_____

I put a lot of time into producing these files which is why you are met with this page when you open the file.

In order to generate this file, I need to scan the pages, split the double pages and remove any edge marks such as punch holes, clean up the pages, set the relevant pages to be all the same size and alignment. I then run Omnipage (OCR) to generate the searchable text and then generate the pdf file.

Hopefully after all that, I end up with a presentable file. If you find missing pages, pages in the wrong order, anything else wrong with the file or simply want to make a comment, please drop me a line (see above).

It is my hope that you find the file of use to you personally – I know that I would have liked to have found some of these files years ago – they would have saved me a lot of time !

Colin Hinson
In the village of Blunham, Bedfordshire.

RESTRICTED

A.P. 3302

STANDARD TECHNICAL TRAINING NOTES

FOR THE

RADIO ENGINEERING TRADE GROUP

(FITTERS)

PART 2

COMMUNICATIONS

MINISTRY OF DEFENCE
FOR USE IN THE ROYAL AIR FORCE
MAY 1964

FOREWORD

'hese notes are issued to assist airmen and apprentices under training as Fitters in the
ɔ Engineering Trade Group. Fitters in this trade group 'require a thorough knowledge of
lectrical and radio principles, and the elementary mathematics, appropriate to the theory
ᴣ specified equipment' (see A.P. 3282A, Vol. 2). It is with the intention of helping to attain
tandard that these notes are written. They are not intended to form a complete text-book,
re to be used as required in conjunction with lessons and demonstrations given at the radio
ɔls. They may also be used to assist airmen on continuation training at other RAF stations.

'he notes, which are based on appropriate syllabuses of training are subdivided as follows:—

Part 1A: Electrical and Radio Fundamentals.

This deals with the principles of electricity, electronics and radio at a level suitable
for the upper technician ranks and for technician apprentices. Because of its bulk
Part 1A has been split into three separate books: Book 1 covers basic electricity;
Book 2, basic electronics; and Book 3, basic radio.

Part 1B: Basic Electricity and Radio.

This deals with the principles of electricity, electronics and radio at a level suitable
for the lower technician ranks and for craft apprentices.

Part 2: Communications.

This deals with the applications of the principles covered in Parts 1A and 1B to
communication systems and is intended to be used as required by all fitters in the Radio
Engineering Trade Group.

Part 3: Radar.

This deals with the applications of the principles covered in Parts 1A and 1B to radar
and is intended to be used as required by all fitters in the Radio Engineering Trade
Group.

n general, fitters employed on communications equipment will be interested mainly in
1A or 1B and Part 2 of these notes. Similarly, radar fitters will be concerned mainly with
1A or 1B and Part 3. However, it is difficult to draw a firm dividing line between the
ⁱledge required by fitters engaged in communications and that required by radar fitters.
ᴣ is considerable overlapping; much of what was once regarded as being exclusively in the
ince of the radar fitter is now a requirement for the communications fitter also, and vice versa.
efore those under training in the radar trades may find much that is useful in Part 2, whilst
ᴣ under training in the communications field may find much of interest in Part 3.

Γhe notes deal with the basic theory and the applied principles of electricity, electronics
·adio in a general way. They do not cover specific details of equipment in use in the Service.
 details are to be found in the official Air Publication for the equipment and this should
ys be consulted during the servicing of the equipment.

Νo alteration to these notes may be made without the authority of official Amendment Lists.

Readers of this publication are asked to report any unsatisfactory features which they may
ᴣ. Such unsatisfactory features may include factual errors in words or illustrations ;
guous, obscure or conflicting instructions ; and omissions. All such unsatisfactory
.res should be reported to Headquarters Technical Training Command (Trg 4c) in the form
ut in A.P. 113A, Section 1, Chapter 1, Paragraph 8.

ΙISTRY OF DEFENCE (AIR)

, 1964

SECTION 1

COMMUNICATION PRINCIPLES

Chapter 2

COMMUNICATION BY RADIO

Radio Communication

Communication by radio implies the use of radio waves as a 'carrier' between transmitter and receiver. As previously mentioned in these notes, radio waves are electromagnetic and travel with the velocity of light. In fact they are of the same nature as light but are of a very much longer wavelength, i.e. they are of a lower frequency. They radiate from a transmitting aerial in much the same way that a ripple radiates from the point where a pebble is dropped into a pond (Fig. 1).
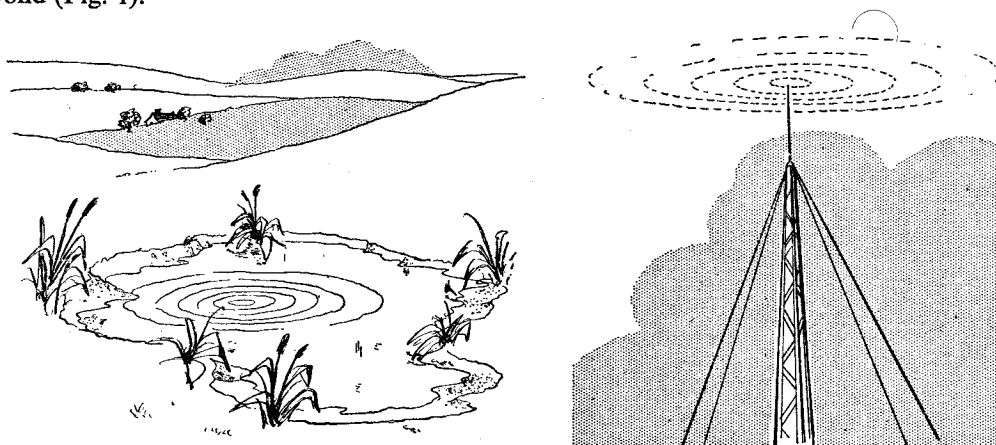

Fig. 1. RADIATION

The main difference between the pond ripples and radio waves is that radio waves are not a movement of a substance whereas water ripples are. In fact radio waves are merely combined electric and magnetic fields which travel through space. They are weakened (or attenuated) by all forms of matter and so they travel best in a vacuum. Thus the basic requirement for an effective radiator of electromagnetic waves (an aerial) is that it should be remote from any large objects which may affect the radiated waves, i.e. it should be high up in open space.

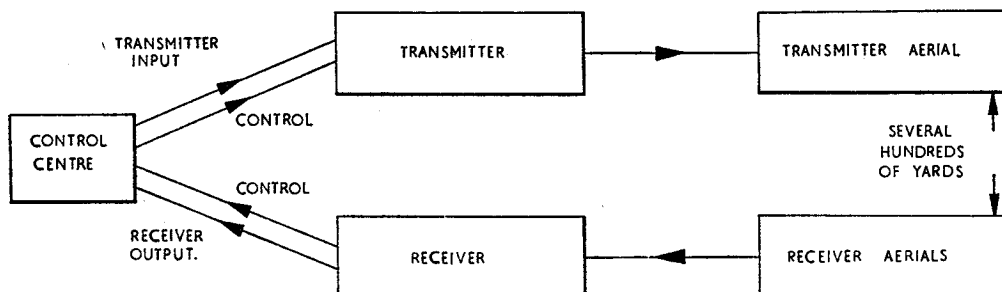Fig. 2 shows the basic layout of a radio communication station.
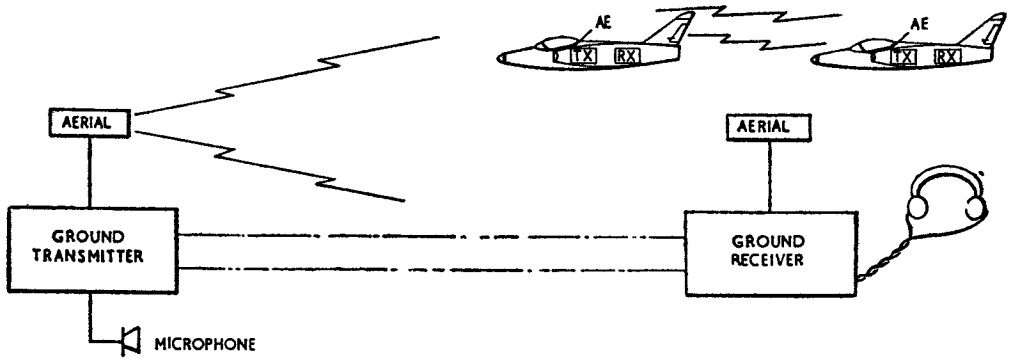

Fig. 2. TYPICAL RADIO STATION LAYOUT

5

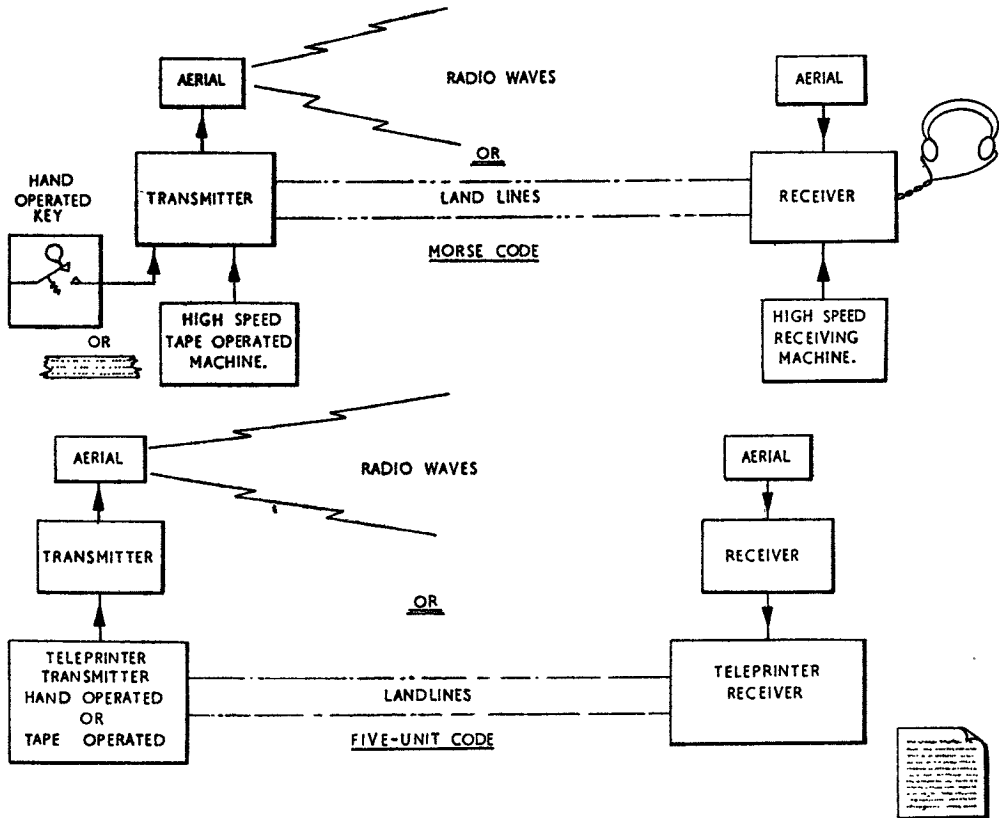Fig. 2. METHODS OF COMMUNICATION BY TELEPHONY

Fig. 3. METHODS OF TELEGRAPHIC COMMUNICATION

2

can be acted upon immediately. For this reason communication between the pilot of a fighter aircraft and ground controllers is by means of radio telephony (r/t), i.e. telephony using radio waves (see Fig. 2).

