by the battery. In effect, the original energy gets pushed off into space. A certain amount of energy has been set free. This is radiated energy.

Now it is a fact that it takes finite time for a current in a coil or in an antenna wire to set up the field at a distant point. Electric waves travel at the rate of 186,000 miles per second, and, although this may seem to be pretty fast, it is still a finite velocity. It takes time to establish an electromagnetic field about a coil or a wire, and, if we find means of reversing the current through the coil or wire at such a rate that the energy at a distant point cannot get back to the original source before a new batch of oncoming energy arrives, this bit of energy will be freed from its originating source.

This is what happens when we put high-frequency energy into an antenna. If the frequency is high enough, the current will reverse so fast that the energy existing at a distance will not have time to collapse back upon the antenna and will be free energy radiated off into space.

The fact that we can detect signals from a radio transmitter perhaps a thousand miles away should be no more strange than the fact that light from the sun pours in on us throughout many hours of the day. Heat, light, and radio radiation are all electromagnetic energy. They have exactly the same form, and the only differences are in the method of their generation, their wavelength, and methods of detecting them. Light is energy of very short wavelength, radiant heat is energy of longer wavelength, and radio waves are still longer.

In an actual antenna the direction of current does not change instantly but less quickly, depending upon the frequency (which is the rate at which the direction of current flow is changed). Some of the distant field does return to the antenna to be neutralized by the new field, but some becomes free. That portion of the energy which returns is called the **induction** field, and the detached portion is called the **radiation** field. The induction field dies out very quickly and cannot be found far from the antenna. The induction field varies inversely as the square of the distance from the antenna; the radiation field varies inversely as the first power of the distance.

Radiation, then, may be looked at as energy which is sent out from the antenna and which is prevented from getting back by production of new energy of opposite polarity.

Radiation resistance. Let us suppose that two wires about 30 ft long and about a foot apart are coupled to a 7-megacycle oscillator as in Fig. 1–19. The current into this double-wire system will rise to a maximum when the tuning condenser has been adjusted so that the product of L and C of the line is equal to the product of L and C of the oscillator. Suppose that this current is 1 amp. Now let us separate the ends of the two wires farther and farther until finally they are stretched out straight. More and more power will be required from the plate battery to maintain a constant value of current in the wires.

At the same time we shall note that the capacity and inductance of the line have changed somewhat so that some minor changes must be made in the tuning condenser to keep the wires in resonance with the oscillator—but these changes cannot account for the greater power required from the battery to maintain the same current in the wires. In the first place, the changes in capacity and inductance are balanced out each time by readjusting the tuning condenser to resonance. At resonance there is only resistance to impede the flow of current. Every time the reactances have been balanced against each other. Clearly the

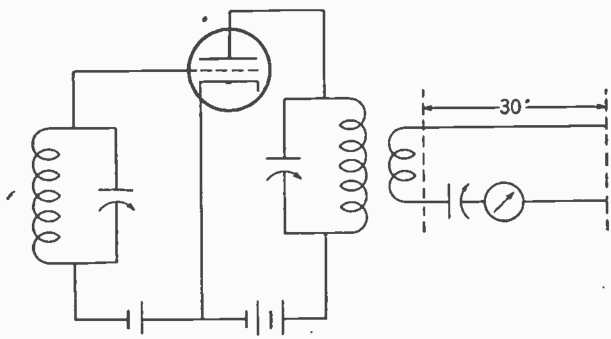


FIG 1–19. Oscillator coupled to a pair of wires to demonstrate concept of radiation resistance.

resistance of the wires has increased. At the same time, if we installed a small receiver near the oscillator and kept moving away from it so that the signals picked up were of constant strength, we should find that, the greater the power taken from the battery, and hence the greater the power into the antenna, the farther away we could hear the signals.

It is apparent that, if the current is still 1 amp but twice as much power is taken from the battery to produce the 1 amp, the resistance of the wires has doubled. The useful part of this resistance is called the **radiation resistance** of the wires, which may now be called the antenna. The power that goes into this resistance is the power that is effective in carrying intelligent communication from the transmitter to the receiver. The resistance we are talking about now is not the d-c resistance which is entirely a function of the diameter and length and material of the conductor. We are talking now about the a-c resistance of the antenna. This is never less than the d-c resistance; it is always greater than this value.

The total resistance of an antenna may be measured by the "added resistance" method used to measure the high-frequency resistance of a coil. If enough resistance is added to our 7-megacycle antenna system to halve the current, this added resistance is equal to the resistance already there. If the resistance is measured at several wavelengths we shall get a curve which will have the general form of that in Fig. 2–10 (taken from the *Proceedings of the Institute of Radio Engineers,* February, 1920). The values of resistance of our 7-megacycle antenna will differ, of course, from those in Fig. 2–19. Antenna resistance is made up of the ohmic resistance of the wires, the losses in the dielectric of the antenna capacity, the loss of energy due to radiation, and other small losses. An efficient antenna is one which has a very high radiation resistance and a low resistance of all other sorts. Then most of the power put into the antenna will be radiated.

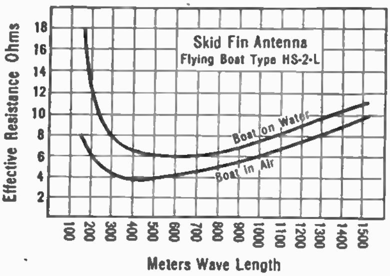


FIG. 2–19. Radiation resistance of an old-fashioned aircraft antenna.

Ground wave—sky wave. Waves of energy are radiated from the antenna in all directions both in a horizontal plane along the earth and in a vertical plane toward the sky. The radiation along the earth tends to follow the earth much as an electrical current follows a copper conductor, for the earth is a conductor, and electrical currents tend to follow a conductor. The **ground wave,** as the energy along the earth is called, is quickly dissipated because its energy is used in setting up fields near the earth, and, since the earth and objects on the earth are not perfect conductors, some of this energy is lost in heating their resistance.

The ground wave from a particular station at a particular point is steady. It is the useful part of the radiation from a broadcast station. It extends out 50 to 100 miles from the antenna, the distance depending upon the power of the transmitter, the construction of the antenna, the resistance of the earth, and the frequency.

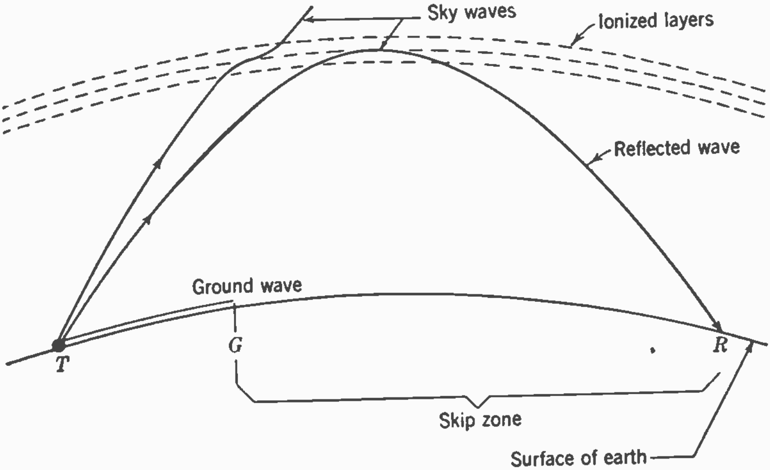


Fig. 3-19. Radiation is bent downward to earth by ionosphere so that it may be detected at great distances from its source. Between the outer limit of the ground wave and the reflected or refracted radiation, no signals can be heard.

The sky wave. Energy radiated into the sky would shoot off into space and be lost except for one fortunate natural phenomenon—a phenomenon discovered through radio. Up in the sky, some 50 to 200 miles above the earth, are several layers of ionized particles produced from the air by ultraviolet radiation from the sun. These layers, acting as reflectors and refractors of radiation, reflect or bend downward radio energy. When, therefore, radio waves shoot off into the upper atmosphere and encounter these ionized particles, part of the radio energy is returned to the earth, part is absorbed by the ionized layer, and part penetrates the ionized layers to escape into outer space.

When the reflected wave reaches the earth again it can be detected just as the direct ray from the radio transmitter or the ground wave can be detected. The reflected wave can reach the earth only at a very great distance from the transmitter. In the intervening distance no signal whatever from the given transmitter may be detected. About the transmitter, therefore, a **skip region** may exist in which the signals are not received. Beyond this region they are returned to earth and may be picked up.

At 7 megacycles the skip distance (distance from transmitter to point where signals are received) is about 200 miles in the daytime; at 15 megacycles it is about 800 miles. The skip distance varies with time of day, frequency, time of year, and position in the 11-year sunspot cycle. By properly choosing the frequency corresponding to these several factors, reliable communication may be maintained between two stations at any distance.

The ionosphere. The several regions of ionization in the upper atmosphere exist in what is known as the ionosphere. The D layer is at an altitude of about 40 km and is effective in reflecting the very low radio frequencies (20 to 550 kc). The E layer, also known as the Kennelly-Heaviside region, about 100 km above the earth, reflects frequencies from 500 to about 1500 kc, but during the day it is so heavily ionized that most of these waves are absorbed. After dark, when the sun is no longer effective in producing ionization, reflections occur and broadcast stations may be heard over much greater distances.

The F layer, some 200 km above the earth, reflects the short waves (1500 to 30,000 kc). Because of this ionized region, communication is possible over very great distances day or night.

These layers of ionization are not stable but vary from moment to moment, and so reflections of radio energy are not stable but vary in much the same way as the ionization varies. Throughout this discussion the word "reflect" is used to indicate the bending downward to earth of radio waves, whether the actual action is reflection or refraction.

Fading. Suppose that a receiving station is within range of the ground wave of a 1000-kc transmitter, say at the edge of it.

Signals come through in the daytime clear and steady. At night, however, when reflections occur, the receiver begins to pick up the reflected wave from the 1000-kc transmitter as well as the ground wave. The waves may be in phase and so reinforce each other, or they may be out of phase and then they will nullify each other. If the ionized layer shifts in height or in density, the sky wave may vary in intensity so that the alternate additions and subtractions of energy to and from the ground wave by the sky wave cause the voltages at the receiving antenna to vary from time to time. This variation is the cause of fading.

On the broadcast band, the sky wave is absorbed so that it has little effect in the daytime. Receivers 50 to 100 miles away from the average transmitter, therefore, get clear signals, but they get very little voltage from distant stations. At night, the sky waves from distant stations are reflected downward to the receiving station and long-distance reception is possible. If the ionized layers are comparatively stable, as during times of weak sunspot activity, reception from a distant station can be fairly stable. Fading will still occur, but the a-v-c system of the receiver will eliminate most of the trouble.

The a-v-c system, however, does not help reception from a station 50 to 100 miles away from which both sky and ground wave are received. Signals from these stations, the best and often the only stations in the daytime, may be badly garbled because of the variations in magnitude and phase of the sky wave.

Sometimes the side bands carrying the modulation fade more than the carrier of the station, resulting in **selective fading.** No a-v-c system and no amount of increase in power of the transmitter can overcome this trouble.

Short-wave transmission. In the broadcast band the ground wave is quite usable; the sky wave is very undependable. Above about 1500 kc the attenuation of the ground wave increases rapidly, and the effectiveness of the sky wave increases. For this reason short-wave communication takes place mainly by means of reflected waves. Antennas are engineered to aim the radiation up into the sky at the correct angle to be reflected downward to the desired receiver station. Frequencies higher than about 30 megacycles are not so easily reflected, tending to shoot on through the ionized layers. Communication at these frequencies depends upon a wave received in a direct line from the transmitter; and, at frequencies of the order of 100 megacycles, very little energy is radiated beyond the distance the transmitter can be seen (line-of-sight transmission).

On the very high frequencies, therefore, it becomes necessary to erect the transmitting and receiving antennas as high as pos-

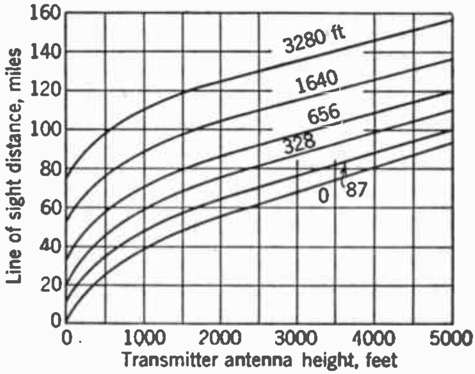


FIG. 4–19. Length of line-of-sight paths for receiving and transmitting antennas of various heights.

sible above the surface of the earth since radiation at these radio frequencies acts like light waves. The transmitter resembles a lighthouse, sending its radiation out at various angles to the antenna, to be picked up as far away as the antenna can be seen. Actually, effective communication can be carried out at greater distances than line-of-sight would indicate, sometimes two or more times this distance.

Types of antennas. All antennas may be thought of as made up of two very simple types, the Hertz half-wave radiator (often called a **dipole** or **doublet**) and the Marconi grounded antenna which may vary in length from something under 0.25 wavelength to about 0.6 wavelength in the broadcast band. The Hertz antenna has such physical and electrical dimensions that one-half of a complete wave of current or voltage appears along it.

In Fig. 5–19 it is apparent that "standing waves" of current and voltage will occur along the wire stretched out straight. That is, the current measured at various places along the wire

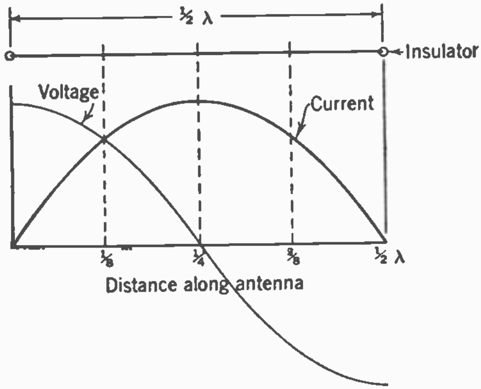


Fig. 5–19. Current and voltage distribution along half-wave antenna.

increases and decreases in regular fashion. If the distance between two maxima is measured it will be found to be one-half of the wavelength of the energy being fed to the wire. If just this much of the wire is used as the antenna, it becomes a Hertz dipole or doublet.

In such an antenna, a maximum of current occurs at the same point as a minimum of voltage. Since there can be no current at the end of the wire, the voltage there must be a maximum.

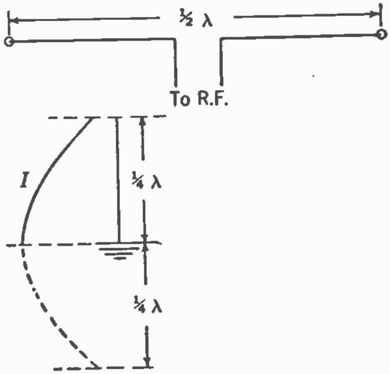


Fig. 6–19. Half-wave antenna fed at center compared with quarter-wave grounded antenna.

The antenna acts like a resonant circuit except that its inductance and capacitance are distributed along its length uniformly rather than being concentrated in isolated pieces of equipment.

It will resonate to a certain frequency or wavelength of energy. The Hertz antenna may be erected vertically or horizontally, and energy may be fed to it in several ways.

A quarter-wave Marconi antenna is either grounded at one end or connected to a mesh of conductors called a **counterpoise.** The current distribution along its length will be such that it will look like one half of a half-wave antenna, the other half being an image of the actual antenna. The grounded quarter-wave antenna acts like a half-wave ungrounded antenna.

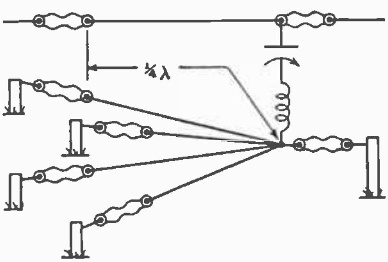


FIG. 7-19. Hertz antenna erected over a counterpoise or artificial ground of low loss.

A long wire may be thought of as a series of half-wave or quarter-wave antennas. Directional antennas are made up of numerous sections of quarter- or half-wave antennas, all properly connected so that the currents are in phase with one another. The elementary sections add up their individual components of radiation, producing a sum greater than the effect of any individual. If, at the same time, the individual elements are directive and if a reflector behind the antenna can be used to cancel the radiation in a direction opposite to the desired direction, or can turn the energy around in the proper phase to go along with the desired radiation, the directive effect may be exceedingly great—equal to one hundred times the power of a single antenna.

The physical length of the antenna depends upon the frequency of the desired radiation and the nearness of the antenna to surrounding objects. But a free wire in space will have a length as shown below. If a wire in free space is excited by a transmitter in such a way that one complete wave of current exists along its length, this length will be

$$\frac{1}{f\sqrt{LC}}$$

where L and C are the inductance and capacitance per unit length of the wire and f is the frequency.

The actual physical length of a half-wave doublet in feet is given by

$$L = \frac{468}{f} \text{ ft}$$

where f is the frequency of operation in megacycles.

The corresponding length for a quarter-wave antenna from the far end to the ground or to the counterpoise is approximately

$$L = \frac{236}{f} \text{ ft}$$

where f is the frequency in megacycles.

Antenna characteristics. Antennas have inductance, capacitance, and resistance. The only useful part of the resistance is the radiation resistance; the rest of the resistance, made up of the ohmic resistance and the loss of energy in dielectrics near the antenna, represents energy wasted. The radiation resistance may be represented by resistance which, inserted in place of the antenna, would absorb as much power as the radiated energy represents. If this radiated energy could be measured, and if the current at the center of the half-wave antenna or at the base of the quarter-wave antenna could be measured, then the radiation resistance would be

$$R = \frac{P}{I^2}$$

where P is the radiated power and I is the antenna current.

Now, the longer the wavelength to be transmitted, the larger must be the physical structure of the antenna to resonate to this wavelength. On the very long waves it is practically impossible to make the antenna big enough to be resonant to the transmitting wavelength, and resonance is secured by "loading" the antenna with inductance. Any I^2R losses in the loading coil are wasted as far as radiation is concerned.

An antenna operated at its natural wavelength (a half-wave antenna) is very efficient because its radiation resistance is high, but an antenna which is small compared to its resonant wavelength is inefficient because most of its resistance is ohmic resistance and losses other than radiation. Therefore most of the power poured into it goes to heating the wire and the surrounding physical objects rather than into radiation.

The short waves are so useful simply because it is possible to make very efficient antennas operate on these frequencies. Most of the power put into short-wave antennas goes into radiation.

A half-wave antenna operating at 40 meters will have a physical length of approximately 67 ft. A similar antenna to operate on 15,000 meters (20 kc) would be 23,400 ft long. The two will have the same radiation resistance, approximately 70 ohms, which is a large proportion of the total resistance of a 40-meter antenna but a much smaller proportion of a 20-kc antenna.

As a matter of fact, the large antennas used at long wavelengths have radiation resistances of the order of 1 ohm or less, and so several hundred kilowatts of energy must be fed into them to get any appreciable radiated power.

Directional characteristics of antennas. The simple dipole antenna is not an all-direction radiator. For example, a vertical doublet will radiate into a doughnut-shaped pattern. No radiation will come off the end of the antenna, and the maximum coming from the sides is at right angles to the antenna. A horizontal half-wave doublet will have the same pattern except that it will be oriented up and down rather than in a horizontal plane.