Telegraphy

Telegraphy is a means of communication using a series of pulses to form a code in which information is contained. The codes most used in telegraphy are the morse code, the five unit code and the seven unit code. Transmissions using the morse code can be made either by a hand-operated key or by a high speed tape-operated machine.

In the five unit code each character is represented by a particular combination of five electrical impulses each of the same duration but in one of two signalling conditions, namely, mark or space. In the seven unit code each character is made up of three mark pulses and four space pulses. The advantage of the seven unit code over the five unit code is that errors due to interference are self-detecting.

These two codes are used with a teleprinter the transmitter of which is connected by landline or radio to the receiver. The transmitter can be manually operated like a typewriter or it can be automatically operated by punched tape. The transmitted message can be received on tape or it can be printed on paper (Fig. 3).

Pulse modulation is widely used in telegraphy. A simple c.w. transmitter is pulse modulated by the operator's key to form the dots and dashes of the morse code; automatic teleprinter systems are pulse modulated to form the five unit or seven unit code. Pulse modulation is also used in data transmission systems.
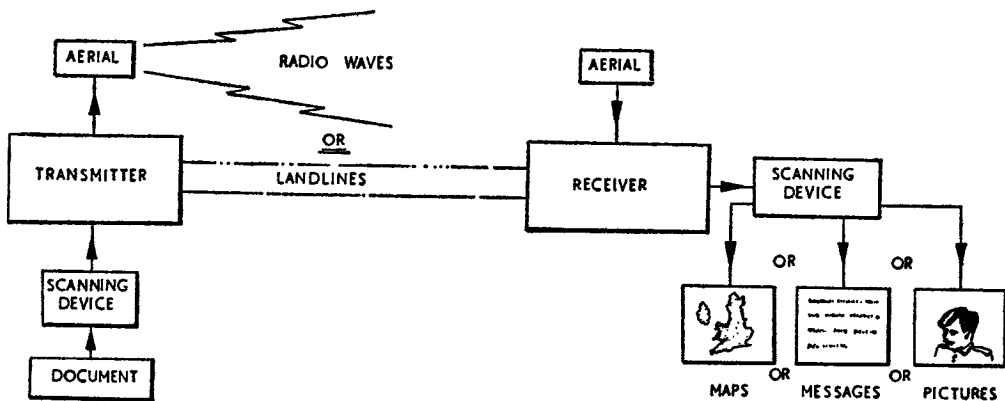
Fig. 4. FACSIMILE COMMUNICATION

Facsimile

With this form of transmission the information is sent out over landlines or radio in the form of pictures, written messages or maps. This method of communication has certain advantages and is dealt with more fully in Section 4. Briefly, the method employed is that the original image is 'scanned' in a series of thin lines by a photo-electric device which translates the various shades into electric currents. These currents are then transmitted, picked up at the receiver, and by a scanning process (synchronised with the transmitter) are reproduced as varying shades on photo-sensitive or electro-sensitive paper. When the completed picture is received, an image of the original subject is produced.

3

Comparison of Telegraphy and Telephony

Although telephony is more convenient to use than telegraphy and little training is required for the operators, telegraphy has several advantages over it. With r/t, information is sent in plain language and it can be intercepted and understood by the enemy. Telegraphic messages on the other hand can be in code and their meaning concealed from the enemy. Radio telegraphy messages, since they occupy a comparatively narrow band, can be received with a much higher signal to noise ratio than can r/t messages, and therefore have a greater range for a given power.

A message can be sent by automatic telegraphic machine much faster than by r/t and, in addition, a written record may be provided. Further, the number of circuits which can be used for r/t is limited by the number of frequencies available, while with telegraphy the number of channels which can be provided on one frequency can be increased by using 'voice frequency' techniques on single, double and independent sideband transmissions. These techniques are considered in Section 4.

Outline of the RAF Telecommunications Organisation

This organisation, in which radio tradesmen play an important part, provides efficient and reliable communication between all formations of the RAF in this country and overseas. All the methods of communication outlined in the previous paragraphs are used. An operational flying unit must be able to communicate with its aircraft at short range in order to give landing instructions and to control the aircraft in its airfield circuit. For this purpose a low power r/t channel, working on v.h.f. or u.h.f. and with a range of 5 to 7 miles, is required. To control aircraft approaching the airfield at a range of about 120 miles a more powerful v.h.f. or u.h.f. r/t channel is used. A radio telegraphy channel is needed for large long-range aircraft. This must be effective over many hundreds of miles to provide contact between the aircraft and a central air traffic control. Over this link are passed meteorological messages, navigation warnings, aircraft position reports and so forth.

For communication between stations in a Command, landlines are normally used to link teleprinter and telephone transmitters and receivers. For communications between overseas Commands and UK, long range radio teleprinters and manual or automatic high speed morse systems are used. Techniques employed to increase the amount of traffic which can be handled include single sideband and independent sideband transmissions. To increase reliability of reception under varying propagation conditions diversity reception is often used. For short range point-to-point communication links where it is uneconomical to lay land lines, u.h.f. and microwave radio relay systems are employed.

Facsimile transmissions are used in addition to telephonic and telegraphic systems. It is the only method of transmitting pictures, maps or charts by radio or landline and is used in the meteorological service.

Conclusion

This chapter has given an overall picture of the methods of communication employed in the RAF. Some of the terms used may sound strange but they will be explained in greater detail in later sections.

COMMUNICATION BY RADIO

Radio Communication

Communication by radio implies the use of radio waves as a 'carrier' between transmitter and receiver. As previously mentioned in these notes, radio waves are electromagnetic and travel with the velocity of light. In fact they are of the same nature as light but are of a very much longer wavelength, i.e. they are of a lower frequency. They radiate from a transmitting aerial in much the same way that a ripple radiates from the point where a pebble is dropped into a pond (Fig. 1).
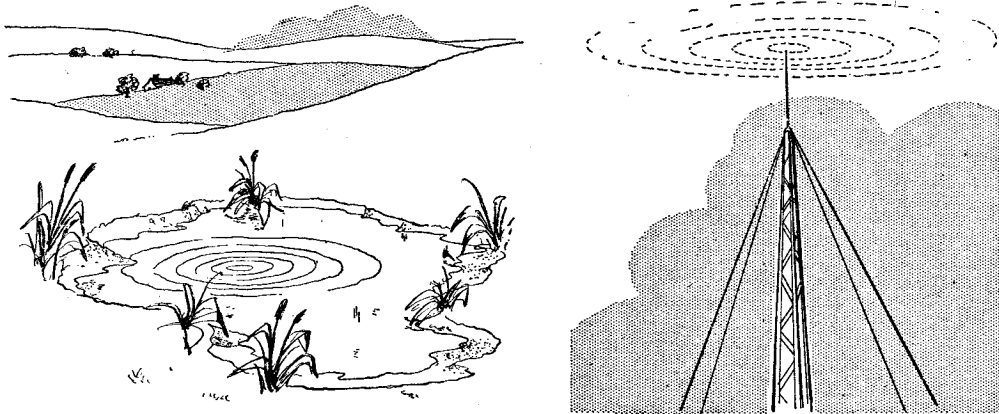
Fig. 1. RADIATION

The main difference between the pond ripples and radio waves is that radio waves are not a movement of a substance whereas water ripples are. In fact radio waves are merely combined electric and magnetic fields which travel through space. They are weakened (or attenuated) by all forms of matter and so they travel best in a vacuum. Thus the basic requirement for an effective radiator of electromagnetic waves (an aerial) is that it should be remote from any large objects which may affect the radiated waves, i.e. it should be high up in open space.

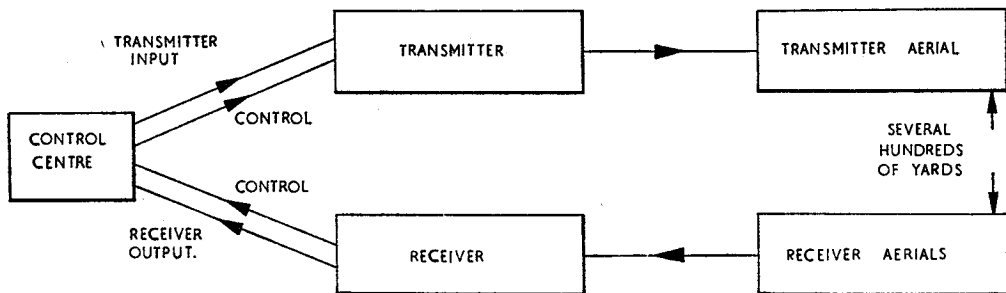Fig. 2 shows the basic layout of a radio communication station.

Fig. 2. TYPICAL RADIO STATION LAYOUT

Basic Radiation Patterns

Different services will require different radiation patterns. For broadcast transmissions uniform all-round radiation is required, while for communication between two fixed points the radiation can be beamed like a searchlight and the system can therefore be made much more efficient.

Typical aerial systems for both broadcast and point-to-point radiation are shown in Fig. 3. Notice that the actual aerials are much the same in both systems, being simply resonant lengths of conductor. It is the parabolic reflector which makes the left hand arrangement appear so different and it is, of course, the reflector which causes beaming.
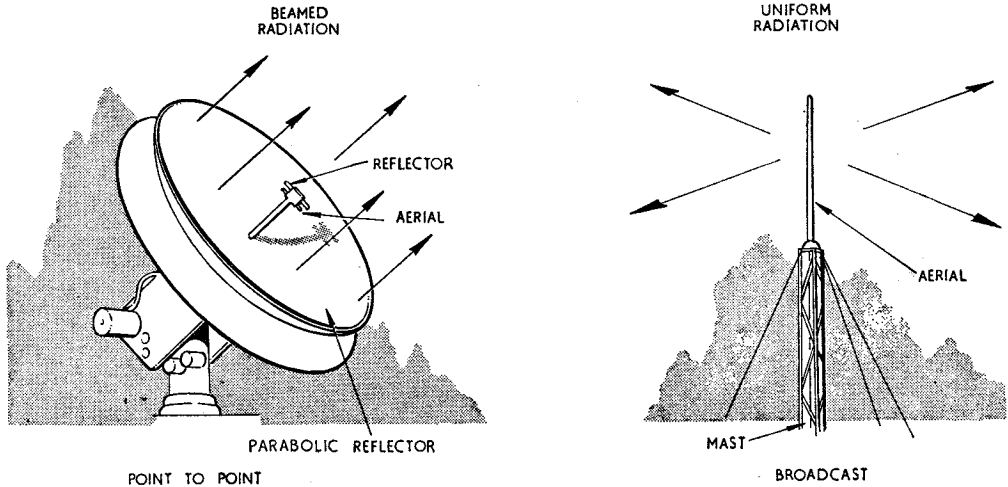

Fig. 3. RADIO AERIALS

Propagation Paths

The path followed by the radio waves will depend upon many factors such as wavelength and type of aerial used but in general one of the three paths shown in Fig. 4 will be used. These three paths are called:
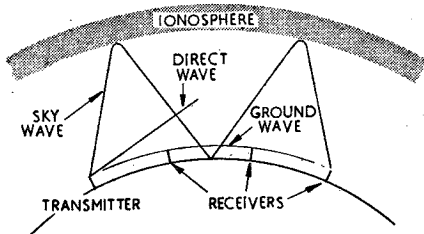
Fig. 4. PROPAGATION PATHS

a. The *direct wave* along the 'line of sight' or 'optical path'.

b. The *ground wave* following the curve of the earth's surface.

c. The *sky wave* 'bouncing' between earth and ionosphere.

Notice that the direct wave cannot provide long-distance communication between ground stations. Thus point-to-point communication from (say) London to Singapore must be by either ground wave or sky wave. Since the former suffers a continuous loss of energy to the earth, the sky wave is preferred.

6

Typical Frequency Usage

Table 1 shows the use of various frequencies in radio communication. The v.l.f. band was developed first and nowadays increasing use is made of this band for complete reliability of communication and for communication with long-range submerged submarines. As component design has improved, and to relieve congestion in the lower frequency bands, higher frequencies have become widely used.

Band	Use	Path	Range
VLF and LF	Broadcasts to ships; weather reports; telegraphy	Ground wave	1,000 miles or more
MF	BBC broadcasting; shipping traffic; telegraphy	Ground wave: sky wave	50 to 250 miles 1,000 miles at night
HF	Intercontinental point-to-point working; aircraft; ships; telegraphy	Sky wave	World wide
VHF and UHF	TV; high quality broadcasting; aircraft r.t.; telegraphy	Direct wave	Optical range
SHF	Microwave links; radar	Direct wave	Optical range

TABLE I.   FREQUENCY BAND CHARACTERISTICS

Choice of System

As outlined in Chapter 1, radio is used in both telegraphic and telephonic systems and both are equally important in modern communications.

The choice between the use of speech and code is mainly decided by the degree of urgency involved. For example, the pilot of an aircraft being 'talked down' by a ground controller must know at once the approach instructions, i.e. direct contact is essential and so telephony is used. On the other hand, speech is 'plain language' information and may compromise security. Thus, where the degree of urgency is less, some form of communication code is then used. Code is also useful when automatic operation is required and has the additional advantage of easy recording at high speed.

Conclusion

This section has served as an introduction to communication by means of radio waves. The relative advantages of two main systems (r.t. and c.w.) have been discussed and the frequency bands commonly used for communication, together with their applications, have been mentioned. The next section will deal with circuit details of c.w. and r.t. transmitters and will discuss the different techniques used in variations of the basic transmitter.

SECTION 2

TRANSMITTERS

CHAPTER 1

THE BASIC TRANSMITTER

Introduction

A simple-three stage transmitter has already been considered in basic theory. Many types of practical transmitter are used throughout the RAF, ranging from low frequency ground installations of considerable size, weight and power, to miniaturised airborne u.h.f. transmitters. Even so, although the appearance and uses of these transmitters may vary considerably, they all use the same basic principles.

This section will consider the requirements of the various practical transmitters in common use.

The Basic Transmitter

The three basic stages of a transmitter are shown in the block diagram of Fig. 1. The master oscillator produces low power oscillations of stable frequency. These are isolated from
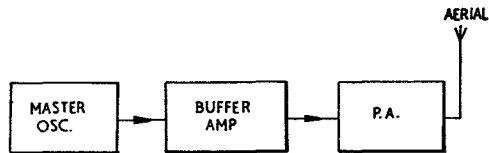
FIG. I. STAGES OF A BASIC TRANSMITTER

load variations, and are also amplified, by the buffer amplifier. The p.a. stage provides the necessary power before the energy is radiated from the aerial.

Frequency Stability

In any practical transmitter it is important to obtain good frequency stability and to provide arrangements which will enable the transmitter to be tuned accurately to the desired frequency. The principal causes of frequency instability in an oscillator are as follows:—

a. Temperature changes can lead to changes in the value of components (L, C and R) and in valve "constants". Transmitters are, in practice, allowed to "warm up" and reach a steady operating temperature before "going on the air" in order to minimise this effect. *Temperature compensated* components are employed for the same reason.

b. Variation in the supply of power to the oscillator can shift the operating point and cause a change in frequency. For this reason the h.t. supplies to anode and screen are often stabilized and so is the heater current.

c. A heavily loaded oscillator is more unstable than a lightly loaded one. In either case variations in the load will affect the frequency. Transmitter oscillators are therefore usually "under-run", i.e. they are operated at a very low power level, and are provided with a light, almost constant, load. This load is the buffer stage.

d. Mechanical movement and vibration of the components in an oscillator will cause frequency drift. Components must be constructed and mounted so that this movement is not possible.

It should be borne in mind that no oscillator is constant in frequency. Some oscillators are more stable than others and the steps which have been outlined would improve the stability of any oscillator, but an improvement is all that can be achieved. For some applications it is adequate to use an oscillator whose frequency does not differ by more than 1% from its nominal value, but for use in a transmitter a stability of about 1 part in 100,000 (or better) is required.

Crystal Oscillators

One way of obtaining good frequency stability is to employ a crystal controlled oscillator. This, of course, requires a suitable crystal, and a crystal oscillator will give only a low power output (less than 5 watts); but an advantage is that it is almost impossible to tune a crystal controlled transmitter to the wrong frequency.

A common form of crystal controlled oscillator is shown in Fig. 2.

Although the frequency of the output from a crystal oscillator is very stable, it does vary if the temperature of the crystal varies. To overcome this the crystal is sometimes mounted in a thermostatically temperature-controlled "oven". The temperature of the oven need not be high, but it must be maintained at a fixed value. Under these conditions the dimensions of the crystal remain constant and frequency stability is improved.
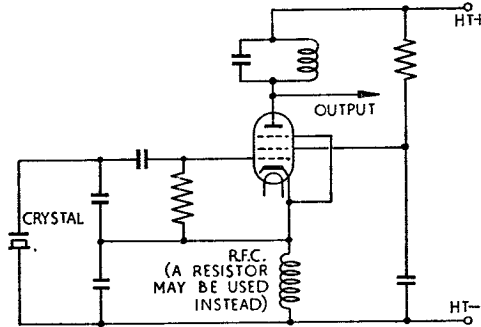

FIG. 2. CRYSTAL CONTROLLED INVERTED COLPITTS OSCILLATOR USING A PENTODE VALVE

Variable Frequency Master Oscillators

Sometimes it is necessary to set a transmitter to a frequency for which there is no crystal available. To cater for this possibility there is provision in many transmitters to switch the oscillator circuit from crystal control to variable frequency oscillator control in which a LC tuned circuit replaces the crystal. The tuned circuit must consist of very high quality components and some of the components may well be temperature compensated to obtain the necessary stability.

The purpose of a master oscillator is to produce a stable r.f. voltage. It is followed by power amplifiers which develop the required power. To assist in obtaining good stability, m.o. valves are "under-run", that is, they are required to develop only a small fraction of the output power which they could provide. For example, if a certain valve can give an output of 20 watts it would be possible to use one such valve in an oscillator circuit to obtain 20 watts directly, but the frequency stability would be too poor for use as a transmitter. To make a reasonably stable transmitter it would be necessary to use two such valves—one as a stable m.o. giving 1 or 2 watts output and the other to amplify the small m.o. output to give 20 watts. Such a transmitter is known as a *m.o.p.a. transmitter*.

The m.o. stage can employ any suitable type of oscillator. The Clapp oscillator circuit shown in Fig. 3 is a typical example. The circuit is similar to that of Fig. 2 with the crystal replaced by a series tuned circuit.
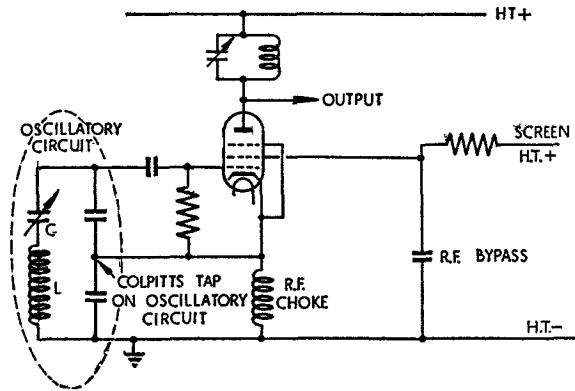
FIG. 3. THE CLAPP OSCILLATOR

Buffer Amplifiers

The more lightly loaded an oscillator is the more stable will be its frequency. Often the m.o. is followed by an amplifier which draws practically no power at all from the oscillator: such a stage is known as a buffer amplifier. In this way very good frequency stability can be obtained since the m.o. is not influenced by any variations in the following circuits. The improvement in stability is indicated diagramatically in Fig. 4.
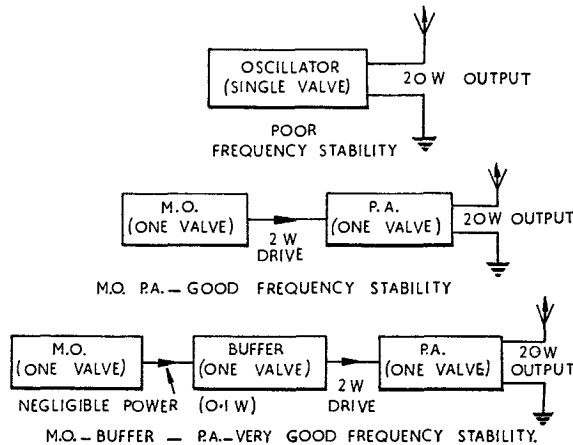
FIG. 4. WAYS OF OBTAINING 20 WATTS RF POWER

Power Amplifiers

Having produced a stable r.f. signal it is now necessary to amplify it to obtain the required power level. The power output which a transmitter is designed to deliver depends upon the service which is to be provided. For local short-range communication less than 5 watts may

11

be sufficient; ground to air radio telephony requires about 50 watts from the ground transmitter; long-range transmitters may have an output of 5 kilowatts or more.

In all cases the power required is obtained by using suitable power amplifiers. The circuit of a p.a. is much the same whether it is to be used to give 10 watts or 1,000 watts; the only real differences are that a high power p.a. valve will:—

a. be large in size.

b. require a large power supply.

c. have special cooling arrangements, either air blowers or circulating coolant.

The anode efficiency of a p.a. stage may be defined as:—

$$\frac{\text{r.f. power output}}{\text{d.c. power input}}.$$

Usually the anode efficiency is simply referred to as the efficiency of the stage. The importance of efficiency can be seen from Table 1 which shows the d.c. power input required to obtain a r.f. power output of 1 kW with different efficiencies.

Efficiency	DC Power required
10%	10 kW
20%	5 kW
40%	2·5 kW
60%	1·7 kW
80%	1·25 kW

TABLE 1 – EFFICIENCY AND POWER INPUT FOR RF POWER OUTPUT OF 1 kW.

The efficiency of an amplifier is determined by the class of operation as shown in Table 2. This is only an approximate indication.

Class	Efficiency
A	20—40%
B	40—60%
C	60—80%

TABLE 2 – EFFICIENCY AND CLASS OF OPERATION.

It will be noticed that the class C amplifier is the most efficient and for this reason the p.a. stages of a c.w. or f.m. transmitter almost always work in class C. In some amplitude modulated transmitters however, modulation is carried out at a low level power stage and the resulting amplitude modulated signal is then amplified by the following p.a. stages. In such transmitters the p.a. stages must be operated in class B in order to avoid distortion.

The following paragraphs describe the function of the class C amplifier and the only significant difference between this and class B is that in class B there is less bias, lower efficiency and less distortion.