If the antenna is lengthened so that it is one wavelength long, instead of a half wavelength, at the operating frequency, the pattern becomes more complex, there being two elongated doughnuts at an angle of 54 degrees to the line of the antenna. If the antenna is still further lengthened, more and more lobes of radiation occur, the strongest radiation occurring in a lobe which gets nearer to the direction of the antenna as the length increases. These same effects may be obtained by operating a given antenna at a higher and higher frequency, so that it is several wavelengths long at the operating frequency.

An antenna operated at a frequency higher than that which makes it a half-wave radiator tends to radiate its energy nearer and nearer the horizon if the antenna is horizontal, or more nearly vertical if the antenna is vertical.

The presence of ground or other objects changes the pattern of radiation from an antenna, and very precise use is made of this effect when erecting highly directive antennas for reception or transmission.

Where it is necessary to receive signals over a wide range of wavelengths, as in the home where signals over the whole broad-

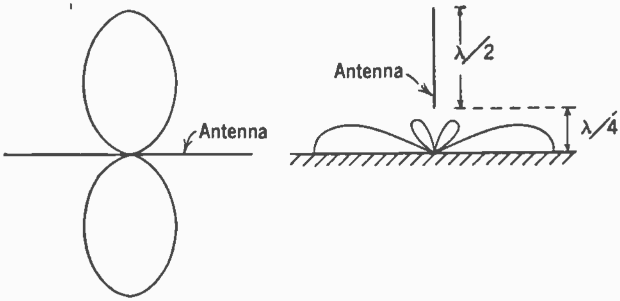


FIG. 8–19. Pattern of radiation from horizontal and vertical doublets located above ground.

cast band must be picked up, or on shipboard where quite a large spectrum must be covered, no attempt is made to use antennas resonant to any particular wavelength. A vertical or horizontal wire, straight or bent, or a wire around the molding of the room —practically any metallic object—may be used. This is not so efficient as an antenna which can be tuned to each frequency to be covered. The loss in efficiency, however, can easily be made up by greater r-f amplification in the receiver.

When two or more lobes of energy are radiated from an antenna, one lies nearer the horizon than the other. This directivity in a vertical plane may be made use of, depending upon whether the *ground* wave or the *sky* is the more useful. For example, in short-wave work where long distances are to be covered, the high-angle radiation is most useful. This wave, directed skyward, is reflected by the ionized layer high up in the atmosphere and it is on its downward path that it is received.

The angle at which the radiation leaves the antenna governs the point on the earth's surface where it returns.

On the other hand, broadcast station radiation is directed along the ground as much as possible so that people near the antenna will get the best possible service. The sky wave is repressed so that it will not return to earth at some distant point where it may be out of phase with the ground wave and cause intermittent fading as the reflected wave varies in strength. Broadcast antennas are usually operated below their natural wavelength, i.e., the antenna is not a half wave long. This keeps the ground radiation more effective than the skyward radiation. On the other hand, short-wave antennas, where the useful wave is to be reflected back to earth, are operated at harmonics of the natural wavelength, which means that the antenna is very much larger than a half wavelength for the desired radiated frequency.

Antenna arrays. When it is desired to concentrate radiation in a given direction, combinations of half- and quarter-wave antennas are used in such a manner that the radiation from each is combined in the desired direction and suppressed in the unwanted direction. For example, Fig. 9–19 shows a "broadside" array made up of a number of radiators each somewhat less than a quarter wavelength long and spaced one-eighth wavelength apart. All the elements are energized in phase; directivity is at right angles to the plane of the array. A similar array may be placed to the rear of the first to act as a reflector.

In another general type of array, a number of identical radiators are arranged in a line carrying equal currents. The phase difference between adjacent elements is equal to the spacing between these elements in wavelengths. If the radiators are a quarter wave apart (90 degrees) there will be a phase difference between adjacent elements of 90 degrees.

Some of these structures may be very complex mechanically, but, on the other hand, in the broadcast frequency spectrum the directive system may be fairly simple. For example, a simple vertical radiator has a circular pattern of radiation. If most of the desired audience lies on one half of the circle, the majority of the energy may be concentrated in this direction by distorting the circular pattern by exciting additional antennas, placed behind the single radiator, and fed in the proper phase with respect to the principal antenna. Such a system automatically reduces the radiation in the undesired direction and creates less interference in this direction so that the same frequency may be used by another broadcast station in another part of the country.

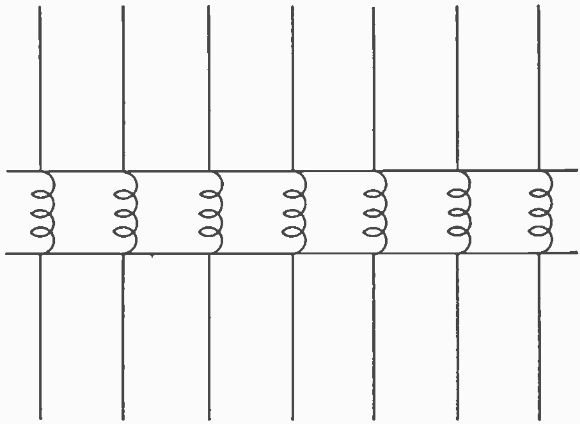


FIG. 9–19. Directive antenna made up of numerous quarter-wave elements properly spaced and connected so that all elements are in phase.

The loop antenna. The loop is a directional antenna. It receives better in the direction towards which the narrow dimension points. For example, its pickup ability is such that when it is pointed east-west it will pick up very little energy from a north-south direction. In connection with a sensitive receiver, it may be used to determine the direction whence the signals come. It is the heart of the radio compass and of the **direction-finding** stations which are situated along the seacoasts of the world. When a ship wants bearings, its signals are picked up by the coastal station which determines the position of its receiving loop which gives the least signal. A compass is attached to the base of the loop, and the indicator then points out the bearing of the vessel. A receiving operator in another location also swings his loop on the vessel and in this way two bearings are obtained. From them the master of the ship can tell where he is with regard to the coast line.

There is one disadvantage of the loop when used alone. Its directivity pattern is composed of two equal loops, so that the operator can tell that a signal is coming from a north-south di-

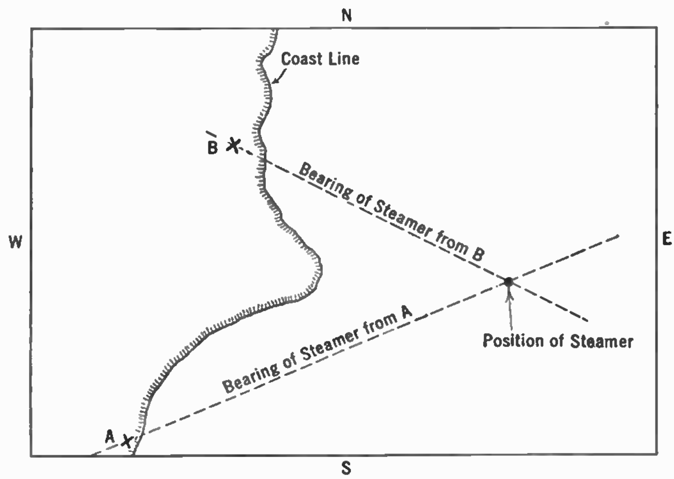


FIG. 10–19. Method of plotting a ship's position by obtaining two bearings from land stations.

rection rather than an east-west direction, but he cannot determine whether the signal is coming from south or from the north. The addition of a single vertical antenna whose output is properly added to that produced by the loop adequately solves this problem. The difference in phase between the signal picked up by one side of the loop and the vertical antenna is used to add to or subtract from the current in this half of the loop. The resultant directional characteristic is shown in Fig. 11–9.

Getting energy into an antenna. Since the best type of simple transmitting antenna is a half-wave dipole high in the air and in free space, getting power into it is a problem. No power should be lost in the connecting system between the antenna and the transmitter. A low-loss transmission line of the proper impedance is required. The impedance of a half-wave antenna at the center is approximately 70 ohms, increasing towards the ends to 2400 ohms. The transmission line, therefore, should have an impedance of 70 ohms or should be coupled to the antenna so that the line presents an impedance of 70 ohms to the antenna.

At the transmitter end the transmission line must be coupled to the power amplifier so that the power amplifier has an output

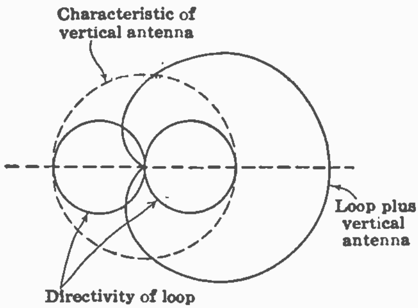


FIG. 11–19. Directional characteristics of loop, vertical antenna, and two combined.

impedance of 70 ohms. A transformer of some sort, perhaps merely a tap on the output-amplifier tank circuit, is required.

Since the impedance of the antenna varies from one point to another, transmission lines of higher impedance than 70 ohms can be used provided that they are connected to the antenna at the proper point. A quarter-wave antenna has an impedance of 37 ohms, approximately, and all antennas shorter than a quarter wavelength have lower impedances than 37 ohms. The antenna current, therefore, will be higher than in a quarter-wavelength radiator, and the losses in ground or wire resistance will be greater.

Transmission lines. Several types of lines are used to get energy from the transmitter to the antenna: single-wire lines, two-wire lines, coaxial cable lines, etc. A single-wire line may be attached to a point of the antenna where the impedance is approximately 500 ohms. Such a single wire will have an impedance of this value.

Two-wire lines have an impedance varying between 400 and 800 ohms according to the formula

$$Z = 276 \log_{10} \left(\frac{a}{b}\right)$$

where Z = the impedance in ohms.
a = the spacing between wire centers in inches.
b = the wire radius in inches.

For example, a pair of No. 14 wires, spaced 5 in. apart, has an impedance of approximately 600 ohms.

Coaxial line is made up of two conductors, one located within the other, maintained in its position by insulating spacers. The impedance of such a line is lower than that of the open wire line; it may be found from the formula

$$Z = 138 \log_{10} \left(\frac{d}{a}\right)$$

where Z = the impedance in ohms.
a = the outside diameter of the inner wire in inches.
d = the inside diameter of the outer conductor in inches.

For example, a/d is equal to 0.31 for a 70-ohm line. If the two conductors have diameters of ¾ and ¼ in., the impedance is 66 ohms. In general, the characteristic or surge impedance of a transmission line is equal to $\sqrt{L/C}$, where L and C are farads and henries per unit length.

The loss in energy in these lines is very low if they are well made. The two-wire line has a loss of $k\sqrt{f}$ db per 1000 ft, where k varies from 0.1 to 0.2, depending upon the insulating material used, etc. The loss in coaxial cable is $0.256\sqrt{f}/d$ per 1000 ft, where d is the outer conductor diameter, and in both formulas f is the frequency in megacycles. The impedance of coaxial and two-wire lines is shown in Fig. 4–18.

In Fig. 12–19 are shown two ways by which two-wire lines may be connected to a half-wave antenna. Of course, a 70-ohm concentric cable line may be connected directly to the center of a half-wave antenna since its impedance is of this approximate value.

Quarter- and half-wave lines. Transmission lines of a particular length for the frequency for which they are designed have very interesting and useful characteristics. A quarter-wave line has such a physical length that it is a quarter wavelength long at the operating frequency. Every line has a characteristic impedance, and for minimum loss in the line the impedances between which it works should be equal to the impedance of the line. If, however, two unequal impedances are to be connected, say a generator and a load, then a quarter-wave line having a characteristic impedance equal to the geometric mean of the two unequal impedances may be used as the connecting link, very much as a transformer is used. The impedance of the line should be equal to $\sqrt{Z_1Z_2}$, where Z_1 and Z_2 are the impedances connected to the two ends of the line.

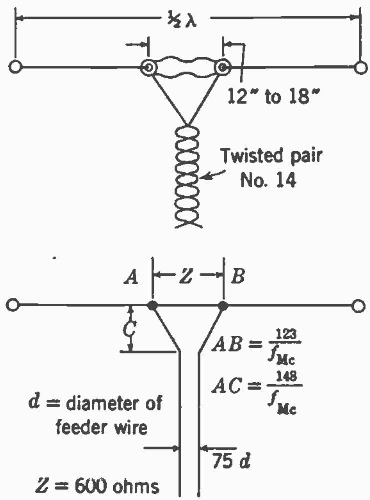


FIG. 12–19. Manner of connecting transmission lines to antenna to produce proper impedance match.

For example, if impedances of 600 ohms and 70 ohms, such as a 600-ohm transmission line and a 70-ohm doublet, are to be connected, a quarter-wave line may be used between them to minimize loss due to impedance mismatch. Its impedance should be $\sqrt{600 \times 70} = 206$ ohms.

A half-wave line, on the other hand, has a quite different quality. The impedance looking into a half-wave line is equal to the impedance terminating the other end of the line. This is true, regardless of the impedance of the line itself. Thus a half-wave line acts like a one-to-one ratio transformer and may be used as such. It is understood that a line is a half- or quarter-wave line at only one frequency. The physical length of a quarter-wave line is equal to

$$L = \frac{246V}{f} \text{ ft}$$

where V varies, depending upon the kind of transmission line used (coaxial, $V = 0.85$; two-wire line, $V = 0.975$); and f is the frequency in megacycles.

A half-wave line open-circuited at the far end has a high impedance looking into the line from the transmitting (near) end. On the other hand, if a quarter-wave line is short-circuited at the far end, the impedance looking into it is high. Either line resembles a parallel circuit tuned to resonance. If, for some reason, it is desirable to connect an impedance across a circuit, say an antenna, having a high value at the circuit frequency but a low value to all other frequencies, the short-circuited quarter-wave line is employed. The analogy to this usage is a parallel tuned circuit shunted across a circuit to drain off voltages of unwanted frequencies, since currents lower in frequency than the resonant frequency will find a path of low impedance through the inductance, and higher-frequency currents will find a path of low impedance through the condenser.

TRANSMISSION LINE CHARACTERISTICS

Length of Line	Impedance at Far End	Z Looking into Line	Equivalent Circuit
$\frac{1}{4}\lambda$	High Zero Z_0	Zero High $\sqrt{Z_i Z_0}$ *	Series tuned circuit Parallel tuned circuit Transformer
$\frac{1}{2}\lambda$	High Zero Z_0	High Zero Z_0 †	Parallel tuned circuit Series tuned circuit 1–1 transformer

* Value of line Z to be used between Z_i and Z_0.
† Independent of line impedance.

An interesting use of a quarter-wave line as a matching transformer is shown in Fig. 13–19. Here it is desired to connect a high-impedance line to a low-impedance antenna, for example. The quarter-wave line is connected to the center of the dipole. At the far end where the current is low the impedance will be

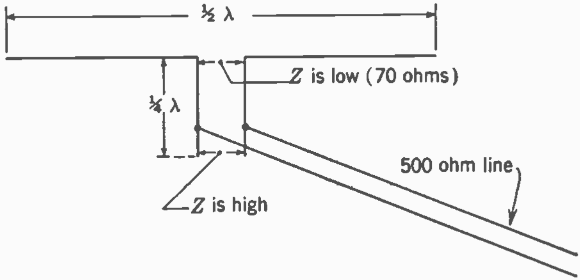


FIG. 13–19. Use of quarter-wave open line to match 500-ohm line to a 70-ohm antenna.

high, and by selecting the proper point to connect the high-impedance transmission line to this quarter-wave matching stub the impedance of the transmission line will be properly terminated and the line will become non-resonant.

Problem 1–19. In Fig. 14–19 is shown a two-wire quarter-wave line connecting a 500-ohm line to a dipole (72 ohms). If the frequency of operation is 7150 kc, determine: (1) the length of the antenna; (2) the surge impedance of the quarter-wave line; (3) the physical length of the quarter-wave line and, from Fig. 4–18, the dimensions of the line.

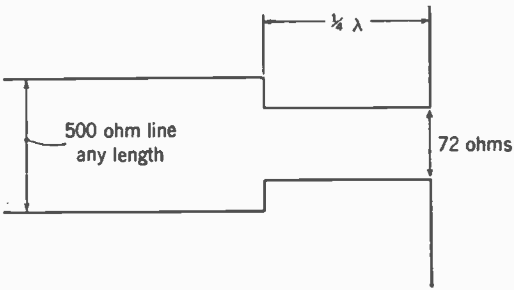


FIG. 14–19. Use of quarter-wave line as a matching transformer between two unequal impedances.

Resonant and non-resonant lines. All transmission lines have a characteristic impedance which is equal to $\sqrt{L/C}$, where L and C are the inductance and capacitance per unit length. If a line is terminated in a resistance equal to its characteristic impedance, this load will take the maximum power from the line. The current through and the voltage across the line measured at intervals along the line will be found to fall off regularly from their original values to consistently lower ones such that the impedance at any point looking toward the load end is equal to the ratio between the voltage and current at that point. Thus, at any point along a line which is terminated in $R = Z_0$, $Z = E/I = Z_0$.

If, however, the line is terminated in some other resistance, and if the current and voltage are measured at regular intervals along the line, it will be found that both current and voltage go through maxima and minima, the maximum of voltage occurring at the point where a minimum of current occurs, the successive maxima (or minima) of current (or voltage) being a half wave apart. In other words, the line resembles an antenna.

Under these conditions the line is said to be **resonant;** the variations of current and voltage are called **standing waves.** The impedance at any point along the line varies between a maximum (when E is high and I is low) to a minimum (when E is low and I is high).

Since instantaneous power is the product of E and I, it follows that in a resonant line some of the power put in at the sending end is absorbed in the terminal resistance and some is reflected back toward the sending end by the mismatch at the receiving end. This reflected power accounts for the standing waves along the line. In a non-resonant line (a line terminated in its characteristic impedance), all the power put into it is consumed by the load resistance except that lost in the inherent resistance of the conductors.

Since the voltage at regular intervals along the resonant line is higher than that at any point along a non-resonant line, more trouble will be experienced in the line insulation. Since the current at periodic intervals along the line is higher than in a non-resonant line, more power will be wasted in the I^2R losses along the line.

A resonant line may be used to connect a transmitter to an antenna, the advantage being that no impedance matching at the antenna end is needed. The line really acts like an extension of the antenna, but, because its two conductors are close together, it does not radiate much energy.

In general, however, it is customary to use non-resonant lines wherever possible. Open-wire lines will have fairly high imped-

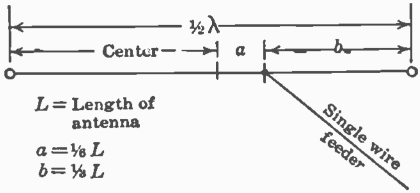


FIG. 15–19. The proper place to connect a single-wire feeder to an antenna for most efficient power transfer.

ance, four-wire lines having impedances of 150 to 400 ohms and two-wire lines having impedances varying from 100 to 800 ohms, depending upon the wire spacing and wire diameters. Concentric lines, in which one conductor is inside the other, have impedances of the order of 30 ohms to 200 ohms.