Class C Amplifiers

To enable a valve to work under class C conditions it is necessary to provide:—

a. a large grid bias, more than enough to cut off anode current; usually the bias is at least twice the cut-off value.

b. a large amplitude signal input at least sufficient to drive the instantaneous grid voltage to zero and usually sufficient to drive the grid positive.

A large valve may, for example, require a bias of —400V to cut off the anode current when only d.c. voltages are applied. A suitable class C bias would be about —800V and the a.c. grid drive would probably be about 950V peak, so driving the grid to +150V on the positive peaks. This is shown in Fig. 5.
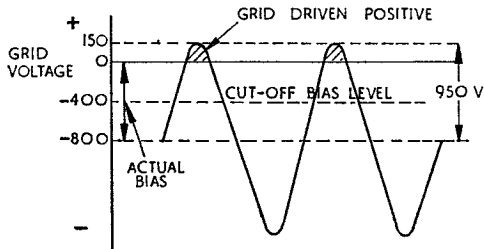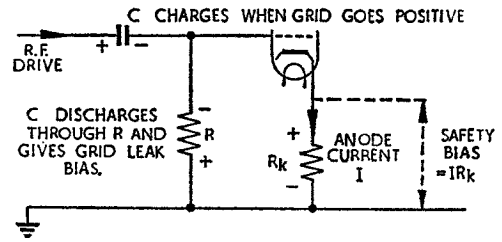

FIG. 5. GRID VOLTAGE IN CLASS C OPERATION

FIG. 6. GRID LEAK BIAS AND SAFETY BIAS

Grid Leak Bias

When the grid is driven positive, grid current will flow because the grid will collect electrons. This grid current may be used to provide grid leak bias (also called self-bias or automatic bias) which is commonly used with class C stages. In addition, safety bias is usually provided by a resistor in the cathode lead so that if the grid drive should fail there will be enough cathode bias to prevent damage to the valve. Both these bias arrangements are shown in Fig. 6.

Under class C conditions the grid of the p.a. valve is driven positive and collects electrons which charge the capacitor C. The charge "leaks" away through the grid leak R and in doing so sets up a voltage drop across R. The current which produces this voltage drop consists of electrons from the grid and in passing through R they make the grid end negative, so giving a negative bias. If the drive increases in magnitude the grid will collect more electrons and the negative bias will increase. A reduced grid drive will give a smaller bias. This is why the arrangement is often called *automatic bias;* it automatically changes as the drive changes. The safety bias is simply cathode bias and under normal conditions is quite small.

Fixed Bias

An alternative arrangement is to provide the grid bias voltage from a power unit similar to h.t. supply units. Such an arrangement is known as fixed bias: as a safety precaution, a

13

cathode resistor may well be included in this circuit, just as with grid leak bias, in case the bias supply fails.

When fixed bias is provided, the d.c. resistance of the grid circuit must be kept small since grid current flow will produce additional bias. In almost all class C stages the grid is driven positive, so that the actual bias is the sum of the fixed bias and the voltage drop due to grid current, together with the cathode bias if used.

Anode Current Waveform

Because of the grid bias and the grid drive, the anode current of a class C stage does not flow continuously, but only for part of the grid input cycle.
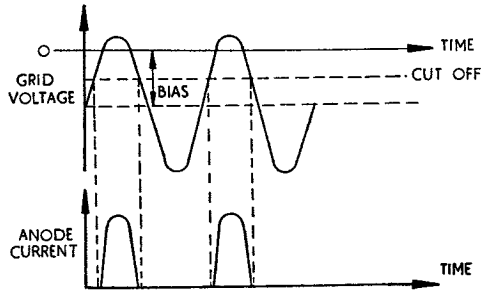

FIG. 7. ANODE CURRENT OF A CLASS C AMPLIFIER

As shown in Fig. 7, the anode current waveform is a very distorted version of the grid input waveform. It has the same fundamental frequency but because of the distortion it is very rich in harmonics; in round figures, the anode current waveform contains about 20% second harmonic and 10% third harmonic. These figures mean, for example, that the amplitude of the second harmonic is about 20% of the amplitude of the fundamental.

A large harmonic content may be very useful in a transmitter as will be explained later, but it does call for a tuned anode load so that only the desired frequency is passed to the next stage. It follows that an a.f. amplifier cannot be worked under class C conditions.

A Class C Circuit

A class C amplifier thus consists of a valve with a large negative bias, a large input signal and a tuned anode load. In the circuits which follow, cathode safety bias will not be shown.
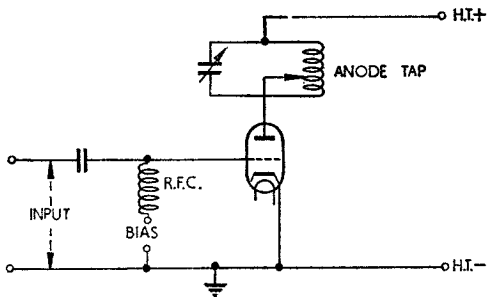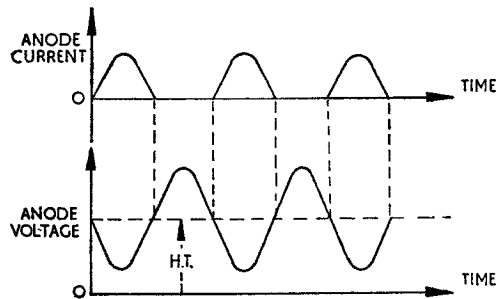

FIG. 8. CLASS C AMPLIFIER

FIG. 9. CLASS C CURRENT AND VOLTAGE WAVEFORMS

The circuit of Fig. 8 shows a class C stage with an anode tap in the tuned circuit. The purpose of the tap is to provide a suitable load impedance match, since unless a p.a. has a suitable value of load impedance it will not develop much power. The adjustable tap performs much the same function as a car's gear box which matches the engine speed to the speed of the wheels.

The tuned anode load resonates at the fundamental frequency of the anode current pulses and develops a voltage at this frequency. The other frequency components of the anode current are outside the bandwidth of the tuned circuit and do not develop a significant voltage. The anode voltage waveform is therefore sinusoidal since only negligible voltages are developed at the harmonic frequencies (Fig. 9).

When the anode circuit is correctly tuned the pulses of anode current flow while the anode voltage is low. This is the real reason for the high efficiency of the class C amplifier: anode current flows only when the anode voltage is low, so that only a *relatively* low power is wasted.

Tuning the Circuit

As the anode circuit is brought into resonance with the grid drive, the voltage developed across the circuit increases, so reducing the anode voltage and therefore reducing the d.c. anode current. By watching an anode current meter, which shows d.c., and tuning the anode circuit for minimum anode current it is possible to tell when an anode circuit is tuned correctly. This is the most common way of tuning a class C amplifier.

In practice the cathode current is usually measured so as to avoid having high voltages on the meter terminals. Instead of connecting a current meter in the cathode lead it is usual to insert a small resistor of a few ohms and then by measuring the voltage across the resistor the current can be determined (see Fig. 10). This has the advantages that the cathode lead does not have to be broken and the use of one meter to monitor more than one valve is made easy.
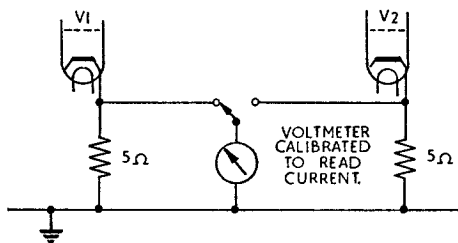

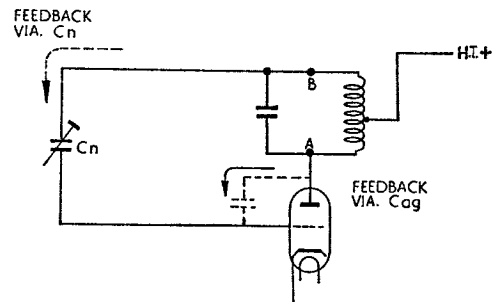FIG. 10. USE OF ONE METER TO READ TWO VALVE CURRENTS

FIG. 11. A NEUTRALIZING CIRCUIT

Neutralizing

Class C stages may employ triodes, tetrodes, beam tetrodes or pentodes. For high power working it is usual to employ triodes since a triode has the least number of electrodes to keep cool. Screen grids are inaccessible and difficult to keep cool if the dissipation is more than a few watts.

A triode has the disadvantage of Miller effect feedback which can cause instability in tuned amplifiers. Even tetrodes and pentodes have sufficient anode-grid capacitance to cause trouble at very high frequencies. To minimise instability, class C amplifiers frequently include a neutralizing adjustment.

The neutralizing circuit is so called because it neutralizes the feedback which occurs via the anode-grid capacitance. One way of neutralizing an amplifier is to connect the h.t. lead to a tap on the anode tuned circuit and then connect a capacitor between the free end of the circuit and the grid.

15

Since the two ends A, B of the tuned circuit (Fig. 11) are in antiphase, current through C_n (the neutralizing capacitor) will be in anti-phase to current through C_{ag}. C_n is adjusted so that the two feedback currents are equal in magnitude and so cancel. There are other neutralizing circuits but they all involve the same principle: feedback through C_{ag} is cancelled by equal and opposite feedback through the neutralizing circuit.

Adjusting the Neutralizing Circuit

The usual method of adjusting a neutralizing circuit (usually referred to simply as "neutralizing") is first to switch off the h.t. supply and set the neutralizer to give no current in the anode circuit. The presence of alternating current or voltage in the anode circuit is checked by means of a simple detector-type circuit.
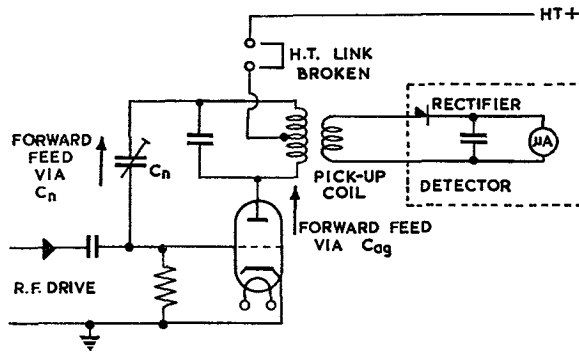

FIG. 12. ADJUSTMENT OF NEUTRALIZING CIRCUIT

If the stage is not neutralized and the h.t. supply is switched off, r.f. power will pass from the grid circuit into the anode circuit via C_{ag}. This will induce a voltage in the pick-up coil (Fig. 12) and will cause the meter to read. The neutralizing control C_n is then adjusted to bring the meter reading to zero when the current through C_n must be exactly balancing (i.e. cancelling and neutralizing) the current through C_{ag}.

Grounded-grid Triodes

The need for neutralizing can be avoided if grounded-grid triodes are employed. In this case the input is taken to the cathode and the output is taken from the anode, while the grid is held at earth potential so far as r.f. currents and voltages are concerned. There can be a d.c. bias connection to the grid as shown in Fig. 13.
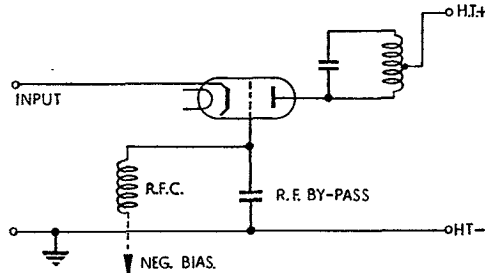

FIG. 13. GROUNDED-GRID AMPLIFIER

16

With this type of circuit the grid acts as a screen between the input and output and so reduces feedback. To minimise feedback through stray capacitances outside the valve, the physical shape of triodes for use in grounded-grid circuits is different from normal triodes. The anode and cathode are at opposite ends of the glass envelope and the grid is connected to a ring around the envelope. Fig. 14 shows a typical valve and its associated anode coil unit.
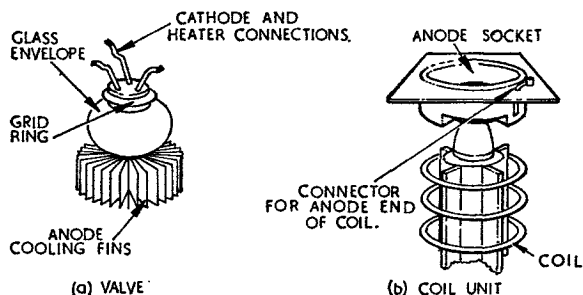
FIG. 14. VALVE AND COIL UNIT FOR GROUNDED-GRID OPERATION

A disadvantage of grounded-grid triodes is that they require much more power in the input circuit than do the more usual arrangements. Further, the physical construction of the associated circuits requires much more care. Because of this, grounded-grid triode valves are specialized types and unsuitable for general use. However, modern transmitter design tends to use grounded-grid triodes in preference to neutralized stages.

Tuning a Transmitter

There are three distinct operations to carry out when tuning a transmitter:—

a. The oscillator must be set to the correct frequency.

b. The buffer and p.a. stages must be tuned to the frequency of the oscillator.

c. The aerial must be coupled to the final p.a. so as to provide a suitable load.

Except in the case of low-power transmitters, adjustments are usually carried out with a reduced h.t. supply so as to minimise the risk of damage due to wrong settings. Some time before a transmitter is required, the heater supply must be switched on so as to allow circuit components to "warm up" to their operating temperature. While the transmitter is warming-up the various controls can be set to the approximate dial readings: these readings can be obtained from the calibration charts supplied with the transmitter.

When the oscillator frequency is controlled by a quartz crystal the correct frequency is obtained simply by inserting the right crystal. When a variable frequency master oscillator is used it can be set to the correct frequency by using a wavemeter (sometimes called a frequency meter). Wavemeters measure the frequency of an oscillator by means of a high quality stable tuned circuit with a calibrated control.

It is usual for transmitters to have a tuned anode load for each stage and for the grid circuits to be untuned, so that the tuning of a transmitter is simply a matter of tuning anode load circuits to resonance. In a transmitter which employs frequency multiplier stages, the anode circuits would be tuned to resonate at the required *harmonic* of the input frequency. In this case it is important to check that the tuning dial is at the approximate setting in order to make sure that the correct harmonic has been selected.

17

To judge when an anode circuit is at resonance requires some form of meter presentation. There are two ways in which an indication of resonance can be obtained:—

    a. tune for *minimum* anode (or usually, cathode) current.

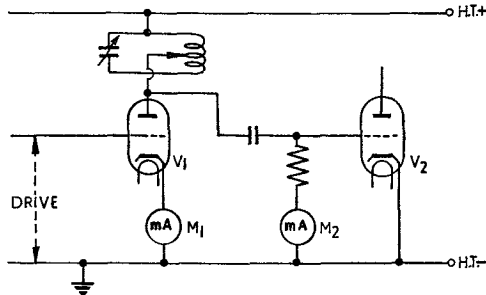   *b.* tune for *maximum* grid current in the following stage.

FIG. 15. CATHODE CURRENT AND GRID CURRENT METERS

The first method has been considered before. The second follows quite simply from it. When the anode current of V_1 (Fig. 15) is tuned to the frequency of the input (or to a harmonic of this frequency) it will present maximum impedance to the pulses of anode current and the cathode current will be a minimum. At the same time, the voltage across the anode load will be a maximum and as this is passed to the grid of V_2, it causes V_2 to draw maximum grid current. Thus the anode load of V_1 is at resonance when either M_1 reading is a *minimum* or M_2 reading is a *maximum*. It is not necessary to use both meters of course: either one can be provided.

The anode circuit of an intermediate p.a. stage is tuned by using a milliammeter in one of the positions shown in Fig. 15. If a neutralizing circuit is provided it must be adjusted as already indicated. The tuning of the final p.a. stage is also carried out by adjusting for minimum anode (or cathode) current but it is further complicated by the need to adjust the coupling to the aerial.

Coupling the Aerial to the Final Power Amplifier

To correctly couple the aerial to the final p.a. stage involves providing a suitable load. For example, a car dynamo may be designed to deliver 120 watts at 12 volts which implies a current of 10 amps. It can do so only if the load is 1·2 ohm: to a load of 60 ohm it can supply only $\frac{1}{5}$ amp. which means only 2·4 watts.

A p.a. stage is designed to work at a certain voltage and to deliver a certain power. Thus there are two adjustments to make:—

    a. the anode circuit must be tuned to resonance.

    b. the load must be coupled so that it will draw the correct current, i.e. present the correct value of load impedance (Fig. 16).
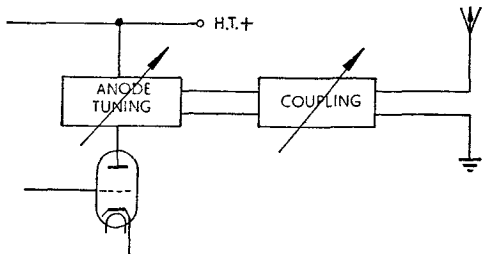

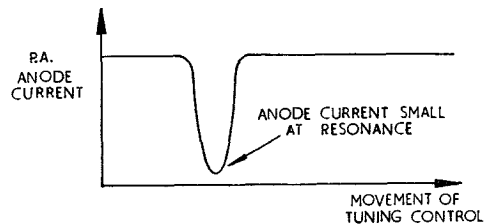FIG. 16. PA TUNING AND AERIAL COUPLING

FIG. 17. LIGHTLY LOADED PA STAGE

The sequence of adjustment is always the same. Initially the coupling is very loose so that the aerial draws very little power and when the anode circuit is tuned to resonance the anode current will be very small. This means that the tuning will be very sharp (Fig. 17), i.e. the dip in the reading on the tuning meter will be very definite and the needle will drop back sharply to a low reading.

The second step is to increase the coupling to the aerial so drawing more power from the p.a. and thus causing more current to flow at the resonant frequency. The "dip" will thus be very shallow and broad: tuning will be less sensitive and it may even be difficult to find a dip on the tuning meter (Fig. 18).
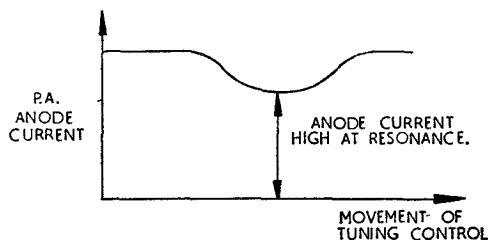
FIG. 18. HEAVILY LOADED PA STAGE

The p.a. stage will be correctly adjusted when the aerial coupling is such that with the circuit tuned for minimum anode current, this minimum anode current is at, or just below, the value recommended in the operating instructions for that particular transmitter.

Aerial Coupling Circuits

The form of the final p.a. anode circuit is governed mainly by two considerations:—

a. the harmonic content of the output power must be low—by international agreement—less than 200 mW.

b. the transmitter must be capable of working with a variety of aerials and transmission lines. This means that the output coupling circuit must be capable of catering for a variety of load impedances: a possible range encountered in practice might well be 100 ohms to 1,800 ohms.

The harmonic content of the output is important only with large transmitters since a small transmitter will in any case have a low output. If a r.f. transformer is used to couple the p.a. stage to the aerial it is necessary in large transmitters to incorporate a filter to eliminate harmonics (Fig. 19).
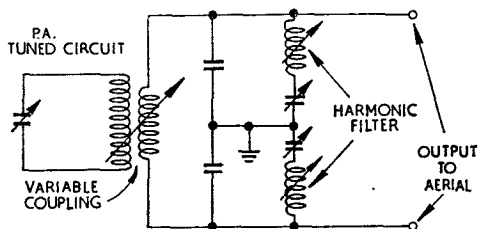
FIG. 19. HIGH-POWER PA OUTPUT CIRCUIT

19

An alternative form of output circuit is called the *pi-coupler* (Fig. 20) in which the tuned circuit is arranged like a low-pass filter.
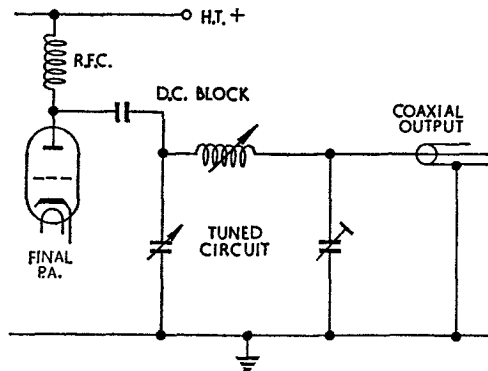
FIG. 20.   PI-COUPLER

A disadvantage of the pi-coupler circuit is that it is suitable only for an unbalanced (co-axial cable) output and not for a balanced output such as the 600 ohm transmission line commonly used in ground communication transmitters.   For a balanced line the circuit of Fig. 21, which is a development of Fig. 19 might be used.
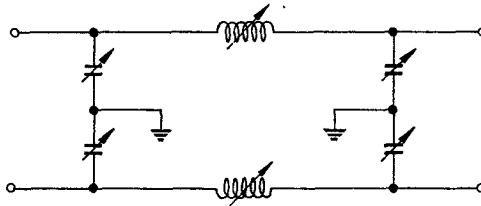
FIG. 21.   BALANCED COUPLER

Tuning Summary

The steps involved in tuning a transmitter may be summarized as follows:—

a. Set the oscillator to the correct frequency by inserting the correct crystal or by checking with a wavemeter.

b. Tune the buffer stages, frequency multipliers and power amplifiers for minimum anode current (or for maximum grid current in the following stage).   Take care to tune to the correct harmonic where necessary.

c. Neutralize the necessary stages.

d. Increase the load coupling from minimum up to the degree of coupling which gives the correct anode current.   Retune the associated tuned circuit for minimum anode current.

e. Check all tuning when full h.t. has been applied and carry out minor adjustments.

Conclusion

This chapter has dealt with the methods of obtaining a high power stable frequency output from a basic transmitter.   The basic method of tuning such a transmitter has been discussed and in Chapter 2 some of the systems employed to remotely tune transmitters and receivers will be considered.

20

CHAPTER 2

REMOTE CONTROL OF TUNING

Introduction

The need often arises to tune a transmitter or a receiver to a desired frequency accurately and quickly, *from a distance*. A communication equipment is often situated in an inaccessible part of an aircraft and it is clearly impracticable for the pilot to carry out detailed tuning whenever he wishes to call a new station. Provision must be made for him to select the required channel without leaving his seat.