The concentric line, therefore, is ideal for coupling a transmitter to a half-wave doublet antenna which has an impedance of approximately 70 ohms at the center. If a high-impedance line, say 600 ohms, must be used, it can be connected to the antenna as in Fig. 12–19, where the two wires of the line are shown attached to the antenna at a spacing off center determined by the line impedance. Such a connection is possible because the half-wave antenna has an impedance which varies from its low value at the center to a high value at the two ends with intermediate impedances at intermediate points.

A single-wire line may be connected to an antenna as shown in Fig. 15–19. It is considered that the antenna capacitance to ground and the ground furnish the return circuit.

Line characteristics. Since a quarter-wave line open-circuited at the far end acts like a series tuned circuit, it follows that the line has an inductive reactance if it is less than a quarter wave long and a capacitive reactance if open at the far end and is less than a quarter wavelength long. A half-wave line closed by a capacitance at the far end has a capacitive reactance when one measures the reactance from the near end and has inductive reactance if it is closed by an inductance at the far end.

CHAPTER 20

FREQUENCY MODULATION

Within recent years it has been found that, if the carrier **frequency** of a transmitter is modulated instead of the carrier **amplitude,** certain advantages will be secured at the cost of some disadvantages. Among the advantages are: (1) an immediate gain in signal strength because the transmitter can radiate at full power at all times regardless of the strength of the audio modulation; (2) a reduction in noise with a resultant effective increase in signal-to-noise ratio, the extent of this gain being dependent upon the width of the frequency band that is employed; (3) a reduction in interference from other stations on the same channel by a factor of 30 to 1 compared to amplitude modulation. All these advantages are in the same direction; i.e., they add up to what could be secured in an amplitude modulation system only by the use of vastly greater power. To these advantages must be added the ability to transmit a wide a-f band with the consequent possibility of high-fidelity reproduction of music.

The disadvantage is that a very much wider frequency band is required, compared to amplitude modulation, to secure all these advantages.

F-m advantages. 1. In an a-m transmitter, the strength of the signal at the receiver is controlled both by the power of the carrier and by the power in the audio modulation. This varies from instant to instant, depending upon the strength of the signal picked up by the microphone. The average modulation of the average broadcast station is, perhaps, 50 per cent. With this degree of modulation there is a 6 per cent increase in transmitter antenna current compared to the unmodulated condition. But if the transmitter could be 100 per cent modulated all the time there would be a 22.6 per cent increase in antenna current and the receiver would pick up more than three times the signal voltage compared to "average" modulation. This would amount to an increase in power at the transmitter of almost two times, so that a 50-kw station would automatically become a 100-kw station as far as the receiver is concerned.

There is another fact, important from the standpoint of the transmitter owner. A 50-kw transmitter must have the ability to handle without distortion as much power as 200 kw during the occasional peaks of 100 per cent modulation. This is because, at full modulation, the peak plate current flowing in the final power amplifier will have doubled, for 100 per cent modulation in an a-m system means simply that there is as much a-f voltage as there is unmodulated r-f voltage in the power stage. Therefore the power, being proportional to the square of the current (or voltage), will have increased four times.

In other words, a 50-kw station must be engineered as if it were to transmit 200 kw although the portion of the time that such power is sent out is extremely small.

On the other hand, an f-m station is modulated by changing the frequency and not the amplitude of the antenna current. There is no reason why the antenna current, and thus the power taken by the final amplifier, cannot be reduced compared to the a-m station to get the same result; or, if the same power is used, to get a fourfold increase in useful power as far as the receiver is concerned. Full power can be radiated at all times from an f-m station and not merely on peaks of modulation as in an a-m station.

2. Greater freedom from interference. In an a-m system, interference must be reduced to one-hundredth of the strength of a desired signal if the interference is to be reduced to 1 per cent. In an f-m system, this same reduction in interference will be attained if the interfering station can be reduced to only one-third the strength of the desired signal. In other words, much stronger signals from interfering stations can be tolerated before any audible interference results. This means that stations on the same channels can be put much closer together geographically.

3. Reduction in noise. In an a-m system, the amount of noise that is permitted to enter the system is a function of the band width. In a narrow-band system, such as a radio-telephone circuit utilizing frequencies from, say, 250 to 2500 cycles, much less noise both from natural and from man-made static will enter the band passed by the antenna circuit, and succeeding circuits, than in a high-fidelity system passing all frequencies from, say, 30 to 15,000 cycles. The band width is roughly seven times greater in the f-m system. In an a-m set-up there is only one way to override noise with signal, and that is to increase the power of the signal. When the signal compared to the noise at the receiver is high, the noise is not heard. In an f-m system, let us assume that noise is both amplitude- and frequency-modulated. But let us modulate our system very much heavier by frequency changes than any noise can be, and then let us remove from our system any and all a-m signals. Now there can be as much a-m noise as ever, but our system will not be responsive to it. The harder we modulate in frequency (wider frequency band used), the less trouble we will have with noise.

It is a fact that static is more amplitude- than frequency-modulated. Now, let us use a "limiter" which removes any change in the amplitude of the desired signal even if modulated by interference, static, or other station's signals. Let us also amplify the higher modulation (audio) frequencies more than the low frequencies by a factor proportional to their frequency. This is done at the transmitter before modulation so that the higher frequencies are relatively stronger, when modulated, than they would be without the additional amplification. This process is called **pre-emphasis.** At the receiver, we will de-emphasize the higher audio components by the same factor so that the over-all effect is as if no pre-emphasis had been used. If, however, somewhere in the transmission system noise enters having a frequency within the pre-emphasis band, it will have less effect by the percentage the higher frequencies are emphasized. Since most of the power is ordinarily contained in the lower frequencies, the peak amplitude of the transmitter is not increased but the signal-to-noise ratio is increased with an accompanying gain to the listener. Such is the reasoning behind modern f-m systems.

Wide-band frequency modulation. That a really high-fidelity transmission can take place is due to the following circumstance. Since the frequency is to be changed, and not the amplitude, it follows from experience as well as from mathematical studies that a much greater band width is required than is necessary in an a-m system. In fact, the wider the band utilized, the greater the advantage of frequency modulation from the standpoint of freedom from noise. Since a wide band is to be used, it might as well transmit high-fidelity signals. F-m transmission could take place on the standard broadcast band, but the band width required for each station would be so great that there would be room for only about seven or eight stations, so that it is necessary to go to the short-wave regions where band widths can be relatively much wider.

F-m broadcasting can now take place on 40 channels, each 200 kc wide, in the band from 42 to 50 megacycles. Each station actually uses a band width only 150 kc wide, 75 kc on each side of the carrier frequency; and the remainder of the separation between stations is a guard band to reduce interference between stations on adjacent channels.

On these frequencies there is less natural static than on the 550- to 1600-kc band allotted to broadcasting; the range of each transmitter is more or less limited to two or three times line-of-sight distance with no interference created by stations on the same channel provided that they are several hundred miles apart; there is no fading; the day and night range is the same; the band width is sufficient so that all audio frequencies between 30 and 15,000 cycles can be transmitted, with the result that broadcast listeners have a chance to hear music in the home exactly as it is produced at the transmitter and not restricted in audio range by the limitations of the 10-kc separation necessary in the standard broadcast band.

Production of frequency modulation. Let us consider a very simple, but unworkable, system first. What is required is some method of varying the carrier frequency of the transmitter linearly in accordance with the strength of the modulation. That is, if 1 volt of audio modulation changes the frequency 1 kc, 2 volts must change it 2 kc, and 3 volts 3 kc, and so on.

A very simple method of varying the frequency of an oscillator is to change the inductance or capacitance in the tuned circuit. If, therefore, we place a condenser microphone across the tuning condenser and talk into it, the capacitance of the microphone varies, and naturally the frequency of the tuned circuit changes accordingly. We have modulated the frequency of the oscillator in our transmitter. The *amount* we have changed the frequency depends entirely upon how loudly we talk into the microphone. The *rate* at which the frequency is changed, that is the number of times per second it is shifted, depends upon the frequencies of the sounds that the microphone picks up.

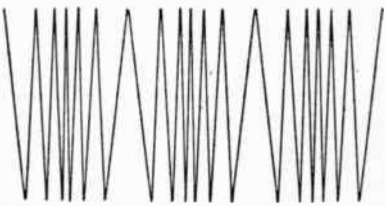


FIG. 1–20. Constant-amplitude, frequency-modulated wave.

The amplitude of the carrier need not change at all. If the microphone picks up a 1000-cycle tone of 1 volt, the carrier frequency of the transmitter will shift from, let us say, 50 megacycles to 50 megacycles plus or minus 1 kc, and this frequency will shift back and forth from its average, unmodulated, value of 50 megacycles to this new value 1000 times per second. If the microphone picks up 2 volts of 1000-cycle tone, the carrier of the station may shift to 50 megacycles plus or minus 2 kc, but the amplitude will be the same as before. If the microphone picks up 2 volts of 2000-cycle tone, the carrier will shift from 50 megacycles to 50 megacycles plus or minus 2 kc but at a rate of 2000 times per second instead of the former rate of 1000 times per second. (Compare Fig. 1–20 with Fig. 10–18.)

In one type of transmitter in common use a tube circuit very similar to the automatic-frequency-control circuits described in the chapter on receivers is employed. This circuit has the ability to act like a varying reactance as the grid voltage of the tube connected to the circuit is changed. This varying reactance (either inductive or capacitive) may be used to control the frequency of the oscillator of the transmitter.

The Armstrong system. Major E. H. Armstrong must be credited with the present interest in and appreciation of the benefits of wide-band f-m transmission. In spite of the greatest skepticism of those well established in a-m broadcasting, he persisted in his belief that frequency modulation had great advantages for broadcasting as well as for communication systems where narrower bands could be used.

His method of producing f-m signals may be described as follows. Suppose that we have an a-c generator in which the stator can be moved with reference to the rotor.* The rotor is driven at a constant speed corresponding to the carrier frequency of the f-m station. If the stator is rocked about its midposition with a *velocity* corresponding to instantaneous values of the modulation, then the output of the alternator will be frequency-modulated. Why is this so?

Consider the moment when the stator is rocked ahead so that it is moving in the same direction as the rotor. In effect the rotor has slowed down, since it requires longer for any point on it to move past a given point on the stator. The frequency output of the alternator is, then, decreased, since the effective rotor speed is decreased. Now, if the stator is moved in a direction opposite to that in which the rotor moves, then two points on the stator and rotor move past each other at a greater speed, and the frequency will increase.

Let us suppose that the stator is moved ahead and back with a speed so that its *position* (not its speed) with reference to its unmodulated position varies in proportion to the amplitude of the modulation. The effect upon the output may be expressed by stating that the phase of the output varies from the unmodulated condition by a value proportional to the amplitude of the modulations. This is known as phase modulation.

* *Applied Electronics,* Electrical Engineering Staff, M.I.T., John Wiley & Sons, 1943.

Phase modulation is closely related to frequency modulation with the following difference. In frequency modulation the amplitude of the frequency change (deviation) is independent of the modulating frequency, but in phase modulation the amplitude of the frequency deviation is directly proportional to the modulation frequency.

In the Armstrong system, phase modulation is employed, but, by means of a filter which gives an output inversely proportional to the modulating frequency, the Armstrong system produces frequency modulation in the end. The effect is the same as if a more direct method of varying the carrier frequency in accordance with audio modulation were employed. A small degree of phase shift is created in the first circuit, but the frequency of that circuit is then multiplied several times in each succeeding stage, and at each multiplication the frequency shift is increased so that, by the time the frequency of a crystal oscillator (about 200 kc) has been multiplied to 50 megacycles, sufficient frequency modulation has been produced. A frequency multiplication of 3000 times is actually required. As this amount cannot be obtained in raising the 200-kc crystal frequency to 50 megacycles, at some stage in the system the frequency up to that point must be heterodyned down. This procedure changes the frequency only (not the deviation), and further frequency multiplication increases the deviation until the full effect is secured. These various functions can be performed in small tubes so that the power required is not great.

F-m terminology. Several new terms will be found in f-m literature. The **frequency deviation** is the maximum change in the average frequency secured in the system. Thus a 50-megacycle carrier shifted to plus or minus 75 kc on modulation peaks has a frequency deviation of 75 kc. The **deviation ratio** is the frequency deviation divided by the highest audio frequency to be transmitted. The greater this ratio, the greater will be the signal-to-noise ratio at the receiver (except on weak signals as mentioned below). A broadcast transmitter having a deviation ratio of 5 to 1 (75 kc ÷ 15,000 cycles) will present the same effectiveness against noise as a police transmitter in which the highest audio tone is 3000 cycles and which is deviated a maximum of 15,000 cycles each side of the average carrier frequency.

It has been determined that, on weak ("threshold") signals, a narrower band width at the receiver will have a better signal-to-noise ratio than a wide-band system because it will admit less noise, but, within the threshold distance, the wide-band system will be better from the standpoint of noise reduction.

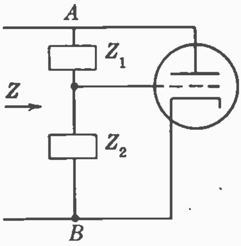


FIG. 2–20. Elements of the reactance tube modulator. Any variation in grid bias changes the phase of Z looking into the points A and B.

In an f-m system the percentage modulation may be considered the ratio of the actual deviation to the maximum deviation and is entirely a function of the strength of the voltages picked up by the microphone.

Reactance tube modulator. It was mentioned above that one simple method of producing frequency modulation would be to vary the frequency output of the oscillator of the transmitter. This is exactly what is done in one system of f-m transmission. It uses a tube and circuit called a **reactance tube modulator,** a most ingenious application of electron-tube principles.

Consider the circuit in Fig. 2–20, in which two impedances are in the plate circuit of an amplifier. Now, no matter where the grid voltage comes from, the current I_p through these two impedances will be approximately

$$I_p = g_m E_g$$

and, if E is the total voltage across the impedances and the impedance looking into these two plate circuit loads is E/I,

$$Z = \frac{E}{I_p} = \frac{E}{g_m E_g}$$

The total voltage across the two impedances is equal to the current through them times their sum, or

$$E = I_p(Z_1 + Z_2)$$

Now let us suppose that the grid voltage is the voltage across Z_2

$$E_g = I_pZ_2$$

and

$$Z = \frac{I_p(Z_1 + Z_2)}{I_pZ_2g_m} = \frac{Z_1 + Z_2}{Z_2g_m} = \frac{Z_1}{Z_2g_m}$$

if Z_1 is much greater than Z_2. Now suppose that Z_1 is a condenser of reactance $1/\omega C$ and $Z_2 = R$. Then

$$Z = \frac{1/\omega C}{Rg_m} = \frac{1}{\omega C} \times \frac{1}{Rg_m}$$

Since the impedance of a capacitance is equal to $1/\omega C$, the impedance of the circuit looking toward AB is equal to that of a capacitance multiplied by a factor $1/Rg_m$. As g_m is varied, the impedance of this fictitious capacitance varies. This amounts to saying that the circuit described has the characteristics of a variable capacitance. We can vary g_m very simply by inserting modulating voltages in series with the bias voltage (or by using a screen-grid tube and impressing the modulation voltages on one grid and the r-f voltages on another grid). In this manner the capacitance of the circuit can be made to vary directly as the amplitude of the modulation and at a rate which depends upon the frequency of the modulation.

All we have to do now is to shunt this varying capacitance across the tank circuit of an oscillator to vary the frequency generated by this oscillator in accordance with a-f modulating voltages. If we want the modulator tube to act like an inductance, we merely interchange R and C.

Let us look at the reactance tube modulator in another way. In Fig. 3–20, the tube and its input network of R and C is shunted across the tank circuit of the oscillator. If R is large compared to the reactance of C, the r-f current through the circuit will be almost in phase with the voltage across the tank circuit since it is this voltage that is forcing current through R and C. The voltage across C will lag the current through it (as is true of any capacitance). The driving voltage for the tube is the voltage across this condenser. The r-f plate current of the tube will be in phase with the grid voltage, increasing when the grid voltage increases. This plate current, therefore, will lag the voltage across the tank circuit by 90 degrees.

A lagging r-f current taken by the tube will flow through the tank circuit. The same effect can be secured by connecting an

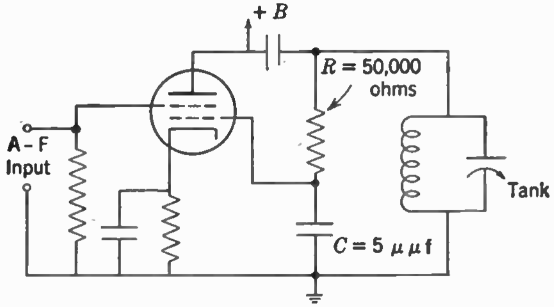


FIG. 3-20. A workable reactance tube modulator which affects the resonant frequency of the tank as though a variable reactance were shunted across the tank.

inductance across the tank circuit, thereby reducing the total inductance of the tank and increasing the frequency generated by the oscillator of which the tank circuit is the frequency-determining part.

If the frequency of the oscillator is increased in a doubler or tripler, as it will be in a practical f-m station, the deviation in frequency due to modulation is multiplied along with the carrier frequency.

Example. An f-m transmitter is to operate on 40 megacycles. The maximum deviation is to be 75 kc. The oscillator frequency is 5 megacycles. How much is this frequency to be shifted to obtain the final desired 75-kc deviation?

Solution. At 40 megacycles the frequency change is 75,000 in 40,000,000. Therefore at 5 megacycles the shift will be one-eighth (40 ÷ 5) as much, or 75,000 ÷ 8 or 9350 cycles.

Frequency stabilization. Since the frequency is to be varied, some means must be provided so that the carrier frequency comes back to its original value when modulation ceases. Otherwise, the average or non-modulated frequency might vary and the receiver could not locate it at the same place on the tuning dial; and of course interference would be created by one station straying into the region of another.

In the reactance tube modulator, a crystal-controlled oscillator cannot be used since the frequency of this oscillator is actually varied by modulation. Therefore an additional oscillator is usually provided which furnishes a reference frequency that is accurately controlled by a quartz plate. Suppose that it differs from the average frequency of the transmitter by 1000 cycles. The two frequencies are allowed to beat and produce the 1000-cycle beat frequency. The beat frequency can be employed to balance a bridge, so that at 1000 cycles nothing happens, but, if the f-m carrier frequency should slowly shift so that a beat frequency of 2000 cycles is produced, the bridge will be unbalanced and any one of several means may be utilized for bringing the f-m transmitter frequency back to its correct value. A practical system is to use automatic frequency control discussed in the chapter on receivers.

F-m receivers. Receivers for f-m signals are superheterodynes of high voltage gain plus two special circuits not ordinarily found in a-m receivers, namely, a **limiter,** and a special detector called a **discriminator.**

In addition it is customary to decrease the higher-frequency audio response at the same rate as that at which the higher audio frequencies are increased at the transmitter. A simple *RC* network which has a response inversely proportional to frequency will produce this "de-emphasis."

The limiter. After the incoming signals are changed in frequency so that they can be amplified in the i-f amplifier (about 4 to 5 megacycles) by means of the same kind of frequency converters that are used in a-m receivers, the signals go through the i-f amplifier and then are fed to a limiter circuit. A limiter circuit is a tube circuit which produces an increasing output for increases in input only up to a certain point; thereafter, the output does not change, even if the input is increased appreciably. The purpose of the limiter is to remove any amplitude variations that may be included in the f-m signals.