The problem is one of turning a tuning control from one angle to another. Originally the controls were moved by means of mechanical links using flexible drive cable, but the control was stiff to operate, fine adjustment was difficult and "backlash" introduced inaccuracies. With modern communication equipment tuning is an operation requiring a high degree of accuracy and involving the precise setting of several controls. This is particularly so in the case of transmitters where it is usual to tune each circuit individually to ensure maximum efficiency at all frequencies. Often the equipment is "pre-set" to a required channel, i.e. it is tuned on a test bench to the frequency on which it will be required to operate later. It is essential that the initial setting up is done accurately and that the pre-set positions of the tuning controls be exactly reproduced when the channel is selected. To ensure this, as much attention is devoted to the choice and design of the tuning mechanism as to the circuit it tunes.

Basic Tuning Mechanisms

The actual tuning device, perhaps a variable capacitor, is usually driven by the tuning control via some reduction gearing. This enables the use of either fast or slow tuning; fast for making quick, large changes of frequency, and slow for final exact setting. Typical reduction ratios are 10:1 (fast) and 40:1 (slow).

Various methods are used to achieve the required reduction ratio: two of the commonest, friction and cog wheel drives, are illustrated in Fig. 1.
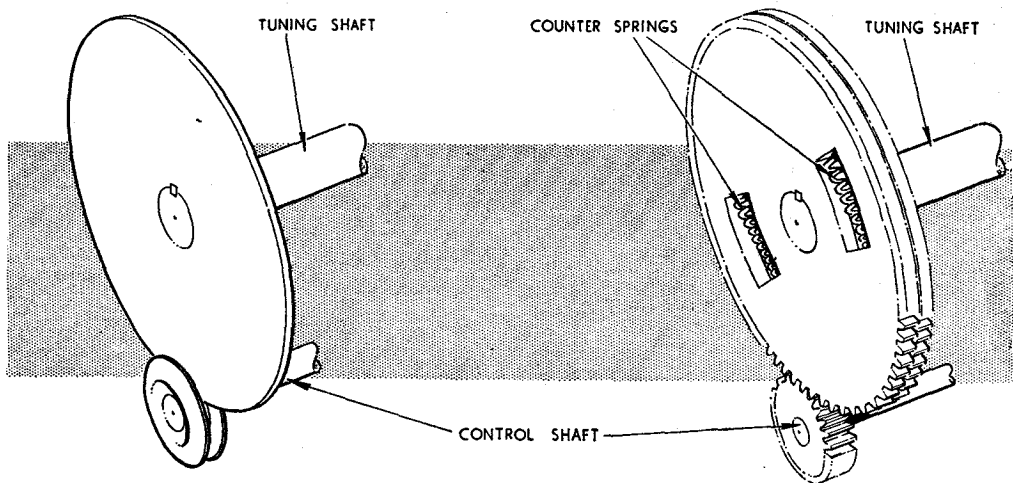

FIG. I. BASIC TUNING MECHANISMS

The cog-wheel drive uses two identical wheels on the tuning shaft. These are held under slight pressure by counter springs so that the teeth are slightly misaligned. The result is that there is no gap between the teeth of the control and the drive wheels *whichever way they tend to move*. Thus there is no "play" on the tuning control and it always causes a definite movement of the tuning shaft, even for very slight control movements. This arrangement of gears is termed "anti back-lash".

Remote Tuning

Where it is necessary to provide remote control of tuning, such as in certain airborne equipments, the tuning controls are usually pre-set to a number of frequencies or channels. Tuning implies selection of one of these channels. Selection is usually either by *switching* the appropriate pre-set tuning control into circuit or by *rotation* of the main tuning control to a pre-set position. The former method has the advantage of simple circuitry but its use is limited by the number of channels required since each channel needs a separate push-button and a separate pre-set control. The latter system provides a large number of channels but the associated tuning mechanism is more complex. A simple example of each is considered in the following paragraphs.

Push-button Tuning

A simple example of pre-set tuning is that often employed in a car radio using push-buttons. In this case however, tuning is only remote to the extent of the push bars and the buttons are really only a convenient form of switch to control the pre-set circuits. Fig. 2 shows a simple four channel arrangement for tuning two circuits.
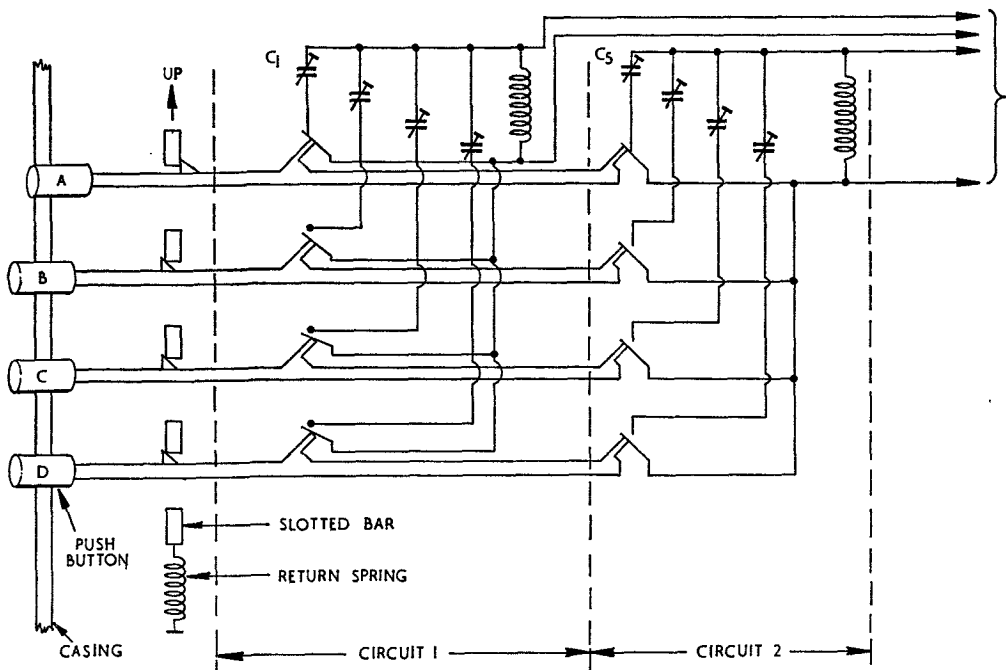
FIG. 2. SIMPLE PUSH-BUTTON UNIT

Each tuned stage includes four pre-set capacitors, one of which is selected by the appropriate push button. When the button is pushed, the slotted retaining bar is forced into the

"up" position, thus releasing the previously selected button. When the button is pushed fully in, the return spring pulls down the slotted bar which then holds in the selected button by means of the stop on that bar.

Movement of the bar switches in or out the appropriate pre-set capacitors and the associated circuits operate at the particular frequency to which the capacitors were pre-set. In the case of a receiver the circuits could be the input circuit of the mixer and the local oscillator circuit. In the case of a simple transmitter the circuits could be the master oscillator and the power amplifier.

Positional Tuning

The other method of remote tuning (rotation of the main tuning control) can employ rotary switches in place of the more bulky push buttons. A simple four channel selector system is shown in Fig. 3 and consists of:—
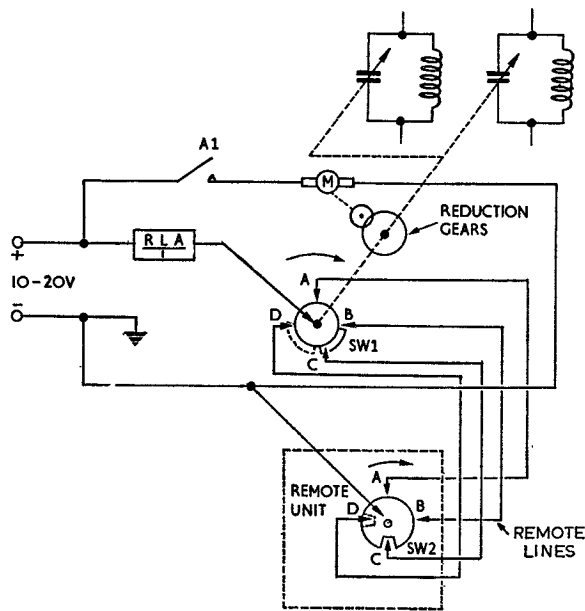

FIG. 3. SIMPLE SELECTOR UNIT

(1) a remote four-position rotary switch (SW2).

(2) relay RLA which is energised from a l.t. supply.

(3) a motor (M) the shaft of which is connected via reduction gears to the circuit tuning capacitors, and which obtains its power from the l.t. supply via relay contact A_1.

(4) a further 4 position rotary switch (SW1) mounted on the reduction gears and connected to SW2 via remote lines.

When the relay is energised the motor rotates, turns the tuning controls until the rotary switch SW1 breaks the relay coil circuit and the motor stops.

With the switches in the position shown in Fig. 3 the relay circuit is broken at SW2 (by the cut-out section) and the relay is de-energised. The relay contact A_1 is thus open and the motor is idle. Channel C is selected.

If the remote control switch SW2 is switched to channel D (shown dotted) the relay circuit is made via contacts SW1 C and SW2 C. Relay contact A_1 closes and the motor is energised. The motor and tuning shaft thus rotate *and so does SW1*. When SW1 reaches the dotted position the relay circuit is broken and the motor de-energised, leaving the shaft set to the pre-set position D.

Notice that the system is easily extended to select ten or twelve channels by using ten-or twelve-position switches and adding the necessary lines.

Continuously Variable Controls

The two arrangements for remote control tuning already mentioned are both really selector systems which enable an operator to select any one of a number of *pre-set* frequencies. In fact, therefore, the operator cannot tune to *any* frequency he wants; he can only select one of a limited choice. When it is necessary to provide un-selected control a different system is required.

The problem is to make a knob turned in one place move a variable capacitor situated in another place perhaps a hundred yards away.

To do this two separate units are used: one sends out information on how much the knob has to be turned and is called the *transmitter;* the other one receives and translates this information into a corresponding rotation of the capacitor and is called the *receiver*. These terms transmitter and receiver are often used in connection with remote control systems and they must not be confused with wireless transmitter and wireless receiver.

The transmitter and receiver are linked electrically by wires and so the only limit to the remoteness of control is that of the length of cable permissible.

The Desynn System

A small permanent magnet is pivoted in a toroidal coil which is fed with d.c. from a remote potentiometer. The permanent magnet assembly forms the receiver, and the remote potentiometer the transmitter. Inter-connections are at 120° intervals as shown in Fig. 4.

As the potentiometer is moved from any rest position the voltage outputs to the remote lines vary and so the currents in the three sections of the toroidal coil vary. These currents give rise to magnetic fields which turn the magnet to its new position.
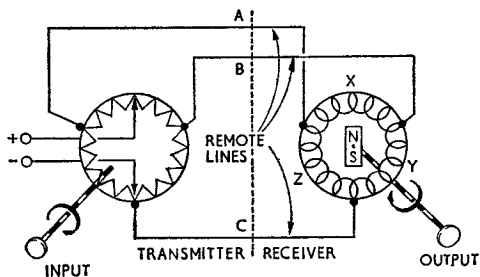

FIG. 4. THE DESYNN SYSTEM

Suppose the input shaft is turned clockwise so that the potentiometer is set to the position shown in Fig. 4. There will be equal voltages on lines A and B (since they are equidistant from the positive input terminal). As a result there will be no current in section X of the coil and equal currents in sections Y and Z. The current in section Z will flow in an anticlockwise direction through the coil producing a magnetic field as shown by vector Z (Fig. 5) and the

24

current in section Y will flow in a clockwise direction to produce a magnetic field represented by vector Y. The direction of the resultant field will be the vector sum of these two fields and is represented in Fig. 5 by the dotted vector.

The pivoted permanent magnet therefore turns to the horizontal position and the tuning shaft follows the movement of the remote input shaft.
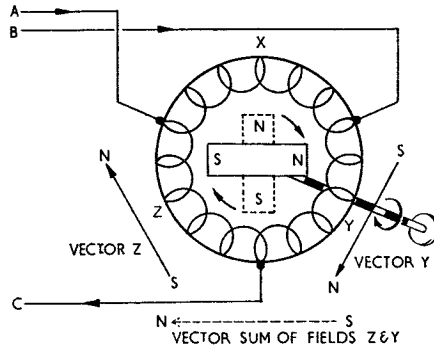

FIG. 5. RESULTANT RECEIVER FIELD

Wheatstone Remote Drive

Another widely used example of continuously variable remote control systems is illustrated in Fig. 6. In this case both transmitter and receiver elements are resistance potentiometers connected to a d.c. supply as shown.
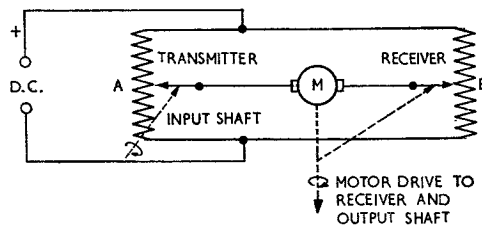

FIG. 6. SIMPLE WHEATSTONE REMOTE DRIVE CIRCUIT

Between the two wiper arms is an electric motor. When the voltage at wiper A equals the voltage at wiper B no current flows through the motor and the system is said to be "balanced". When the transmitter wiper A is moved from this position current flows through the motor and the motor starts to run. The direction of the motor rotation depends upon the direction in which the transmitter slider is moved. When the motor is running it drives the output shaft and at the same time it turns the receiver slider towards a balance position. When balance is reached the motor stops and the output shaft has moved through the same angle as the input shaft.

A disadvantage of this simple system is that for small variations of the input shaft only a small current flows through the motor and a very sensitive motor with correspondingly small power is needed. However, if the motor is replaced by a sensitive polarised relay, then the relay can be used to switch power to a more powerful motor as shown in Fig. 7.

25

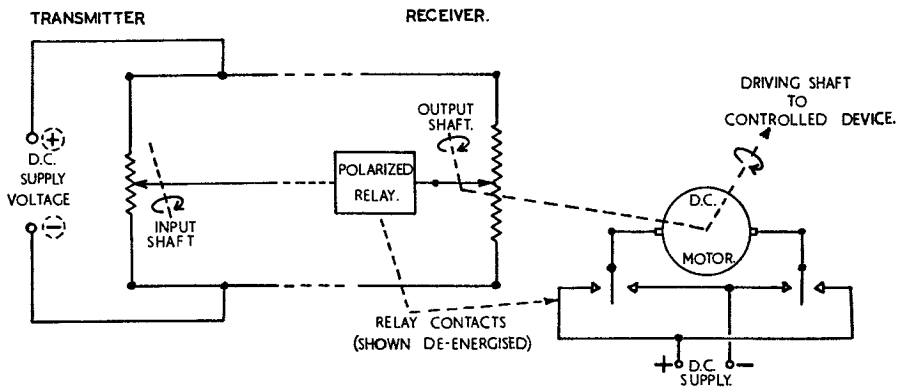TRANSMITTER   RECEIVER.

FIG. 7. WHEATSTONE REMOTE DRIVE CIRCUIT WITH POLARISED RELAY

Since the relay is polarised the direction of current through the coil governs the direction in which the contact arm moves and so controls the direction of rotation of the motor.   At the balanced position the relay is de-energised, the contact arm is in the central position and the motor stops.

CHAPTER 3

FURTHER CIRCUITS

Introduction

As shown in Chapter 1 a transmitter consists of three basic stages: the m.o., buffer amplifier and p.a.; some of the circuits which could be used in these stages have been described. However, other circuits may be used when they offer a particular advantage or when the transmitter has a special function. This chapter deals with some of the circuits which can be used in the r.f. stages of transmitters.

Further Oscillator Circuits

If the components of the tuned circuits of oscillators are allowed to vary in temperature the frequency of oscillation will drift; this is due to expansion and contraction of the materials from which the components are made.

For maximum frequency stability, therefore, the tuned circuit components of an oscillator are maintained at a constant temperature by enclosing them in a thermostatically temperature-controlled oven similar to that already described for a crystal.

The Franklin Oscillator

This circuit gives good frequency stability and is a common form of m.o. circuit (Fig. 1). It can be considered as a valve amplifier together with a positive feedback valve. Both valves, in fact, amplify and because of the high gain of the combination it is possible to obtain oscillations
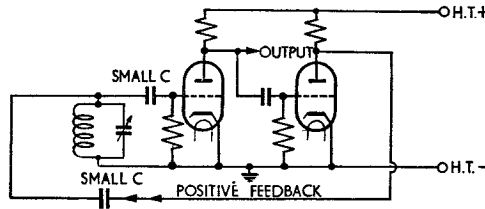

FIG. I. BASIC FRANKLIN OSCILLATOR

with extremely loose coupling to the tuned circuit. Because of this loose coupling the tuned circuit is virtually independent of outside influence and the frequency of oscillations is stable. The stability can be improved by enclosing the tuned circuit components in an oven as previously described.

The Butler Oscillator

This is a two-valve crystal controlled oscillator used in u.h.f. and v.h.f. transmitters. The basic circuit is shown in Fig. 2.

The crystal is operated in what is called an *overtone mode*. A quartz crystal has a fundamental frequency of oscillation determined largely by its thickness. The higher the required frequency the thinner must be the quartz and with reasonable crystal thickness the frequency is usually below 20 Mc/s. However, it is possible, by cutting and mounting a crystal in a special way, to make it oscillate at a frequency much higher than the fundamental frequency. For example, a crystal can be cut with a thickness corresponding to 10 Mc/s and then, by means of a special mounting, be made to oscillate at about 40 Mc/s. This higher frequency is known as an *overtone*.
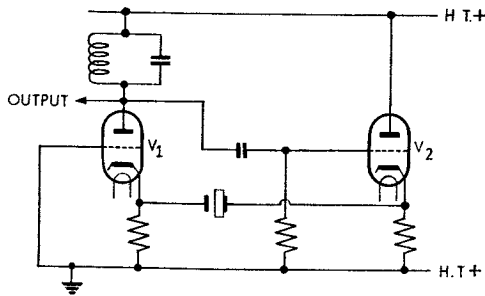
27

FIG. 2. THE BUTLER OSCILLATOR

Overtone oscillations are more easily obtained if the crystal is connected in a low impedance circuit. In the Butler oscillator this is achieved by connecting the crystal between two cathodes which are connected to earth through low value resistances.

The valve V_1 acts as a grounded-grid amplifier and V_2 as a cathode follower. Feedback from V_2 to V_1 is via the crystal which can thus control the frequency. It is possible to take the output from V_2 anode by placing a suitable load between the anode and the h.t. supply.

Combined Oscillator-multiplier

It is often convenient to combine the functions of oscillator and frequency multiplier in one circuit, especially when using a crystal oscillator. Fig. 3 shows a typical arrangement of crystal oscillator-trebler. The grid circuit oscillates at the crystal fundamental frequency but the anode circuit is tuned to the third harmonic of the crystal frequency. A triode valve is unsuitable in this type of circuit and a pentode valve is required. Because of this, the power output of an oscillator-trebler is low.
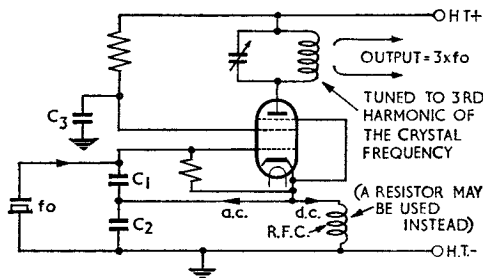

FIG. 3. OSCILLATOR-TREBLER

The oscillator circuit is a crystal form of Colpitts oscillator, feedback being determined by the ratio C_2/C_1. C_3 is the screen decoupling capacitor which connects the screen (the effective anode) to the earthed side of the crystal. The cathode is returned to earth via the r.f. choke thus providing a d.c. path for the anode current. Grid leak bias is provided by the grid resistor in conjunction with C_1 and the capacitance of the crystal holder.

Double-valve Amplifiers

A common method of increasing the power output from a given valve is to use a pair of such valves in parallel or push-pull. The basic principles have been covered for a.f. amplifiers

28

and the arrangements are equally useful for r.f. amplifiers. In the case of triode r.f. amplifiers there is the additional advantage that the centre-tapped anode coil is useful for neutralization.

Push-pull Triode Power Amplifiers

Fig. 4 shows a basic push-pull triode r.f. stage with C_{n_1} and C_{n_2} as the neutralizing capacitors.
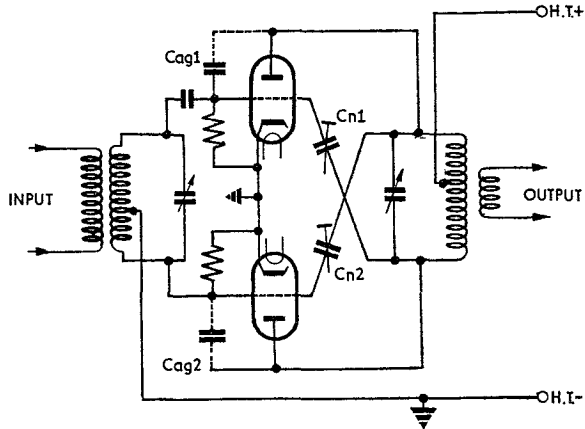

FIG. 4. PUSH-PULL TRIODE PA's

The capacitor C_{n_1} is adjusted to neutralize C_{ag_1} and C_{n_2} neutralises C_{ag_2}. With symmetrical construction and identical valves, C_{n_1} and C_{n_2} can be ganged.

Notice that the push-pull arrangement amounts to feedback via C_{n_1} and C_{n_2} from one anode to the other grid. The circuit is in fact often drawn as outlined in Fig. 5.
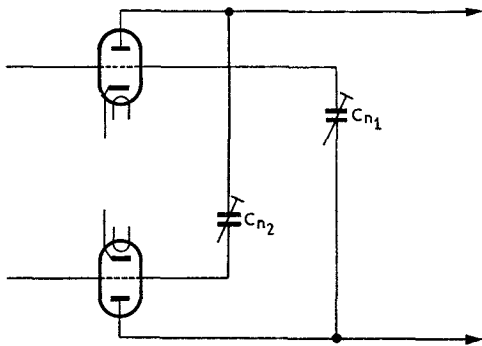

FIG. 5. EFFECTIVE NEUTRALIZING CIRCUIT
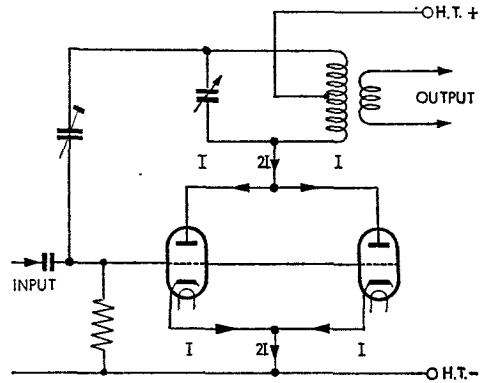
FIG. 6. PARALLEL TRIODES PA STAGE

Parallel Triode Power Amplifiers

A frequent alternative to the use of push-pull triodes is the parallel combination shown in Fig. 6.