For example, note in Fig. 4–20 the appearance of an unmodulated f-m signal. The individual r-f cycles are equidistant and of constant amplitude except in the region of A, where the amplitude varies. This variation in amplitude may be due to noise, static, or some other disturbance not in the original signal.

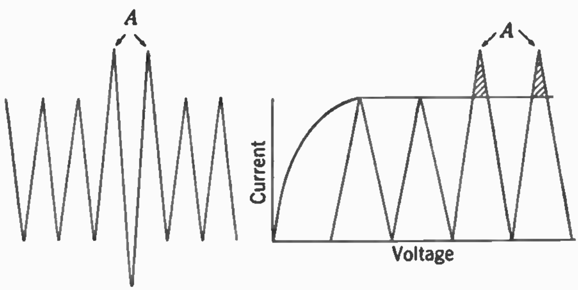


Fig. 4–20. How a variation in amplitude is eliminated by a limiter.

If amplitude variations are permitted to go through the f-m receiver, they will produce audible noise. Let us pass the output of the i-f amplifier through a circuit which has a characteristic like that in Fig. 4–20. Now the variations in amplitude will not get through, since they will be cut off by the limiter. The output of the limiter will not look like the input—there will be distortion—but as long as the distance apart of the individual signals of the incoming f-m wave does not change, the receiver will be quiet. Remember that the f-m system modulates the rate at which the individual waves follow each other (the number per second), and not their relative height.

The purpose of the limiter, therefore, is to chop off any amplitude modulation that may exist in the incoming signal or that may be created within the input circuit, the converter stage, or the i-f amplifier.

What is desired is a tube and circuit so arranged that the input-output characteristic looks like that in Fig. 4–20. This can be secured with a pentode by using low screen and plate voltages. The tube operates in a saturated condition, increases in input grid voltage producing little or no output plate-current increase.

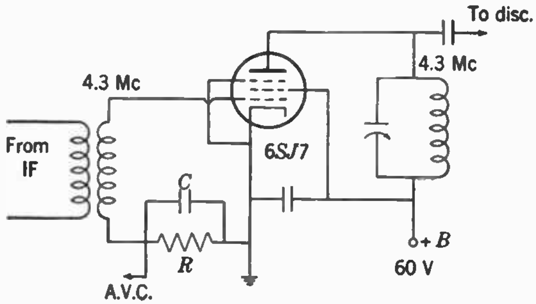


FIG. 5–20. Circuit of simple limiter using low plate and screen voltages.

If the antenna signals are so weak that, after amplification in the i-f system, they do not reach the saturated portion of the limiter characteristic, they will still produce audio tones, but there will be noise in the output since the amplitude variations have not been removed. Thus, an f-m receiver will work with-

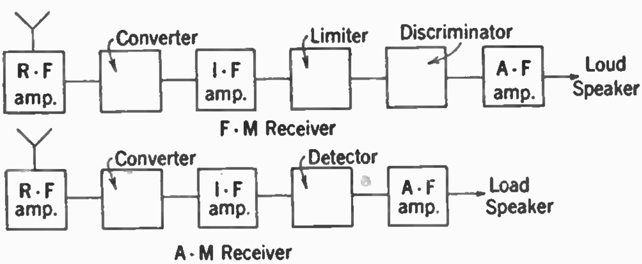


FIG. 6–20. Comparison of f-m and a-m receivers.

out a limiter, but the full effects of frequency modulation will not be secured. The receiver can be wide band and more immune to noise and static than an a-m receiver, but the full effect of the system will not be secured. An f-m receiver, properly designed, manufactured, and operated, and reproducing strong incoming signals, will be almost totally without sound (tube hiss, static, etc.) during moments when the f-m transmitter is not being modulated.

The output of the limiter is a constant-amplitude, variable-frequency signal, the amount the frequency varies from the unmodulated frequency being proportional to the microphone voltages at the transmitter, and the rate at which the frequency varies being proportional to the *frequency* (rate) at which the microphone voltage changes. Now we must convert these variations in frequency back into amplitude variations that will correspond to the amplitude variations in the original microphone voltage. This is done in a discriminator circuit.

The discriminator. One simple manner of performing this function is to use a tuned circuit as shown in Fig. 7–20. Suppose that the output of the limiter is impressed upon the curve at the point A. When the frequency increases, the output of the circuit whose characteristic is shown in Fig. 7–20 will increase; when the frequency decreases, the output in current will decrease. Thus a variation in frequency produces a variation in amplitude. The variations in current strengths may be amplified and fed to the loud-speaker system.

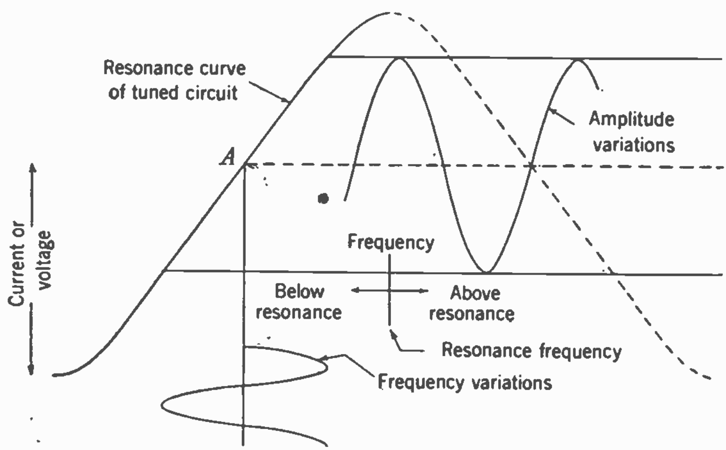


FIG. 7–20. Use of simple resonant circuit to convert frequency variations into amplitude variations.

After the discriminator, the f-m receiver does not differ at all from an a-m receiver except that the a-f amplifier and speaker must be wider in response and freer from noise and distortion.

In practice, a more complex discriminator than that shown in Fig. 7–20 is employed. It may be seen in Fig. 8–20, where will be found a typical i-f transformer primary which is fed signals from the limiter. The secondary of this transformer is split, each half of the winding being connected to a diode rectifier.

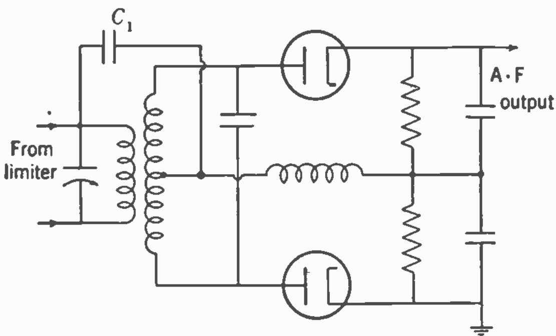


FIG. 8–20. Discriminator circuit for producing audio frequencies from f-m voltages.

R-f current is also fed to the secondary from the limiter through C_1. Primary and secondary circuits are tuned to the center of the i-f pass band.

When the incoming signals are unmodulated, that is when the incoming signal is at the frequency for which the primary and secondary circuits are tuned, the diodes get equal voltages which are opposite in polarity since they are connected to opposite ends of the center-tapped secondary. If, however, the output of the limiter is higher (or lower) in frequency than the unmodulated signal frequency, voltages induced across the secondary are added to (or subtracted from) the voltage secured from the C_1 path. Thus, one diode will have more voltage on it than the other. This is true because of the phase relations existing between the two voltages in the secondary circuit—the induced voltage and the C_1 voltage. The voltage across the diode output load is equal to the difference between the two diode voltages so that, under the unmodulated condition, no audio signal results; but, when one diode gets more input voltage than the other, the output voltage will increase. On the other hand, if the opposite diode gets more input voltage, less audio output will be produced.

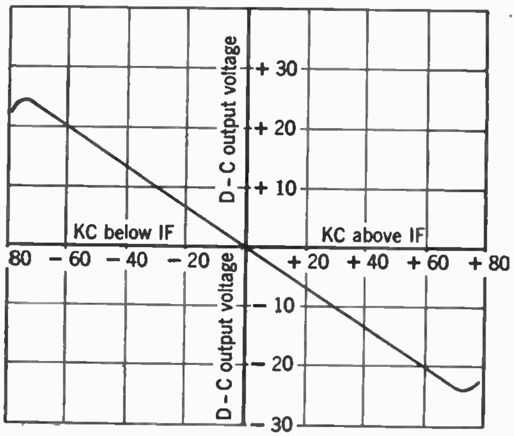


FIG. 9–20. Discriminator characteristic.

The audio amplitude variations will be proportional to the extent the limiter voltages vary in frequency from the unmodulated frequency, and the frequency of the audio output voltages of the discriminator will depend upon the rate at which the limiter output voltage varies in frequency. The over-all effect of the discriminator may be seen in Fig. 9–20, which shows that, over a considerable range, there is a linear relation between frequency variation and output voltage.

The discriminator, therefore, is a converter which changes frequency variations into voltage variations. These voltage changes may be amplified in typical audio systems and applied to the loud speaker.

CHAPTER 21

ULTRA-HIGH-FREQUENCY PHENOMENA

As one proceeds to generate, transmit, and receive radiation of higher and higher frequencies, certain effects which were more or less negligible at the lower frequencies begin to assume new importance.

Transmission. Let us first consider the actual transmission of radiation. The first experiments by Hertz and Marconi were concerned with very short waves. Hertz demonstrated with centimeter-length waves that "wireless" and other forms of electromagnetic radiation—light and heat—had identical characteristics. That is, "radio" energy could be *transmitted* through space, *reflected* just as light waves are reflected; it could be *refracted*, as light rays are refracted (bent) by a lens, and *diffracted*, as light is broken up into interference patterns by a grating.

Marconi, however, wanted to send wireless messages to a great distance, across the ocean, in fact. He reasoned naturally, and correctly, that the antennas must be high in the air. A high antenna meant a long wavelength because transmission lines to connect a short-wave antenna, high in the air, to a transmitter on the ground had not been developed. Marconi's antenna and the ground lead both contributed to the radiation, and the total length governed the wavelength.

Commercial practice was to use long waves for the above reason and for other reasons as well. It was not until much later in the development of the art, after the first World War, that shorter wavelengths were employed. Even then, it was assumed that there was a lower limit to the wavelength (or upper limit to frequency) beyond which it was useless to go, because of the rapid dissipation of the ground wave in short-wave radiation.

Then it was discovered that short waves could be sent long distances, contrary to theory. This discovery, in which American amateur wireless operators played a part, followed the theoretical work of Kennelly and Heaviside and led to the discovery of the ionosphere which makes long-distance, short-wave radio communication possible. All the time, however, communication could have taken place at wavelengths vastly shorter, but no one wanted to communicate for only short distances, or do the many other things now performed by the ultra-short waves.

As one increases the frequency (decreases the wavelength) to the neighborhood of 30 megacycles, the sky wave no longer is a useful part of the communication system because it no longer is returned to earth. The ground wave is very quickly absorbed by ground losses. Thereafter only the direct rays between transmitter and receiver are useful.

Direct ray only useful. Why are not the very high frequencies reflected by the ionosphere? To be useful as a reflector, the individual particles which make up a reflecting surface must be close together compared to the wavelength of the radiation to be reflected. Let us think of a series of pickets in a fence. Let us place them in a pond and then stir up the water and see what happens when the waves strike them. If we make long waves (that is, waves that are far apart), they will strike against the pickets and be reflected back. Very little of the wave will go through the interstices between the pickets. But, if we produce waves of short wavelength, we shall see that they go through the spaces between the pickets and appear on the other side. If the wavelength is small, individual waves are closer together, more of them pass a given point per unit of time, and the frequency is higher. The higher the frequency of the waves (the shorter the wavelength), the closer together the pickets of the fence must be to cause reflection.

Now consider the ionosphere. It is made up of ionized particles with some sort of irregular distribution. But, whatever that distribution is, there is some wavelength of radiation small enough so that most of the radiation escapes through the spaces between particles, and only part is reflected. At still shorter wavelengths (higher frequencies) very little or none of the radiation is reflected to the earth.

On the very high frequencies, say above 30–50 megacycles, there is no sky wave which can be depended upon for communication like that at lower frequencies, say 2 to 30 megacycles. Since the ground wave is highly attenuated or decreased in strength as it moves out from the antenna, only the radiation which shoots straight through space from the transmitting to the receiving antenna can be depended upon for communication.

At the lower end of the very high-frequency region there is some reflection of radiation from the earth, and some bending downward of the radiation by the lower atmosphere, but at still higher frequencies these effects no longer are apparent, and then only the direct ray is useful. The rays that proceed from a lighthouse can be seen at greater distances the higher the lighthouse and the higher the point of observation. Thus a man in the crow's nest of a ship can see a light on shore farther than a man on the deck. Similarly, the higher the transmitting and receiving antennas for these high frequencies, the greater can be the distance between them for reliable communication.

Radiation increases as the frequency increases. Now it is a fact that radiation increases as frequency increases, so that a small amount of radiated energy at high frequencies is just as useful as much more energy at lower frequencies. Why should this be so?

Let us go back to our picture of how radiation occurs. Suppose that an antenna of flexible dimensions is connected to a transmitter of variable frequency. As the frequency is increased, the dimensions of the antenna decrease. The antenna energy or power is in the form of a-c energy or power. When the transmitter is first started up, a magnetic field is produced in the vicinity of the antenna just as a field is set up near an inductance carrying a current. This field moves out from the antenna at the speed of light, 300×10^6 meters per second. At the end of the first half cycle, the field collapses about the antenna, and then moves out again in opposite polarity as the generator of the transmitter produces energy on the second half cycle.

If the generator reverses its polarity quickly enough, the magnetic field cannot totally collapse and change its direction. Some of this field energy is "pushed off" by the oncoming energy of opposite polarity. Now how far can the energy travel and still get back before antenna-current reversal "pushes it off" into space? The distance must depend upon how fast the generator changes its direction of current flow; that is, the distance is a function of the frequency of the currents produced.

Consider a generator which produces 50-cycle current; that is, it produces 50 complete cycles or 100 half cycles per second. Each half cycle, therefore, lasts $\frac{1}{100}$ sec. The magnetic lines of force about a coil carrying 50-cycle current move out in a positive polarity for $\frac{1}{100}$ sec, then reverse their polarity and move out in the negative polarity in the following $\frac{1}{100}$ sec. In other words, the positive magnetic field (we are speaking in very inexact terms now) has $\frac{1}{100}$ sec to reverse and move out in the opposite polarity. This seems to be a comparatively short time —$\frac{1}{100}$ sec.

Let us increase the generator frequency 1000 times to 50 kc. A half cycle lasts only 1/100,000 sec now. At 50 megacycles, a half cycle lasts only 1/100,000,000 sec.

Since all these fields are set up at the rate at which light, heat, and electricity travel, 300×10^6 meters per second, it is not difficult to see how far the field will have traveled in one half cycle. The distance decreases just as fast as the frequency increases. At 50 megacycles the field travels only 3 meters (10 ft) before the reversal in direction of the antenna current has occurred.

Thus, at higher and higher frequencies, the field has less and less time in which to follow the rapid reverses of antenna current, and more and more energy is pushed off into space to radiate to distant parts.

Band width. There are many reasons why the very short waves are so useful. Consider the matter of band width. In the standard broadcast region, stations each require a band width of 10 kc, and they can be placed 10 kc apart. Between 500 and 1600 kc (the approximate total band width) there are 1100 kc. Room exists in any one large region for only 110 stations since modern receivers will not separate stations on adjacent channels in the same locality as the receiver. Now let us consider the region between 100 and 320 megacycles in which the ratio of the highest to the lowest frequency is the same as in the broadcast band. There are 220 megacycles band width, or 220,000 kc, and if each station still requires only 10 kc, we can put 22,000 of them in this region. The number of channels goes up directly as the frequency increases.

There are many services, however, which require wide bands: television and frequency modulation, for example. The only possible place to put them is in the high-frequency regions where there is plenty of space.

BAND-WIDTH REQUIREMENTS

Service	*Band Width*
Continuous-wave telegraphy	Equals telegraph speed in bauds. (1 baud = 0.8 word per minute for a telegraph code having 8 dots or blanks per letter.)
Modulated continuous-wave telegraphy	Add twice the modulation frequency to the above.
Commercial telephony	6 to 8 kc.
Broadcasting	10 to 30 kc.
Television	6 megacycles
Wide-band frequency modulation	200 kc.
Facsimile	Number of picture elements ÷ time of transmission in seconds.

Antennas. Consider now the half-wave dipole antenna, the most efficient radiator. At 15,000 meters (about the longest wavelength used commercially) a half-wave antenna would be 7500 meters or approximately 25,000 ft long.

At 1000 kc (middle of the broadcast band), or 300 meters, the antenna need only be 500 ft long (approximately); at 100 megacycles the antenna need only be 5 ft long. The inherent resistance of the wire will be so little that the 70-ohm radiation resistance represents nearly the total resistance. Nearly all the power put into the antenna will be radiated.

Furthermore, the fact that the antenna can be so much smaller has other advantages. It requires not only less space but also less expensive supporting structures, and it will be so much easier to make antennas that are highly directive by the simple expedient of using numerous half-wave doublets properly spaced from each other and with the currents in them properly phased with respect to each other.

It is no wonder, then, that short-wave radio is so effective on airplanes, in automobiles, or in portable transmitters.

Difficulties at high frequencies. The advantages of the high frequencies are commensurate with the difficulties in the way of generating and receiving them. Let us consider some of the difficulties.

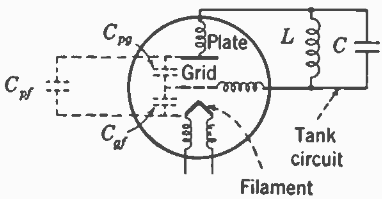


FIG. 1-21. Interelectrode capacitances and inductances inherent in tube.

The frequency to which a circuit responds is proportional to $1/\sqrt{LC}$. To respond to higher and higher frequencies, the product of L and C must get smaller and smaller. Soon we reach the point where the individual values of L and C are merely those which are inherent in the circuit or apparatus—the inductance of the connecting wires shunted by the stray capacitances between the elements, in the sockets, etc. For example, 0.1 μh shunted by 10 $\mu\mu$f will be resonant at approximately 160 megacycles. Now 10 $\mu\mu$f is the input or output capacitance of the average radio tube. Therefore if we connect grid and cathode (the input circuit) or plate and cathode together by a wire having an inductance of 0.1 μh, we shall have an antiresonant circuit tuned to 160 megacycles. This amount of inductance is equal to the self-inductance of a single No. 10 copper wire about 5 in. long.

Even here we have not taken into account the inductance of the grid, cathode, or plate leads inside the tube, or the capacitance between the socket or base terminals. And yet we want to reach still higher frequencies. How can this be done?

One simple way is to reduce the dimensions of all the tube elements, making grid and plate smaller and the spacing of the elements inside the bulb smaller. In this way we shall reduce the capacitances inherent in the tube, say by a factor of 10. But the power output of the tube is now correspondingly reduced. Fortunately, high power is not so necessary at these high frequencies, but lack of power is undoubtedly a limitation.

In working with short waves, every precaution must be taken to see that unnecessary inductance and capacity do not exist.

Impedances at high frequencies. We know that the reactance of a condenser of given capacitance decreases directly as the frequency increases, while the reactance of an inductance increases directly as the frequency increases. Consider a capacitance as small as 10 $\mu\mu$f, the input capacitance of the average radio receiving tube. It has a reactance of only 160 ohms at 100 megacycles. Therefore, any capacitance of this order may represent a very low-impedance path for alternating currents.

On the other hand, suppose that an inductance of only 1 μh exists in part of the circuit where it is not wanted. The voltage drop across this inductance due to its reactance will be 0.00063 volt per microampere of current, at 100 megacycles. Appreciable feedback voltage might be produced in such a wire by permitting currents from two points in an amplifier to flow through this common reactance.

Since even small capacitances have low reactance at the high frequencies, trouble is experienced in making a choke coil, for example. On paper it looks like an inductance, but its distributed capacitance may have less reactance than its inductance at the frequency at which it is to be used, and so we do not have an inductance but actually a capacitance. If the inductive and capacitive reactances are equal, and they will be at some frequency, then our inductance turns out to be an anti-resonant circuit presenting a very high impedance to currents of the resonant frequency.

Condensers, on the other hand, look like capacitive reactances on paper, but the leads to the conducting plates have inductance and at some frequency the inductive reactance will be such that once again we have a parallel tuned circuit. If this condenser is employed to by-pass currents of the given frequency, the usefulness of the component is completely vitiated by the fact that it has both inductance and capacitance.

Distributed constants vs. lumped constants. In low-frequency circuits we have little difficulty in making good coils and condensers to supply inductive and capacitive reactance, but at the higher frequencies these components become so small that we have trouble in making them small enough. We find that we must begin to think of supplying these reactance elements not by building up lumped inductances (coils) or capacitances (condensers) but in other ways.

For example, two parallel wires have a certain capacitance per unit length. Let us say that it is 1 $\mu\mu$f per in. Then, if we want 10 $\mu\mu$f, we simply use 10 in. of the two wires. But, as these two wires also have a certain inductance, a tuned circuit can be made up of them. Since one form of transmission line is made up simply of two parallel wires, we may soon discover that we are using transmission lines, made either of parallel wires or of coaxial cable, as tuning elements. We have avoided the necessity of using lumped inductances and capacitances and instead we are employing capacitance and inductance spread out over a very large area (compared to a coil or condenser), that is, "distributed" inductance and distributed capacitance.

The problem of Q. An engineer has no trouble securing high impedances for tubes to work into and out of at, say, 1 megacycle. He simply makes up an anti-resonant circuit by shunting a condenser by an inductance. The lower the resistance of the circuit, the higher the Q, and the greater the impedance of the circuit. Suppose that the engineer decides to use a condenser of 300 $\mu\mu$f shunted by the proper coil and the combination shunted across the input to an oscillator tube.

The condenser will have a reactance of 530 ohms. The impedance of this tuned circuit at resonance is equal to the reactance of the condenser (or the inductance) multiplied by the Q of the circuit. Thus

$$\text{Impedance} = QX_c = QX_L$$

working with, no trouble develops. If, however, before the electron has arrived at the plate, the plate has become negative with respect to the cathode owing to the reversal of voltages in the plate load circuit, then the electron is not going to arrive at the plate. It will turn about and go back to the cathode.

This may seem far-fetched, but at a frequency of 50 megacycles the time of a half cycle is a hundred millionth second (10^{-8} sec). Can the electron get from cathode to plate in that time?

The time of transit can be calculated as follows. If the electron starts from rest at the cathode (no initial velocity), and if the voltage across the tube is E volts, then the final velocity with which the electron hits the plate is given by

$$v = 0.595 \times 10^8 \sqrt{E}$$

in which E is the potential in volts and v is the velocity in centimeters per second. Since the original velocity is zero and the final velocity is v, the average velocity is $v/2$.

The time required for an electron to travel a given distance is the distance divided by the average velocity or

$$\text{Transit time} = \frac{\text{Distance}}{\text{Average velocity}} = \frac{D}{v/2} = \frac{2D}{v}$$

Suppose that the distance from cathode to plate is 0.3 cm and the plate-to-cathode voltage is 300 volts. What is the transit time?

$$\text{Transit time} = 2D/v = 0.6/v$$

$$v = 0.595 \times 10^8 \sqrt{300} = 0.595 \times 10^8 \times 17.3 \quad \text{cm per sec}$$

$$\text{Transit time} = \frac{0.6}{0.595 \times 10^8 \times 17.3} \cong \frac{1}{17.3 \times 10^8} = \frac{1}{1730} \quad \mu\text{sec}$$

This is the time required to complete one cycle of current at a frequency of 1730 megacycles.

New concepts at high frequencies. Now we must develop some new ideas, or what may seem new to those familiar with the technique at lower frequencies. We have, up to this time, considered an electric current as a completed motion of electrons

If Q is 100

$$\text{Impedance} = 100 \times 530 \text{ or } 53{,}000 \text{ ohms}$$

This is a very respectable impedance, and considerable amplification could be obtained by using it as an amplifier load.

Let us try to do the same thing at 150 megacycles. To secure the same reactance, the tuning capacitance would have to be reduced to 2 μμf, which is less than the input or output capacitance of the acorn type of pentode tube. How are we going to get out of this trouble? Suppose that we try to swamp the input tube capacitance with lumped capacitance, which is what we do at 1 megacycle. Let us increase the tuning capacitance by 30 times, to 60 μμf. Now the reactance is only 18 ohms at 150 megacycles. To get an impedance of 53,000 ohms we shall have to secure a Q of approximately 3000 ($Z = Q\omega L$).

Problem 1–21. Let the reader calculate the r-f resistance of the tuned circuit which will have a Q of 3000 at 150 megacycles if C and L each have a reactance of 18 ohms.

It is a fortunate fact that the two parallel wires mentioned above, used as a tuning element, will have Q values much higher than those of lumped inductances and capacitances; concentric transmission lines will have higher values still, and certain types of "resonators" described later will have still higher Q values. And so the situation is not hopeless; it is only different, and new ideas must be developed, new agencies must be brought into use.

Transit time. Another difficulty at high frequencies comes about from the fact that an electron requires a definite time to cross the space within a tube from cathode to plate. The time required depends, of course, upon the distance between cathode and plate, and upon the speed. Both are under some sort of control. By making the distance to be traveled short and by increasing the voltage across the tube, which increases the speed of the electron, the time of transit can be decreased.

First, let us see what difference the **transit time** makes. As long as the time to get from cathode to plate is short compared to the time of a half cycle of the alternating currents we are positive potential of the grid, but because of the open structure it may not hit a grid wire but may pass through. Now it comes into the field of the plate and will be repelled. If, again, it goes through the grid, it may hit the cathode and cease its motion, but when it comes near the cathode, which is negative with respect to the grid, it may turn about and again go through the grid. Thus the electron executes oscillations in the space between cathode and plate. These motions of the electron induce corresponding changes in electron distribution on the plate, and the corresponding motion of electrons in the plate circuit constitutes an electric current.

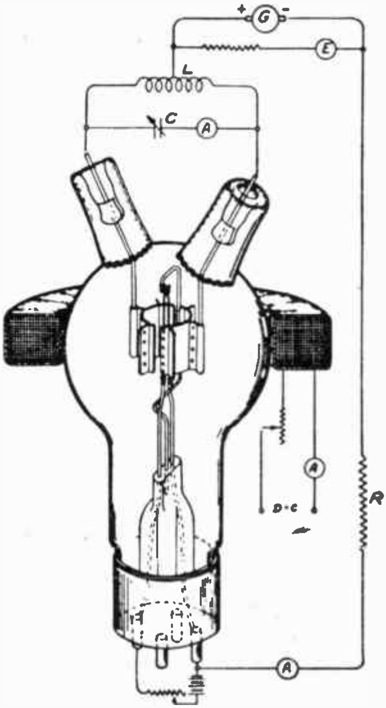


FIG. 2–21. Split-anode magnetron surrounded by electromagnet supplying an external field.

In some of these oscillators the frequency is determined by the tube constants much more than by the external L and C. Wavelengths as low as 10 to 20 cm are obtained by generators of the positive-grid type; they are named after their originators, Barkhausen-Kurz or Gill-Morrell.

Magnetron oscillators. Short-wave energy may be produced in a tube with the general structure shown in Fig. 2–21. This has a cathode and a plate divided into two sections. An external magnetic field is supplied by an electromagnet or by a permanent magnet. If an external alternating voltage is applied between cathode and plate, and if the frequency is such that at about the moment the electron is to hit the plate the external voltage from one place to another. Thus, within the tube, electrons go from cathode to plate and no current is considered to flow until the electron arrives at the plate.

But let us consider a simple condenser. We might think that alternating current flows through it, that is, that electrons pass from one electrode through the dielectric to the other plate. Actually this is not so at all. Electrons, in spite of their small dimensions and their other properties, are not omnipotent. They do not go through the dielectric of the condenser.

What actually happens is that a *motion* of electrons *to* one condenser plate from the outside circuit and *away from* the other plate to the outside circuit causes an effect which we think of as an actual passage of electrons from one end of the circuit to the other. The shift in electrons on the two sets of condenser electrodes creates a strain within the dielectric (and may break it down), but there is no necessity for a single electron actually to move across the gap between electrodes, any more than there is a necessity for electrons to pass between primary and secondary of a transformer for a voltage to appear across the secondary when the primary current is changed.

Any motion of electrons, then, may produce a motion of electrons in a near-by circuit. If, therefore, we can produce an electron motion, say a surging back and forth, we shall have produced a surging or rhythmic current in the external circuit.

Now consider one other point. Suppose that we have a flow of electrons from cathode to anode. Suppose that in the space between we speed up the electrons. This requires an additional supply of energy since more energy is needed to cause the electrons to flow at a high speed than at a low speed. On the other hand, if we extract energy from the electron stream, the speed of the electrons will decrease. Any acceleration or deceleration of the electrons, then, represents a change in energy; and this change in energy can be imparted to the external circuit.

Positive grid oscillators. One of the earliest generators of very high-frequency energy consisted simply of a triode with a positive grid and a negative plate. Consider an electron leaving the cathode. It is accelerated toward the grid because of the reverses, the plate will become less positive (the steady B voltage minus the alternating voltage) and the electron will turn around and head back to the cathode.

In other words, the frequency of the applied voltage is controlled by the transit time of the electron from cathode to plate. The transit time can be changed by adjusting the external steady magnetic field. If there is no field the electrons move straight from cathode to plate, but, if the field is applied in such a way that the electrons tend to move in a curved path toward the plate, the transit time is increased. By proper adjustment of the magnetic field, the electrons can be prevented from reaching the plate, or they can barely graze it before they turn around and leave the region of the plate.

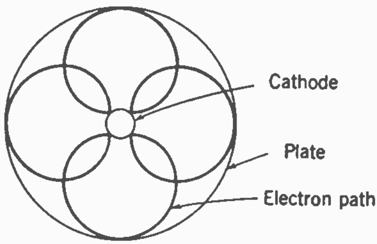


FIG. 3-21. Path of electrons when they just graze the anode.

In practice no external a-c field is applied but a tuned circuit is connected between cathode and plates as shown in the circuit diagram. A property of a resonant circuit is that it tends to produce an alternating voltage across its terminals corresponding to its own natural resonant frequency, no matter what the frequency of the exciting voltage. Similarly, a pendulum will tend to oscillate at its natural frequency even if the impelling force is off frequency, though it will not oscillate so easily under these conditions. The motion of the electrons toward the plate causes a motion of electrons in the tuned circuit, and an alternating voltage appears across this tuned circuit.

In this manner, oscillatory currents will flow in the tuned circuit and energy may be extracted from it by coupling it to a load circuit, such as an antenna. Transit-time magnetrons have been made to oscillate at wavelengths as short as 6 mm; several watts at wavelengths of 9 cm were obtained in 1936, and during the war considerably more power has been developed from magnetrons.

Negative resistance oscillators. If the magnetic field of the magnetron is adjusted so that cut-off occurs (electrons do not reach one of the plates), then the current to the other plate as its voltage is varied will be as shown in Fig. 4–21. Note that the plate current increases with increase of plate voltage up to a

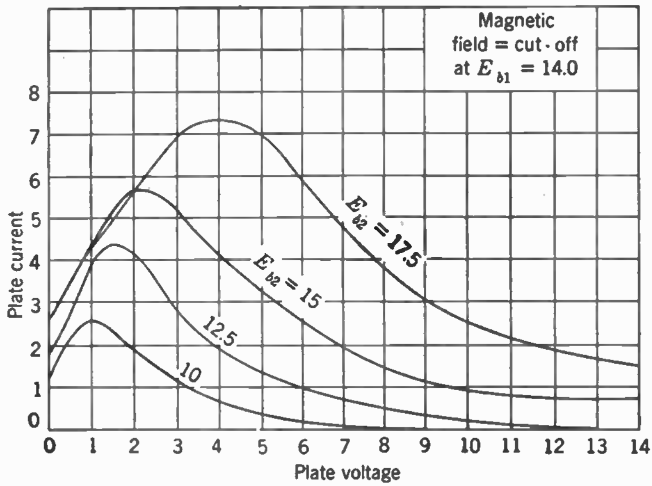


FIG. 4–21. Current to one anode of magnetron when electron flow to other anode is cut off.

certain point and then begins to decrease. After this turn-over point, the slope of the E_b-I_b curve is negative.

Now the resistance of the tube may be determined by dividing the change in plate voltage by the corresponding change in plate current effected by the plate voltage change. Thus

$$r_p = \frac{\Delta E_b}{\Delta I_b} \text{ up to the turn-over point}$$

and

$$r_p = -\frac{\Delta E_b}{\Delta I_b} \text{ after the turn-over point}$$

After the turn-over point the tube is said to have a negative resistance.

Feedback from output to input of an ordinary oscillator may be thought of as supplying negative resistance to the tuned circuit, so that the losses in the tuned circuit due to its positive resistance are overcome. Then the circuit will oscillate.

Similarly, a negative-resistance tube with a characteristic like that shown in Fig. 4–21 will produce oscillatory current in a tuned circuit connected to it. The magnetron can be operated in this manner to produce very short waves.

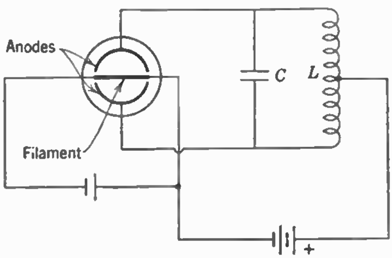


FIG. 5–21. Connections to a split-anode magnetron.

Velocity-modulation tubes. Suppose that a beam of electrons of uniform velocity is required to pass between two grid-like structures (**buncher**) upon which an alternating potential is placed. If the spacing of the grids, the frequency of the alternating potential, and the transit time of the electrons between the

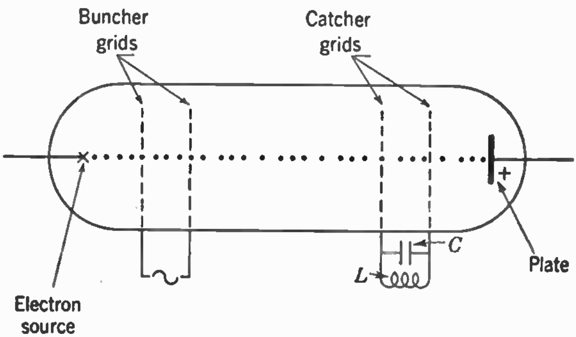


FIG. 6–21. Elements of velocity-modulation tube. Note that electrons flow in bunches after passing the field of buncher grids.

grids are properly arranged, some electrons will be speeded up and some slowed down by the fields between which they pass. In this manner the uniformity of the velocity of the electrons is destroyed, and, instead, the electrons are bunched together into groups. If the bunched electrons pass through another pair of grids, on their way toward the final positive electrode or collector, voltages will be produced upon the "catcher" electrodes.

This is the basis of several types of tubes very useful at centimeter waves, of which the Klystron is typical. In one tube of this general type, 110 watts may be produced at a frequency of 450 megacycles.

To make the tube self-generating it is only necessary to use a feedback system by which some of the energy in the catcher circuit is fed back to the buncher circuit.

Transmission systems. At very high frequencies, the same problem of connecting the generator to the antenna occurs as at

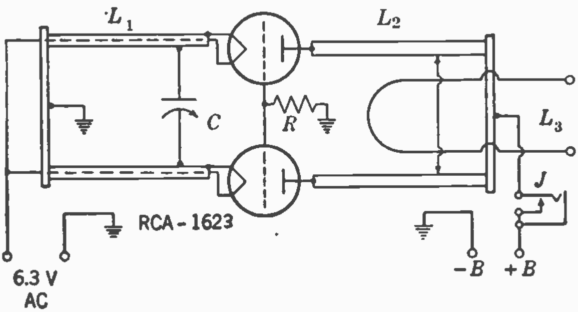


FIG. 7-21. Push-pull 2½-meter, 45-watt oscillator, RCA-1623 tubes; $C = 15\ \mu\mu f$; $R = 1000$ ohms, 5 watts; L_1 and L_2 = 12-inch lengths of ½- and ⅝-inch metal tubing, spaced 1 inch and 1⅛ inches between centers, respectively; plate voltage = 500 for CW; plate current = 200 ma.

lower frequencies. A simple two-wire line may serve both as the tuned circuit and as the line to connect the tube to the antenna. If the wires are fairly close together, say 1/10 wavelength, the wires will not radiate appreciably. At the same time they will have a high Q and so the circuit can be fairly efficient. Greater efficiency may be secured by using a concentric line. Then no radiation occurs until the energy arrives at the antenna, because the outer conductor completely shields the inner wire. As is true at lower frequencies, a quarter-wave line can be used as a matching transformer coupling the tube to its load.

Wave guides. Suppose that we make up a transmission line of two conductors but that for wires we substitute flat parallel plates as shown in Fig. 8–21. Now let us place two other parallel plates alongside the opening between the first two plates. If these two sets of plates are connected as shown, we will merely have placed two sets of conductors in parallel—nothing remark-

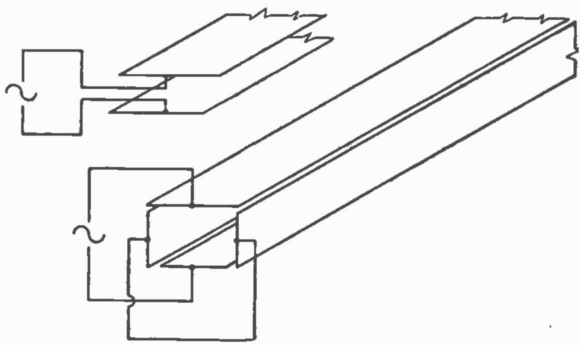


FIG. 8–21. The beginning of a wave guide—two parallel plates. If additional plates are added as in the lower picture, and if all four plates are soldered or welded together (or if a pipe is employed instead) electric energy will flow through the guide without the necessity of a "return" conductor.

able about that. But if the frequency and the dimensions of the conductors are correct, the four plates may be soldered together and we will have a pipe through which high-frequency energy will flow.

Let us look at this matter from a rather different viewpoint. Suppose that we enclose a conductor within another, larger, conductor. We now have a coaxial cable, perhaps of somewhat large dimensions, but a coaxial conductor, nevertheless. If the frequency is increased, the current flows closer and closer to the surface of the inner conductor, and finally, at some frequency, the current no longer flows on the inner conductor at all but only on the inside wall of the outer conductor. The inner wire may now be removed, but energy will still flow through the pipe.

Such conductors of r-f energy are called **wave guides** since they act as guides for electric energy.

Wave guides are useful only at very high frequencies because the physical dimensions must be large compared to the wavelength of the energy to be conducted. At long wavelengths the dimensions would be too great to construct, support, or maintain the guide. But at high frequencies, a wave guide can be an extremely efficient conductor. The losses in metallic pipes of cir-

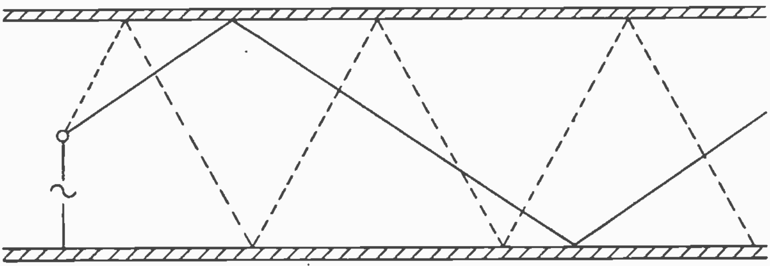


FIG. 9–21. How radiation is reflected by inner walls of a wave guide just as light may be sent down a glass pipe or tube.

cular or rectangular cross sections are of the order of a few decibels per mile. In a typical 5-in. cylindrical copper wave guide, the actual attenuation for 5000-megacycle signals is 10 db per mile, and over a band several thousand megacycles wide the attenuation does not differ much from this value.

It can be seen at once that wave guides provide the high-frequency engineer with an extremely simple and efficient conductor. To many engineers it may seem remarkable that no inner conductor is necessary, that energy will flow along the inside of the pipe and appear at the other end just as though there were a return conductor. It must be remembered, however, that light or heat energy will flow through a pipe or solid cylinder, say of glass or fused quartz, and that radio waves in the 1000-kc region are reflected down from the ionosphere and upward from the earth. The inner walls of the wave guide, therefore, act like reflectors, and as long as they are of high conductivity there will be small losses since the only flow of currents is along the inner surface, which is of considerable area.

It will be found that the wave guide acts like a high-pass filter, providing low-loss transmission above a certain frequency known as the cut-off frequency. This cut-off frequency is determined by the dimensions of the cylinder, which may be rectangular or cylindrical. Thus in Fig. 10–21 the wavelength for cut-off

$$\lambda = 2b$$

where b is the long dimension of the wave guide.

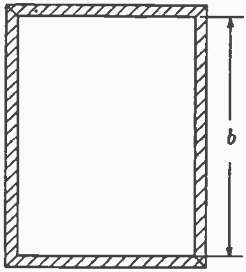


FIG. 10–21. Rectangular wave guide. The longest wavelength that will go through it is equal to $2b$.

For a cylindrical tube the longest wave that will pass down the pipe varies from 1.64 to 3.46 times the radius, depending upon the type or "mode" of propagation.

As in using any transmission line, it is desirable to match the impedance of the input and output devices to that of the line. The lowest impedance of a round pipe is about 250 ohms, but, by varying the small dimension of a rectangular pipe for a given value of the large dimension, the impedance may be made practically any value desired up to 465 ohms. By varying both the dimensions correctly the impedance may be made to have practically any value.

Practical aspects of hollow pipe lines. In a coaxial line, the attenuation or loss of energy per unit length depends upon the high-frequency resistance of the inner wall of the outer conductor and the outer surface of the inner conductor. Since the outer conductor has a large inner surface, most of the losses occur in the central wire or pipe. In a wave guide, there is no inner conductor and these losses do not occur. All the energy dissipation takes place in currents flowing along the inner wall of the tube. If the lowest losses are to be experienced, low-resistance material (copper) should be used. Silver plating will still further reduce the loss. It must be remembered that, at the frequencies under consideration, currents flow on the surface or penetrate the metal to only a very small extent, owing to skin effect.

For 3000-megacycle waves, brass pipe having an outside diameter of 3 in. and a ¹⁄₁₆-in. wall may be used. A rectangular pipe for this frequency (0.1 meter) may be made of 0.081-in. material. Dimensions of 1¼ by 2½ in. are suitable.

Cavity resonators. If a two-wire transmission line of the proper length (¼ wavelength) is short-circuited at the receiving end, it acts like a parallel resonant circuit made up of *L* and *C*.

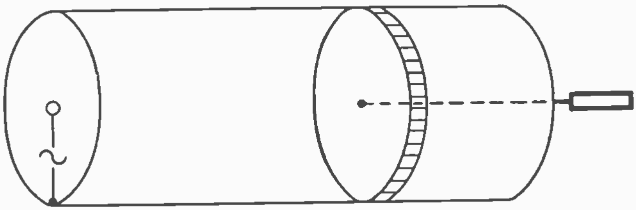


FIG. 11-21. Elements of a cavity resonator. If the frequency of the energy introduced and the position of the plunger are correctly related, standing waves will be set up in the chamber which then acts like an anti-resonant circuit of very high *Q*.

To the generator it presents high impedance. In other words, the line is resonant. Now suppose that we choose the proper length of hollow pipe, close it tightly at one end, and place a movable piston in the other end. In one of these closing surfaces we make a small hole and admit energy of a certain frequency. If, now, the movable piston is adjusted properly it will be found that the pipe acts as a resonant circuit, presenting high impedance to the generator. Such a device is known as a **cavity resonator.** Since it acts like an anti-resonant circuit of extremely high *Q*, it may be looked at as though it were a tuned circuit of extremely low resistance.

A copper sphere having a radius of 17.5 cm, resonating to a wavelength of 40 cm, has a *Q* of about 50,000 and a shunt resistance of about 4 megohms. Since the dimensions of other spheres for use at other wavelengths are proportional to the wavelength, it is a simple matter to design a whole series of resonators, once one has been properly constructed. The value of Q and the shunt resistance vary as the square root of the wavelength.

In such a resonator, standing waves will exist along the sides, and so the device may be considered merely a section of wave

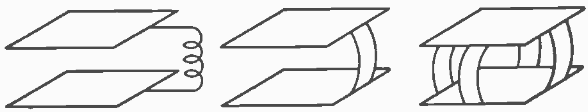


FIG. 12–21. How a cavity resonator may be built up by completely closing in the sides of a two-plate condenser.

guide in which standing waves appear. The dividing line between resonant transmission lines and cavity resonators is rather vague.

Another way of looking at a cavity resonator is as follows. Suppose that we have a condenser made up of two flat metallic plates. If we connect these two plates together by a coil, we shall have a tuned and resonant circuit, whose resonant fre-

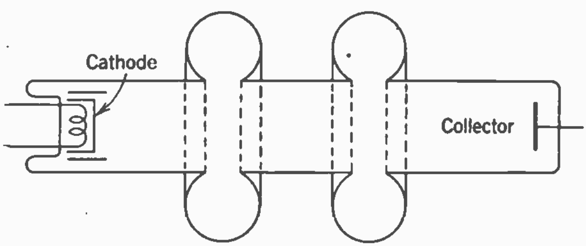


FIG. 13–21. Elements of the Klystron. The circular devices along the tube are cavity resonators.

quency is a function of L and C. Suppose, however, that we use a single turn of wire or even a straight wire or a whole series of connecting wires, we still have capacitance and inductance. In fact, if we add metallic sides to the two condenser plates, the device will still resonate to some frequency. We have constructed a cavity resonator.

The cavity resonator (sometimes called a Rhumbatron) is very useful in the construction of certain types of high-frequency tubes, the Klystron, for example.

Antennas for ultra high frequencies. Any general-type antenna suitable at the lower frequencies may be employed at the high frequencies. In addition, the physical dimensions of short-wave antennas are so small that it becomes possible to employ structures that could not be used at long waves. For example, arrays made up of several antennas, with reflectors to the rear and "directors" in front of the radiating antenna, having highly directional qualities become easily possible.

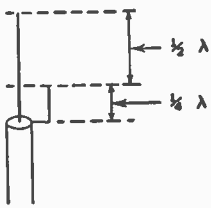


FIG. 14–21. Flagpole antenna using quarter-wave coupling line to connect concentric cable to half-wave antenna.

A very simple type of antenna is made up of a half-wave vertical rod connected to a concentric cable transmission line through a quarter-wave line acting as an impedance adjusting transformer. Such an antenna can be erected on top of a flagpole, high above the earth, so that the transmission range of the station can be large.

If a wave guide is left open at the far end, energy will escape into space and will be radiated as though from an antenna. If, however, the open end of the pipe is flared (Fig. 15–21), the energy will leave the guide in a highly directive pattern. Thus a flared wave guide acts much like a r-f megaphone. These structures are known as "horn" antennas.

The flared wave guide can serve to indicate to the reader how closely related transmission lines and antennas are. A simple two-wire line will radiate some energy. If, however, the wires at the far end are flared out at right angles to each other and are of the correct length (½ wavelength), then we have a transmission line connected to an antenna. This antenna is directional in two oval-shaped lobes. If the angle at the flared end is less than 90 degrees with respect to the line, the directional pattern of the antenna becomes a very elongated oval.

Parabolic reflector. Everyone is familiar with the fact that a source of light placed at the focus of a parabolic reflector can be made to produce a highly directive beam of light. Similarly, a source of r-f energy can be properly located with respect to a parabolic metallic reflector with the result that a highly directive beam of r-f energy is produced. At the short wavelengths we have been discussing, the entire structure, antenna, reflector, and generator, can be easily portable. The angle of radiation may be extremely narrow, and a telescope is often used to line up transmitter and receiver.

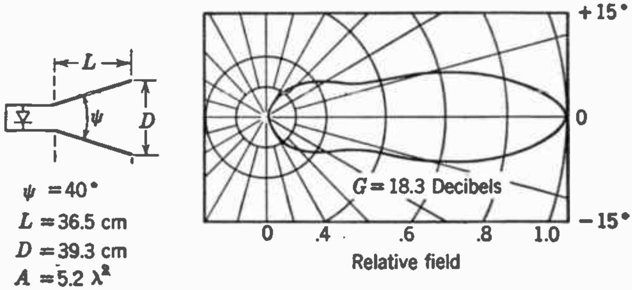


FIG. 15–21. Directional characteristics of a metal horn-type antenna operating on 15.3-cm wavelength. The flared part of the horn is 36.5 cm long; the opening is 39.3 cm.

Much simpler to construct is a simple "corner" reflector made up of two sheets of metal placed at an angle to each other with the antenna at the apex. The angle between the two plates and the location of the antenna control the width of the beam produced.

Of course all these antennas may be used for receiving as well as transmitting, and the gain experienced over a simple dipole may be 10 to 15 times. The effect of a few watts in a dipole may be increased 10 to 15 times when the same power is radiated from a directive antenna. If both transmitting and receiving antennas are directive a gain of 100 times (20 db) is possible.

The future will see many applications of the ultra-high-frequency techniques described briefly here. Centimeter waves are destined to play a most important role in the extension of radio to our communication system. In addition it is highly probable that wave guides, resonators, and all this new apparatus may tend to replace other electronic devices in non-communication applications. As a simple example, it is quite feasible for centimeter-wave apparatus to replace a photoelectric tube and light source for some of the many industrial uses they now serve. Radio waves are not affected by fog, dust, smoke, or rain. They can be guided around corners much better than light. Our imagination is the only limit to the ultimate uses of the high frequencies now produced and received with relative ease.

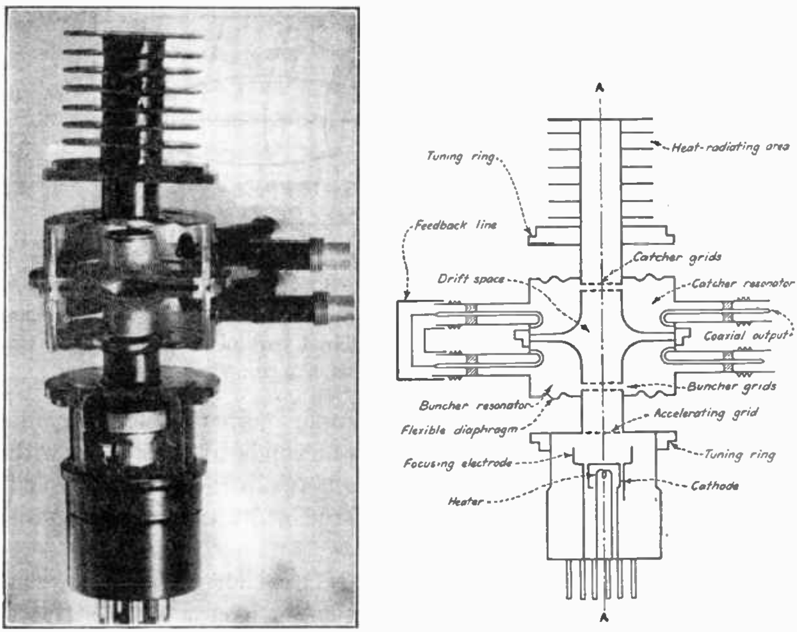


Fig. 16–21. Cutaway and sectional views of the 410-R/2K30 Klystron of Sperry, a tube of the velocity-modulation type which makes possible generation, amplification, and conversion of frequencies in the centimeter wavelength region. The tube is 6 inches long overall.

CHAPTER 22

ELECTRONIC INSTRUMENTS

Vacuum-tube voltmeters. In the study of vacuum-tube detectors as used in radio receivers, it was learned that a d-c component of current was always present in the output of the detector. It was also found that the magnitude of this d-c component depended upon the magnitude of the voltage impressed on the detector. From this, it appears that a d-c milliammeter inserted in the plate circuit of a detector would give an indication of the a-c voltage impressed on the grid circuit. If known voltages were impressed on the grid and the milliammeter scale marked correspondingly in volts, the device would read voltage; in other words, it would be a voltmeter. Also, since it uses a vacuum tube in its operation, it would be a vacuum-tube voltmeter.

A simple vacuum-tube voltmeter as described above operates on the principle that a small voltage applied to the grid of a triode controls a relatively large amount of plate current without drawing any appreciable power from the source of voltage. This type of voltmeter has three distinct advantages over the ordinary D'Arsonval type of d-c instrument, or the iron-vane or electrodynamometer types of instruments used in a-c measurements.

1. The vacuum-tube voltmeter (VTVM) has high sensitivity: a relatively small change in input voltage may produce a relatively large change in the current flowing through the indicating meter. The high sensitivity is due primarily to the amplification taking place in the tube, and it is a very desirable feature when low voltages are to be measured.

2. The meter movement used in a vacuum-tube voltmeter is much more rugged than the movements used in, for example, a D'Arsonval meter having the same sensitivity. This is because

the voltage being read is in effect amplified in the vacuum-tube voltmeter so that more current is available for actuating the indicating meter. For example, a d-c voltmeter having a sensitivity of 20,000 ohms per volt requires a movement which reads full scale with a current of only 50 μa, whereas a vacuum-tube voltmeter of similar or greater sensitivity would probably have a 0–1 ma (1000μa) movement for the indicating instrument. The 0–1 ma movement is much more rugged than the 0–50 μa movement.

3. The vacuum-tube voltmeter has a much higher input impedance than the ordinary type of voltmeter, particularly when the a-c ranges are compared. The advantages of this characteristic are tremendous. With this instrument, it is possible to measure voltages while drawing only a negligible amount of power from the voltage source. Thus it is possible to measure voltages accurately where the use of an ordinary voltmeter would so seriously disturb the circuit by lowering the voltage as to render any measurements unreliable. For example, it is possible to measure the a-c voltage appearing on the grid of a detector tube with a vacuum-tube voltmeter. Such a voltage measurement would be very difficult by any other method.

Other advantages will soon become evident. For example, a well-designed vacuum-tube voltmeter may be calibrated at a low frequency, and this calibration will be accurate at much higher frequencies.

The question may be asked: why do we use any but vacuum-tube voltmeters if they are so superior? The answer is that they are not superior in every respect. In general, they are not so accurate as the standard types of voltmeters. They may require more frequent calibration, owing to changes in circuit or tube constants. They may be sensitive to line voltage variations, if the power to operate them comes from the 110-volt a-c lines. However, these disadvantages are gradually being overcome, and the vacuum-tube voltmeter of today is an accurate and reliable instrument.

Plate-circuit voltmeters. Vacuum-tube voltmeters using plate rectification to convert the a-c input into direct current may be divided into three general classes, depending on the initial grid bias used: (1) the full-wave square-law type is biased midway in the lower curved portion of the I_b-E_c characteristic curve; (2) the half-wave square-law type is biased approximately at cut-off; and (3) the peak type is biased past cut-off.

In the full-wave square-law type (Fig. 1–22), plate current flows during most of the cycle of input voltage. The change in

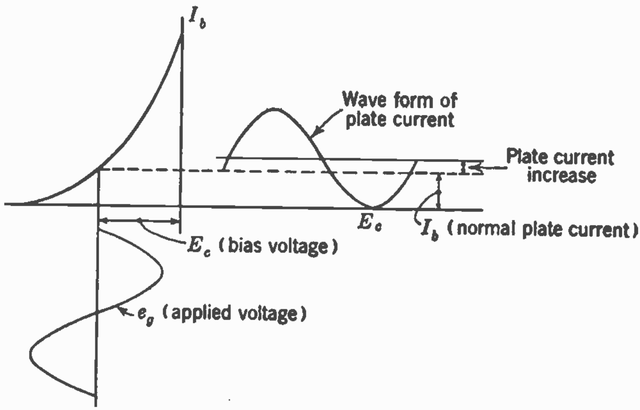


FIG. 1–22. Full-wave square-law vacuum-tube voltmeter.

plate current produced by the rectifying action when an a-c voltage is impressed on the grid is almost exactly proportional to the square of the effective value of the applied voltage. Both halves of the applied voltage wave affect the plate current.

If the grid is biased at cut-off (Fig. 2–22) the negative alternations of the applied voltage have no effect on the plate current, and the change in plate current will be very nearly proportional to the square of the positive alternations. If this type of voltmeter is used to measure an unbalanced voltage wave, it may be necessary to reverse the input terminals for a second reading, then average the two readings to obtain the correct value of voltage.

If the grid is biased well past cut-off (Fig. 3–22) the change in plate current is determined primarily by the peaks of the posi-

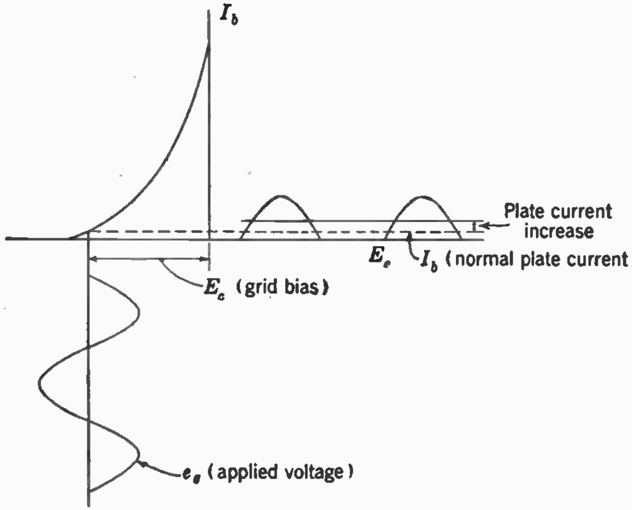


FIG. 2–22. Grid biased at or near cut-off in plate circuit detector used as voltmeter.

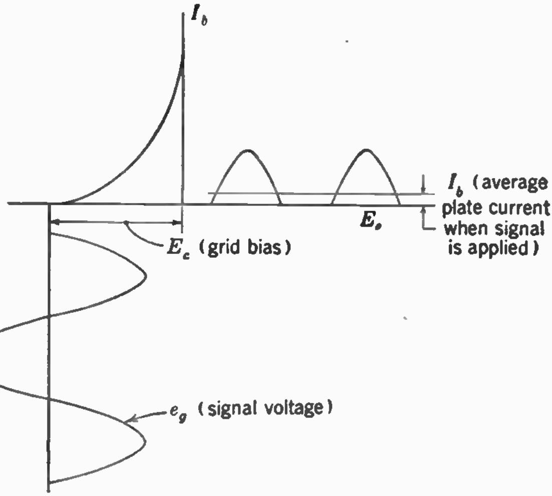


FIG. 3–22. When bias is well beyond cut-off, peaks of input voltage determine plate current.

tive alternations, and the meter becomes a peak-reading voltmeter. The wave form of either alternation has very little effect on the reading, since the voltage reading depends on the peak value only. If the peak values of the negative alternations are different from the positive peaks, a different reading would be obtained by reversing the input terminals.

Bucking-out circuit. There may be some plate current flowing in a vacuum-tube voltmeter even when no signal is applied,

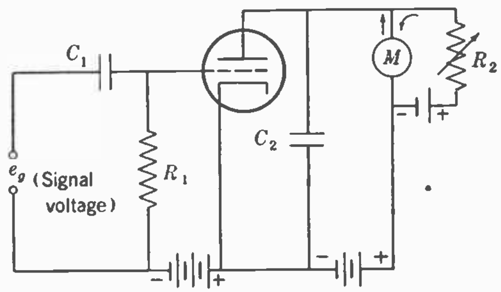


FIG. 4–22. Use of battery and rheostat for bucking out of the plate current meter the steady no-signal current.

unless the grid is biased at or past cut-off. This initial current is of no value and, when flowing through the plate milliammeter, reduces the amount of the scale which may be used for indicating voltages. This disadvantage may be overcome by balancing out the initial plate current so that the plate milliammeter reads zero when zero voltage is applied. A typical circuit for accomplishing this is shown in Fig. 4–22. The variable resistor (R_2) is adjusted so that the current flowing through the auxiliary battery is exactly equal to that flowing through the main battery, with no signal applied to the grid. Both these currents flow through the meter, and since they flow in opposite directions the meter reading will be zero. Various other circuit arrangements may be made to accomplish this same purpose, some not requiring the use of an auxiliary battery.

The **slide-back voltmeter** (Fig. 5–22) indicates positive peaks of the a-c input wave. The grid bias is adjusted so that a small current flows in the plate circuit. The unknown voltage is then applied and the bias readjusted so that the plate current is the same as before. The peak value of the unknown voltage is then equal to the change in grid bias required.

For small values of input voltage, the slide-back vacuum-tube voltmeter is not very accurate and a correction must be applied. However, for larger values of input voltage, this meter gives very accurate readings. The accuracy depends primarily on the ac-

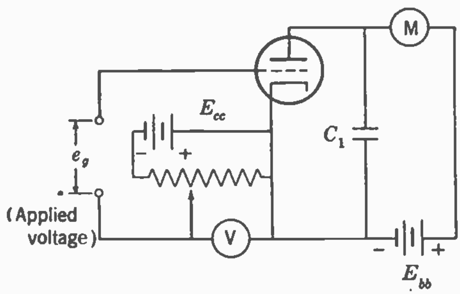


Fig. 5–22. In operation, grid bias is adjusted for some chosen value of plate current. Then the signal to be measured is applied increasing the plate current. Finally the bias is increased to return plate current to its original value.

curacy of the voltmeter employed in the grid circuit and not on changes in circuit constants or tube characteristics.

Diode voltmeters. The discussion so far has dealt only with voltmeters in which the unknown voltage is impressed on the grid of a tube, implying the use of a triode. Diodes may also be advantageous in vacuum-tube voltmeters. The purpose of the diode is to rectify the a-c voltage so that a d-c voltmeter may be used as the indicating instrument. Since the diode does not amplify, the diode vacuum-tube voltmeter is no more sensitive than the voltmeter employed. However, a d-c voltmeter and rectifier for measuring a-c voltages are desirable, since d-c voltmeters in general have much higher sensitivity and higher input impedance than the corresponding a-c voltmeters.

A typical diode vacuum-tube voltmeter is shown in Fig. 6–22. The condenser serves to by-pass the alternating current so that only direct current passes through the meter. The meter-resistor combination may be a high-sensitivity d-c voltmeter, or the meter may be a 0–50 μa movement with the appropriate multiplying resistor. The meter indicates the peak value of the positive half cycle of the input voltage.

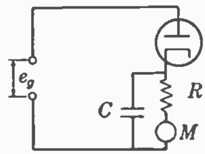


Fig. 6–22. Simple diode voltmeter.

Voltmeter amplifiers. It is possible to combine a diode rectifier with a d-c amplifier to produce a vacuum-tube voltmeter having many desirable characteristics. The General Radio vacuum-tube voltmeter, the circuit of which is shown in Fig. 7–22, is a typical example. The detector and associated resistors are mounted in a probe at the end of a cable, which reduces the input capacitance and stray pickup to a minimum. An acorn tube having a very low input capacitance is the detector. Because of this low input capacitance, the calibration is independent of frequency up to about 50 megacycles.

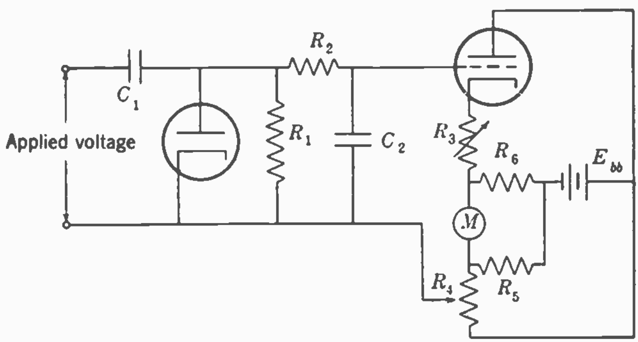


Fig. 7–22. Diode rectifier plus d-c amplifier to enable the use of a more rugged indicating instrument.

Resistor R_3 introduces a large amount of inverse feedback in the d-c amplifier, making the gain nearly independent of tube characteristics. Several ranges of full-scale voltage are also provided by this resistor. The other resistors provide bias and serve to balance out the zero-signal plate current from the meter circuit.

The input impedance of this instrument is about one-fourth the resistance of R_1, which in the commercial instrument is 50 megohms. Thus, the instrument draws a negligible amount of power from the circuit under measurement.

Because of the low detecting efficiency of a diode for inputs of less than 1 volt, this instrument may not have sufficient sensitivity for some applications. If greater sensitivity is desired, a preliminary amplifier, consisting of one or more stages of a-f or r-f amplification, may be added.

Calibration. A well-designed vacuum-tube voltmeter may be calibrated at 60 cycles, and this calibration will be accurate at frequencies up to several megacycles. However, at the higher frequencies, certain factors become prominent which tend to reduce the accuracy of the meter. For instance, the inductance of the input leads may resonate with the input capacity, causing an increase in reading due to the step-up in voltage of the series resonant circuit. For a type 955 acorn tube connected as a diode, and with the shortest possible leads, this effect is negligible up to about 100 megacycles. The transit time of the electrons may also enter into the picture, since some of the electrons leaving the cathode near the end of the cycle may not reach the plate before the cycle reverses. In the tube mentioned above, when reading voltages of the order of 10 volts, the transit time error causes the readings to be about 3 per cent low at 100 megacycles.

The input capacitance of a vacuum-tube voltmeter remains essentially constant regardless of the input frequency. The input impedance thus decreases as the frequency increases, and the decreased input impedance may cause the voltmeter to draw excessive power from the circuit under measurement at higher frequencies.

The effect of wave form on the voltmeter readings has already been discussed briefly. Vacuum-tube voltmeters are ordinarily calibrated with sinusoidal voltages, and any departure from sinusoidal wave form may affect the readings.

Reversing the input terminals of a vacuum-tube voltmeter sometimes changes the reading. This effect is known as "turnover," and it is greatest in peak voltmeters. It is present to a certain extent in the half-wave square-law type but is nonexistent in the full-wave square-law and linear instruments. When turnover is present, averaging the direct and reversed polarity readings will give approximately, but not necessarily, the correct value. The presence of turnover indicates that the amplitude of the positive voltage peaks is different from the amplitude of the negative peaks.

CATHODE-RAY OSCILLOGRAPHS

The cathode-ray oscillograph is probably the most valuable electronic instrument available. It has a multitude of uses, and new ones are being found daily.

Within the cathode-ray tube, a beam of electrons is focused on the fluorescent screen at the end of the tube. When this beam strikes the fluorescent material, visible light is produced. The beam may be deflected by causing the electrons to pass between parallel plates on which potentials are placed or by causing the beam to pass through a magnetic field. These are called respectively electrostatic and electromagnetic deflection. Of the two, electrostatic deflection is used more widely at the present time.

The beam of electrons is produced by an electron gun, consisting of a cathode which supplies the electrons; a grid, negative with respect to the cathode, for controlling the number of electrons in the beam, thereby controlling the intensity; and one or more anodes, positive with respect to the cathode, for accelerating and focusing the electron beam. Two pairs of deflection plates are placed at right angles to each other, so that one will produce a deflection of the beam at right angles to the deflection produced by the other pair. Thus, if an alternating voltage is applied between one pair of plates, the electrons will be attracted to the plate which is positive at the moment, and repelled by the negative plate. The beam is therefore deflected up and down as the voltage varies, and the luminous spot varies correspondingly on the screen. Then any voltage on plates which are at right angles to the first plates will produce horizontal deflections.

Time bases. If a sine wave of voltage is applied to the vertical set of plates, the beam will be deflected up and down from its normal position. A straight line will be seen on the screen. If, however, the beam is moved to the right at the same time that the sine wave of voltage tends to make it move upward, then a pattern which is a sine wave will be traced on the screen. What is desired is a deflecting voltage that will move the beam horizontally at a uniform speed so that at any instant of time, say a particular point in an a-c cycle, the beam will be at a particular

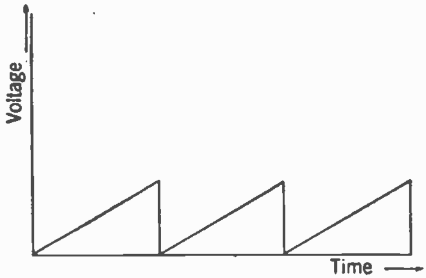


FIG. 8–22. Sawtooth waveform for time base of oscilloscope.

spot on the screen. The sine wave of voltage on the vertical plates will be "spread out" by the horizontal deflecting voltage. This voltage is called a "time base" since it enables the operator to get the time relations of the pattern he is seeing on the screen.

There are many forms of time bases and time-base circuits. In all of them, the beam is swept across the screen (sometimes around the screen in a circle), and then the beam is brought back to its starting point. Sometimes the return trace is blacked out by overbiasing the tube during the return period.

The **sawtooth** waveform is widely used for time bases since it is linear, that is, it sweeps the beam across the screen at uniform velocity, then returns it back to the starting point very quickly.

The ideal sawtooth wave form can be approached quite closely by means of a simple relaxation oscillator, the circuit of which is shown in Fig. 9–22. Direct current flows through the resistor, charging up the condenser until the voltage across the neon bulb is sufficient to cause the bulb to "break down," discharging the condenser. The charge and discharge cycle is then repeated at a rate depending on the size of the resistor and the condenser.

As the condenser in the above oscillator charges exponentially rather than linearly, the rate of charging is not constant. The charging rate may be made constant by replacing the resistor with a constant-current device, such as a pentode vacuum tube. Then, since the charging current must be constant, the rate of charge of the condenser must be uniform, and the voltage across the condenser will increase at a uniform rate. A type 884 or 885

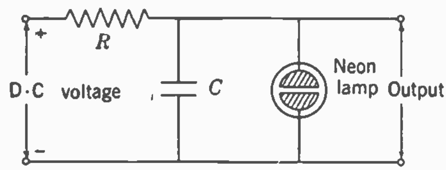


FIG. 9–22. Sawtooth generator circuit.

grid-controlled gaseous discharge tube is generally used in place of the neon bulb, as the frequency of the sweep may be controlled to a certain extent by the application of a small voltage of the proper frequency to the grid. The frequency of the neon bulb oscillator cannot be controlled in this manner.

A circuit of a sweep frequency oscillator using a constant-current pentode is shown in Fig. 10–22. Such a circuit is capable of producing a wave form very close to the ideal sawtooth wave.

Quite frequently the amplitude of the voltage to be studied is so low that, if impressed directly on the oscillograph plates, it would have no noticeable effect on the electron beam. It then becomes necessary to amplify the voltage. The amplifier used for the purpose must be especially designed to have a fairly high gain and to amplify uniformly a very wide band of frequencies, in fact, a band which includes all frequencies for which the oscillograph may be used. Such an amplifier incorporates a high-gain resistance-coupled pentode in its make-up.

It is desirable to be able to control the sweep frequency amplitude, and occasionally an external signal is employed in place of the sweep frequency. For these reasons, an amplifier is commonly provided for the horizontal as well as the vertical plates.

Oscillograph controls. Most oscillographs are equipped with a minimum of eight controls, and it is necessary that the functions of these controls be understood before the instrument can be used intelligently.

The horizontal and vertical amplifiers are both provided with gain controls to adjust the amount of amplification to the de-

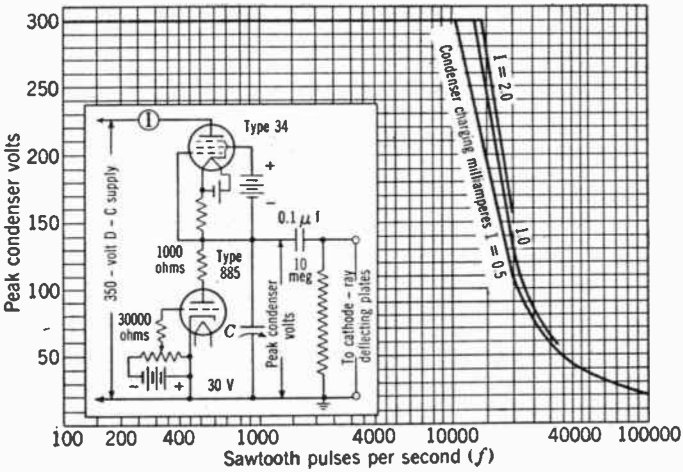


Fig. 10–22. Complete sweep voltage oscillator circuit and its characteristic. (From RCA.)

sired value. Focusing and intensity controls are provided to produce a sharply defined spot or line of the correct brilliancy. The focusing control determines the voltage applied to the first anode, and the intensity control determines the amount of negative bias on the control grid. A fine and a coarse sweep frequency control are usually provided. The coarse control switches in various sizes of condensers which are charged up and discharged through the grid-controlled discharge tube, while the fine control adjusts the amount of resistance through which the direct current must flow to charge the condenser. Spot locating controls adjust the steady d-c voltage on the deflecting plates so that the rest point of the spot may be adjusted to the desired position on the screen. A synchronizing control may be provided for feeding part of the input voltage to the grid of the sweep frequency oscillator tube, so that the sweep frequency may be "locked in" with the input voltage.

Other controls, such as "Z axis" control, may be added on the larger and more expensive instruments. Each instrument should be carefully studied before it is used so that the function of every control is properly understood.

Oscillograph applications. One of the principal uses of the cathode-ray oscillograph is to study the wave form of alternating voltages. If the sweep circuit is sufficiently linear, the picture of the voltage wave appearing on the screen will be a very good indication of the exact form of the wave.

The oscillograph may be used as a voltmeter, since the deflection on the screen is directly proportional to the impressed voltage. By properly connecting the oscillograph to the output of an amplitude-modulated transmitter, the trace on the screen will indicate the percentage modulation of the transmitter. It will indicate phase shift in reactive circuits and, by means of Lissajous figures, may be used to compare the frequency of two oscillators, or to calibrate an oscillator by means of a standard frequency. These and a myriad other applications make the cathode-ray oscillograph an extremely valuable instrument.

Electronic switch. A device frequently used with cathode-ray oscilloscopes is the electronic switch, which makes possible the simultaneous observation of two voltages on a single screen. It consists essentially of two amplifiers, alternately biased at cut-off, with both amplifiers connected to one set of oscillograph plates. The two voltages to be compared are impressed on the amplifiers, and the trace on the screen corresponds first to one voltage and then the other, depending on which amplifier is in operation. If each amplifier operates half the time, then each voltage will appear on the screen half the time.

Bias for the cut-off voltage on the amplifier tubes is obtained from a square-wave generator, usually built around two gaseous discharge tubes. Each amplifier tube is biased at cut-off for one-half the time, and the actual period of cut-off depends on the frequency of the oscillator. This device is also very valuable for determining the phase shift between two voltages.

SIGNAL GENERATORS

The r-f signal generator is an instrument which provides a convenient source of modulated or unmodulated r-f voltage. The frequency and amplitude are variable over very wide limits, adding to the usefulness of the device. It is, in fact, a miniature radio transmitter with an accurately measured output.

The signal generator consists essentially of a very completely shielded oscillator which can be modulated, together with an attenuator for varying the output. Some signal generators are provided with output meters, so that the magnitude of the output voltage may be read directly on the meter. The output may be continuously variable from as low as 1 μv to as high as 0.5 volt. Usually the percentage modulation is also variable over wide limits. The frequency range may cover the r-f spectrum from 10 meters to 550 meters.

The a-f oscillator provides a convenient source of a-f voltage. As in the r-f signal generator, the frequency and output are continuously variable over rather wide limits. The design is usually such that the output stays fairly constant for a particular setting of the output control, even though the frequency is varied. The output is very nearly a sine wave.

The so-called beat-frequency oscillator consists of two r-f oscillators operating at different frequencies. When these two frequencies are mixed, the difference frequency or "beat" frequency is produced. In this type of oscillator, the frequency of one r-f oscillator is fixed and that of the other is variable. The usual practice is to make the frequency of the fixed oscillator of the order of five times the maximum beat frequency desired.

Many electronic instruments other than those mentioned are in use today, but their description is beyond the scope of this book. These include the *Q* meter, harmonic wave analyzer, signal tracer, audolyzer, and chanalyst, to name only a few.

CHAPTER 23

RADAR

Radar was one of the supreme contributions of electronics engineers to the war. By its means, it became possible accurately and quickly to determine the direction of an enemy airplane and at the same time to measure its elevation and its distance. The word radar was invented from the terms "radio direction and ranging," and it means simply the determination of direction and range of any object which will reflect back to the transmitter a portion of the energy received from the transmitter.

Basically, radar depends upon three simple phenomena, all well known in the radio art before the war. These are: (1) the ability to aim a narrow band or cone of radio energy by means of directive antennas; (2) the fact that radio waves travel at a known velocity in space, namely 300,000,000 meters or 186,000 miles per second; and (3) the fact that radio energy striking an object such as an airplane induces currents in the object. The currents cause a reradiated pulse of energy to be sent out or reflected. These facts were known to Hertz, the first and third being demonstrated by him and the second being part of the basic theory developed by Maxwell. Thus the basis for radar existed and was known before 1900. Only in recent years, however, have means been available for measuring the time required for a radio wave to go out to a reflecting object and be returned to the sender. This portion of a radar system had to wait for modern techniques.

A simple radar system consists of a transmitter which sends out periodically pulses or short bursts of energy, an antenna array designed to concentrate the energy into as narrow a beam as possible, a similar receiving antenna, and a means of measuring the time of transit of the radio waves. The transmitting and receiving antennas are rotated in synchronism. The direction of the antennas when a return echo is picked up indicates the direction of the target.

Since the received echo is weak compared to the transmitted signal, it is necessary that the transmitter send out its energy in short pulses with a comparatively long period between pulses.

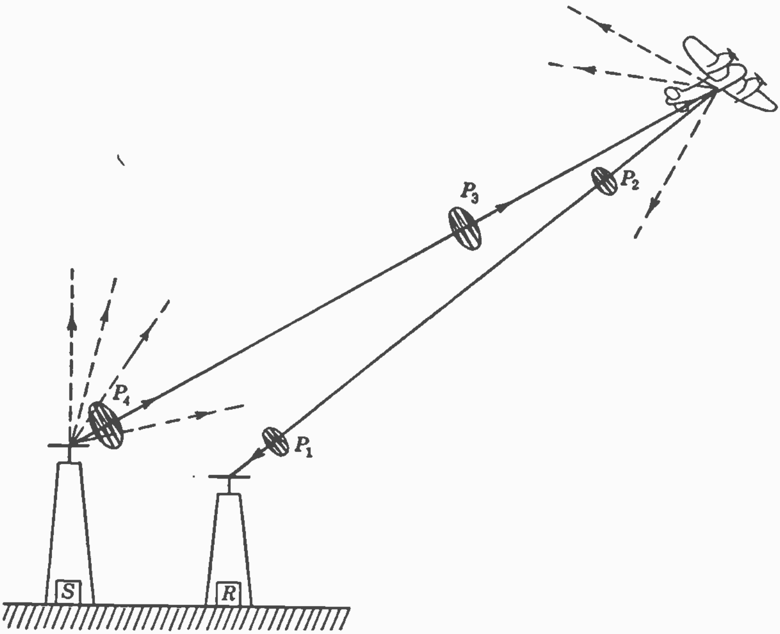


FIG. 1–23. Energy sent out in pulses from transmitter S is reflected in pulses from the plane to the receiver R. (From *Wireless World*.)

During this quiet period the receiver is enabled to pick up any returning signal. If the transmitter were on all the time, the receiver would be overwhelmed by the local signal and would be unable to pick up the weak echo.

Since the transmitter is energized at intervals and only for a very short time (a microsecond or so), the peak power may be very great without the average power being above the tolerance of the tubes. This peak power may be of the order of hundreds of kilowatts.

The distance or range of the target is measured by cathode-ray apparatus. When the pulse is sent out, a signal is sent to the horizontal plates of the cathode-ray tube. This starts a sweep voltage across the tube. When the echo is returned, a signal is sent to the vertical plates so that a slight kick (called a "pip") upwards from the horizontal trace on the screen is produced. If the electrons

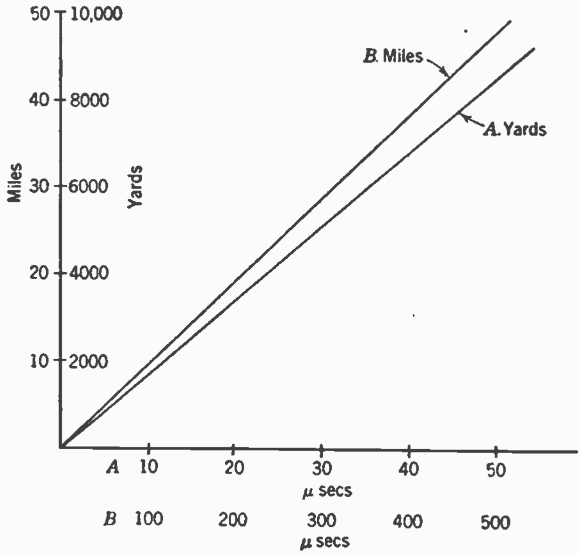


FIG. 2–23. Relation between transit time from radar to plane to radar and distance between plane and radar.

are swept across the screen at a constant speed, the place on the screen where the pip occurs is a measure of the time required for the signal to go out from the transmitter to the target and to be returned by the target.

For example, radio waves travel 186,000 miles per second or 0.186 mile per microsecond, out and back, or a one-way trip of 0.093 mile (491 feet) per microsecond. Suppose, now, that the electrons are swept horizontally across the screen a distance of 4 inches in 400 microseconds, or 1 inch for every 100 microseconds. If, therefore, a pip occurs 1 inch away from the starting point of the horizontal trace, the elapsed time is 100 microseconds between transmission of the signal and reception of the echo. In 100 microseconds the signal went out to and returned from a distance of 100 × 0.093 mile or 9.3 miles. Thus the reflecting target was 9.3 miles away.

If echoes are obtained when the antenna is pointed at a bearing of 185 degrees with respect to true north, then the target lies 9.3

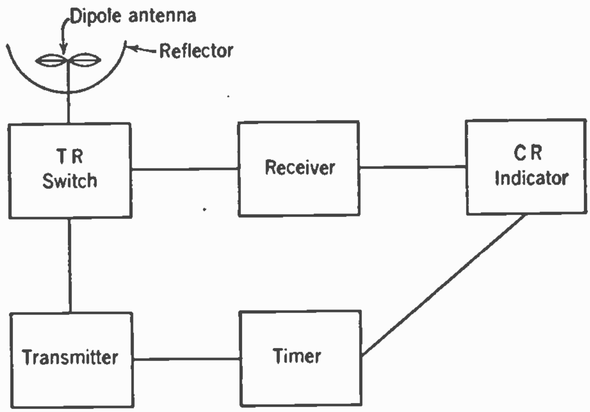


FIG. 3–23. Block diagram of typical radar. *TR* switch throws antenna from transmitter to receiver.

miles away on the 185-degree line. Furthermore the elevation above earth can be determined by tilting the antennas away from the horizontal until echoes are recieved. The angle of the antenna with the horizon is, then, a measure of the height of the target, which can be determined by elementary trigonometry.

The radar transmitter. Since highly directional antennas can be constructed within reasonable dimensions when the wavelength of the transmitted energy is short, radar systems employ the very high frequencies, ranging from hundreds to thousands of megacycles. The transmitter, therefore, is a short-wave transmitter. It is not continually on the air like a broadcast transmitter. It resembles a code transmitter more than a voice transmitter, and a short-wave code station could be employed as the radar transmitter by merely making the dots very short in duration and at a regular rate.

Means must be provided for furnishing requisite instantaneous power during the pulses, and there must be some sort of keying system which allows the transmitter to produce high-frequency energy at regular intervals and for the required length of time during the individual pulses. This keying system must be accurate; if the timing of the pulses gets out of order, the returning echoes may be obscured by the more powerful transmitter pulse signals. The timing mechanism must be synchronized with the indicator in the receiver so that the trace of the cathode-ray tube starts across the screen when the pulse leaves the transmitter.

Since the transmitter is essentially a short-wave station, any of the well-known methods of producing short-wave signals may be employed. Thus a magnetron, the Klystron, or the positive grid oscillator may be utilized at the very high frequencies, and at the lower frequencies an ordinary negatively biased oscillator tube may be employed.

Transmitter power. The average amount of power required depends upon the fraction of the time in each second that the power is on and the amount of power taken during this time. In this situation, a radar transmitter resembles an electron-tube-controlled welding machine where high-powered pulses of energy are supplied to the weld at regular intervals, each pulse lasting only a fraction of a second.

As an example let us suppose that a 200-watt transmitter is to be keyed with pulses lasting 2 microseconds each, there being 500 pulses per second. What fraction of a second of time is the transmitter actually consuming power from the power source? How much power may be packed into each pulse without overloading the transmitter or the power supply?

If there are 500 pulses each 2 microseconds wide, the total time the power is supplied to the transmitter is $500 \times 2 \times 10^{-6}$ or 1000 microseconds, and, since there are 1 million microseconds in each second, the transmitter is "on" only $\frac{1}{1000}$ of the total time. During the pulses, therefore, the power may be as high as 200 kilowatts (200×1000 watts).

As in welding technique, these factors can be related as follows. Let the term "duty cycle" taken from welding terminology be the relation between the width of the pulse and the time between pulses:

$$\text{Duty cycle} = \frac{\text{Pulse width}}{\text{Time between pulses}}$$

If there are 500 regularly spaced pulses (cycles) per second, the time between pulses is, obviously, $\frac{1}{500}$ second or 2000 micro-

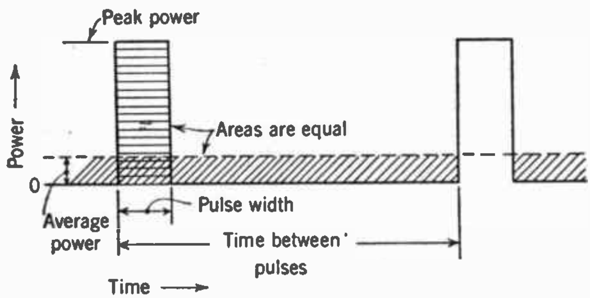


FIG. 4–23. High peak power does not mean high average power. In radar sets the time between pulses may be quite long compared to pulse time.

seconds. If the pulse is on for 2 microseconds, the duty cycle is $\frac{2}{2000} = 0.001$.

The duty cycle is also the ratio between the average power and the peak power. Therefore,

$$\frac{\text{Average power}}{\text{Peak power}} = \frac{\text{Pulse width}}{\text{Time between pulses}}$$

In the example above, where the transmitter had a normal rating of 200 watts, $\frac{200}{\text{peak power}} = 0.001$ or the peak power is 200,000 watts or 200 kilowatts.

Since the amount of energy reflected by a small object at some distance from the transmitter is very small, the radar system is inherently inefficient, and it is of advantage to endow the pulses with high instantaneous power.

Pulse frequency. How many pulses per second should be sent out? Let us consider the example above, where 500 pulses are sent out per second. How far away can the target be to reflect one pulse back to the receiver before the next pulse is emitted? The pulses are 2000 microseconds apart in time. The energy will travel a one-way distance of 0.093 × 2000 or 186 miles. This will

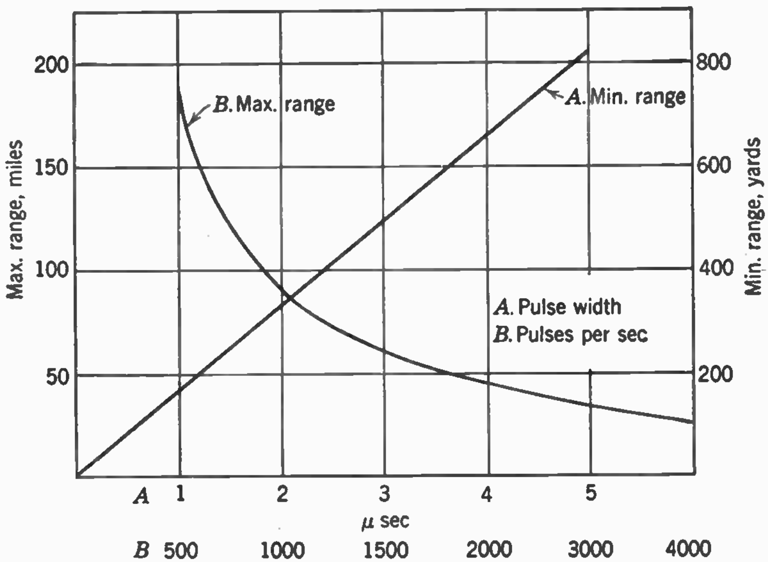


FIG. 5-23. The minimum range increases as pulse width in microseconds increases; the maximum range, however, decreases with wider pulses.

be the *maximum* range of a radar transmitter using a pulse frequency of 500 per second.

How long should the pulse last? If it lasts too long, the returned signal will arrive before the transmitter signal is turned off. If the pulse lasts 2 microseconds, the signal will travel (one way) 2 × 0.093 mile or 0.186 mile, or 328 yards. This is the *minimum* distance over which a radar using a 2-microsecond pulse would be useful.

The pulse *rate* determines the maximum effective distance, a higher rate giving the signals less time to go out and return and thus lowering the maximum distance; the pulse *width* governs the minimum distance over which echoes can be received successfully. A wide pulse may return more energy from the distant target but will increase the minimum distance over which the radar is useful. Numerous pulses per second increase the possibility that several pulses may hit the target and be returned, with the result that a brighter trace on the cathode-ray tube will be secured.

In practice, the pulse repetition rate is much greater than the rate at which the antenna is to rotate about the horizon or about the sector the radar is exploring. In this manner, there is sufficient time for several pulses to go out to the target and to return to the radar site before the antenna has moved on to a new location. If proper pulse control is attained, the received echoes will superimpose upon one another so that a visible "pip" may be produced.

The lowest effective pulse rate, therefore, is determined by the maximum range desired, by the rate at which the antenna is rotated, and by the persistence characteristics of the cathode-ray tube screen.

Modulator. Since the transmitter is to be turned on and off at regular intervals and for short periods of time, some means must be provided for modulating the power to the tube, in this case modulating it completely and abruptly. This modulation may be accomplished by the oscillator tube itself, or additional tubes may be employed for the purpose.

Radar receiver. Since the energy returned to the vicinity of the transmitter by a target may be exceedingly minute, the receiver must have the maximum possible sensitivity and considerable stability. The incoming energy is lowered in frequency by conventional superheterodyne technique and is amplified as much as necessary prior to the detection process in which the pulse envelop is secured free from the r-f and i-f voltages. The pulse voltage may be increased further in a video or wide-band amplifier and applied to the vertical plates of the cathode-ray tube.

The indicator. As is true of many modern applications of electronics, the cathode-ray tube is an important element in a radar system. It serves as a visual indicator of range (calibrated in yards or miles), of height (calibrated in degrees from the horizontal), or bearing with respect to north (calibrated in degrees). In each case the movement of the electrons across the cathode-ray screen in a horizontal direction is synchronized with the movement of the antenna in a horizontal direction for bearing or in a vertical

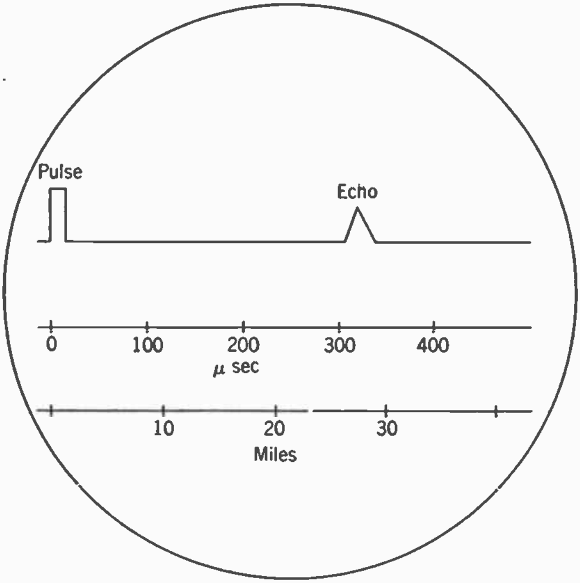


FIG. 6–23. How echoes appear on radar indicator, which may be calibrated in units of time or distance.

direction for height. The screen can be used to indicate bearing or vertical angle, or it can be used merely to indicate when a response is received from the distant target, a compass card synchronized with the antenna movements indicating the actual bearing.

If the radar is to be used on shipboard only to locate ships at sea, the dimension of height may be eliminated since all that is needed is the distance and the bearing of the target. If the radar is to be used for gun control, much greater accuracy is required than if the system is merely to indicate the vicinity and range of a target. When employed for fire control, the movements of the antennas are synchronized with the movements of the guns or searchlights.

Antennas. If a long-distance detection system is required, a comparatively long wavelength is necessary, since the microwaves are more or less limited by line-of-sight transmission and since it is not possible to get as much power from the tubes in the microwave region as at lower frequencies.

The long-wave system may employ half-wave dipoles with reflectors and directors to concentrate the energy into a narrow cone. A very common antenna is quite similar to the Jagi structure used by amateurs in their 5- and 10-meter bands. If the wavelength is of the order of 1.5 meters (200 megacycles), 16 dipoles with a similar set of 16 dipole reflectors, all properly spaced and connected, will produce a power gain of 400 times. Thus, the 100-kilowatt signals described above become in effect 400×100 or 40,000-kilowatt signals. The directive property of such an antenna may be indicated by stating that it is 30 degrees wide at a point where the radiated energy is one-half that radiated along the axis of the antenna. Ordinary transmission lines conduct the power from transmitter to antenna.

In the microwave region, the antenna may be very small and may be located in parabolic or spherical reflector structures with considerable power gain. Wave guides serve as connecting links between antenna and transmitter or receiver.

Thus it will be seen that a radar system employs known techniques taken from the radio, television, and electronic-welding arts. The frequencies used, the usage to which the system is put, the power outputs, the kinds of tubes employed, and the fine details of power generation and control are still military secrets. But, basically, there is nothing mysterious about a radar system to any one having an average knowledge of present-day radio technique.

Advantage of radar over audio systems. Before radar reached its present state of development, detection of the presence of airplanes in time to do anything about it was most difficult. Numerous audio systems were developed, one of the commonest being a large assembly of directional horns or inverted megaphones which collected sound emanating from the airplane engine. These sounds were amplified and delivered to a pair of headphones worn by the operator. The horns were aimed until the maximum loudness of the sound was secured. This indicated the direction from which the sound was arriving.

Sound travels at a very slow rate compared to radio waves. Thus, sound in air has a velocity of approximately 1100 feet per second. If the vicinity of the sound detector was relatively quiet, an airplane could be detected as far away as 25 miles. At 1100 feet per second, the time required for the sound to arrive at the detector from the plane was of the order of 2 minutes. Thus the audio sound detector indicated merely where the airplane had been and not where it was at the time the sound of its engines arrived at the detector site. The sound detector, therefore, was not of much use against modern high-speed planes since it did not give sufficient warning for defensive planes to get into the air.

On the other hand, radar signals travel at approximately 1000 feet per microsecond or nearly a million times as fast as sound waves. For this reason radar reveals much more accurately than sound methods the actual location of airplanes and gives defense crews early warning of the approach of enemy aircraft.

INDEX