The total anode current is the sum of the two valve currents and with two identical valves, power output is doubled. Unfortunately, the total C_{ag} is also doubled, and because of this the arrangement is not suitable for use at frequencies where inter-electrode capacitances are

limiting factors. In fact, the parallel combination is merely a convenient method of obtaining a larger valve.

Thus, with identical valves:—

r_a is halved

g_m is doubled

μ is the same

all inter-electrode capacitances are doubled.

In other words, the combination amounts to one big valve (Fig. 7).
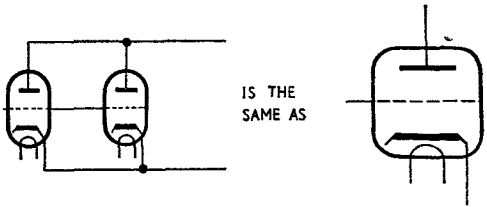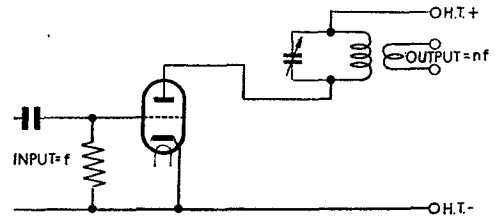

FIG. 7. EFFECT OF PARALLEL VALVES   FIG. 8. BASIC FREQUENCY MULTIPLIER

Frequency Multipliers

The basic idea of frequency multiplication has already been covered in Part 1 and is summarised as follows:—

The circuit of Fig. 8 is simply that of a class C amplifier with the anode circuit tuned to a multiple of the input frequency (usually between 2 and 6 times). If a higher multiple than 6 is required, more multipliers are used, e.g. 3 doublers and one trebler will give 24 times the original frequency.

Note that with the output circuit tuned to a multiple of the input frequency, neutralization is unlikely to be necessary.

A frequency multiplier stage is often called a harmonic amplifier—a name which describes its function.

The Push-push Doubler

This arrangement provides a very efficient frequency doubler in that it uses both half cycles of input in spite of working in class C. In fact, it is an amplifier which corresponds closely to a full wave rectifier. Fig. 9 shows the circuit and Fig. 10 the waveforms. A normal push-pull input circuit is used but the valve outputs are connected in *parallel*.
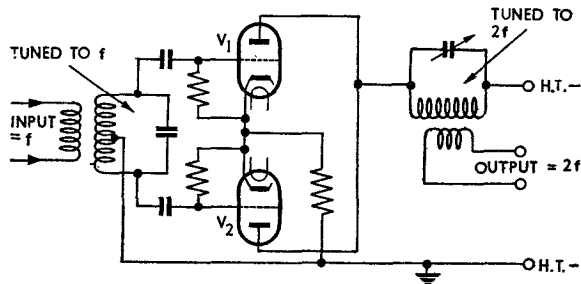

FIG. 9. PUSH-PUSH DOUBLER

30

The circuit is called a "push-push" arrangement since both valves produce in-phase outputs.

A comparison of push-push and push-pull circuits reveals:—

a. The input (grid) circuits of both push-push and push-pull amplifiers are identical, the grid voltage being fed to the two valves in antiphase in each case.

b. The anode circuit of a push-pull circuit must be tuned to the fundamental or to odd harmonics of the fundamental, while for a push-push circuit it must be tuned to an even harmonic.

c. While class A, B or C bias can be used in a push-pull circuit class C bias must be used for push-push.

The Push-pull Trebler

If the anode circuit of a normal push-pull amplifier were tuned to twice the input frequency there would be no output. If the anodes of V_1 and V_2 in Fig. 9 were connected to the output circuit in the normal push-pull arrangement, V_2 anode current waveform would be reversed and the output from V_1 and V_2 would cancel *at the double frequency*.
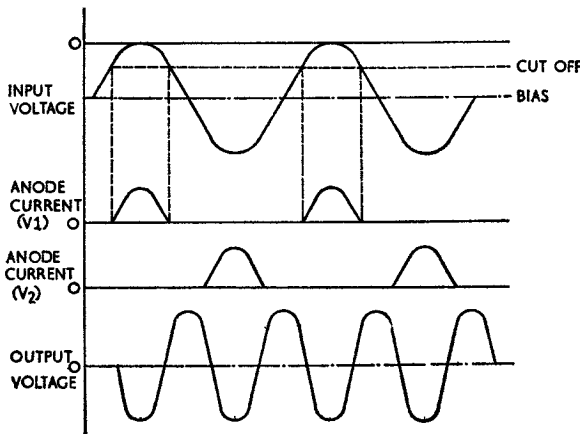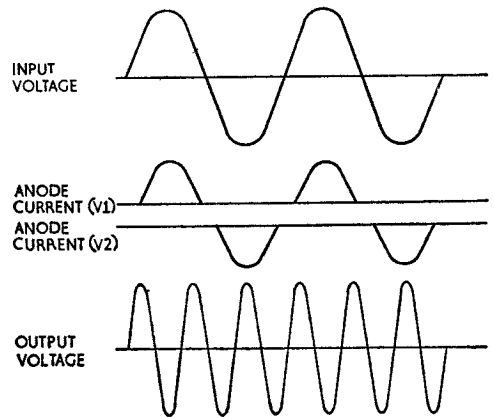
FIG. 10. PUSH-PUSH DOUBLER WAVEFORMS   FIG. 11. PUSH-PULL TREBLER WAVEFORMS

However, if the anode circuit is tuned to three times the input frequency the normal push-pull amplifier becomes an efficient trebler. This is shown in Fig. 11. The pulse of anode current from each valve is in the right phase to maintain oscillations at the *third* harmonic of the input frequency.

Parasitic Oscillations

In most oscillator and amplifier circuits, particularly where appreciable power is involved, there is enough stray coupling and reactance to cause undesirable oscillations at very high frequencies. Fig. 12 shows how these stray reactances in a neutralized p.a. stage can combine to form a conventional oscillator circuit which will produce unwanted v.h.f. oscillations.

Since such oscillations can exist only by drawing power from the normal supply at the expense of the desired amplification or oscillations, they are termed *parasitic*. If they are permitted to occur they represent not only a loss of power but also a source of interference in nearby radio receivers.

(a) R.F. AMPLIFIER

(b) R.F. AMPLIFIER WITH VALVE L&C

(c) EQUIVALENT OF (b) AT V.H.F.

(d) CONVENTIONAL FORM OF (c) - COLPITTS OSCILLATOR
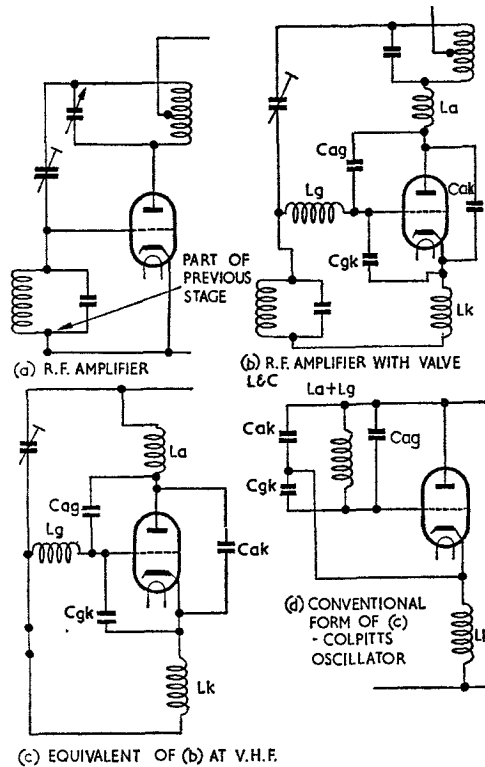
FIG. 12. EFFECT OF VALVE AND WIRING INDUCTANCE AND CAPACITANCE

Suppression of Parasitic Oscillations

Although the troublesome stray reactances cannot be eliminated, they can be reduced by keeping the length of wiring between components to a minimum and by careful planning of wiring and component layout. Long straggly leads are not only untidy but can cause parasitic oscillations. Even earth leads are important and should not form long loops. Fig. 13 shows the correct and incorrect connections in a single r.f. amplifier stage.

FIG. 13. CONNECTIONS IN RF CIRCUITS

Parasitic oscillations can be minimised by using resistors. These can be connected as damping resistors in series or parallel with the parasitic oscillatory circuit, or as v.h.f. potential dividers in the input (grid) or output (anode) leads of the valve. Any oscillatory circuit can be made to stop oscillating if a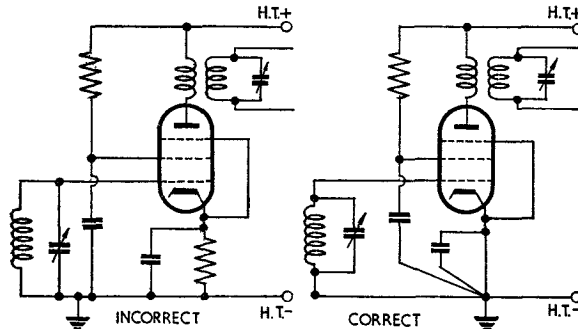 resistance of sufficient value is included in it. Since it is difficult to identify the effective oscillatory circuit, the potential divider arrangement is most commonly used; the resistors are termed grid or anode *anti-parasitic* resistors. Fig. 14 shows a typical grid anti-parasitic resistor and the equivalent amplifier input circuit at normal and parasitic frequencies.
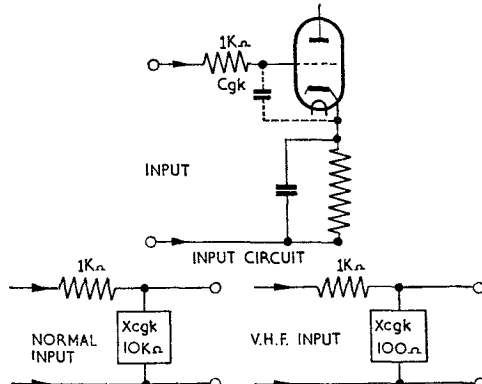
FIG. 14. GRID ANTI-PARASITIC RESISTOR

If at the normal frequency (say about 1 Mc/s) $X_{cgk} = 10k$ then at the parasitic frequency (say 100 Mc/s) X_{cgk} will be only 100 ohms. Thus the actual input to the valve will be 90% of the normal input, but only about 10% of the parasitic input.

Similar reasoning applies if a resistor is placed in the anode or screen lead, but in this case the value of resistor used is limited to a few hundred ohms. Attenuation of both normal and parasitic frequencies is therefore less with this arrangement.

Notice that in either case, in order to function as a potential divider the resistor must be connected direct to the grid or anode. Fig. 15 shows the right and wrong positions for anti-parasitic resistors.

FIG. 15. POSITIONING THE ANTI-PARASITIC RESISTOR

33

The shading shows the area which is important so far as parasitic oscillations are concerned, and in the right hand diagram the anti-parasitic resistor is obviously outside the effective circuit.

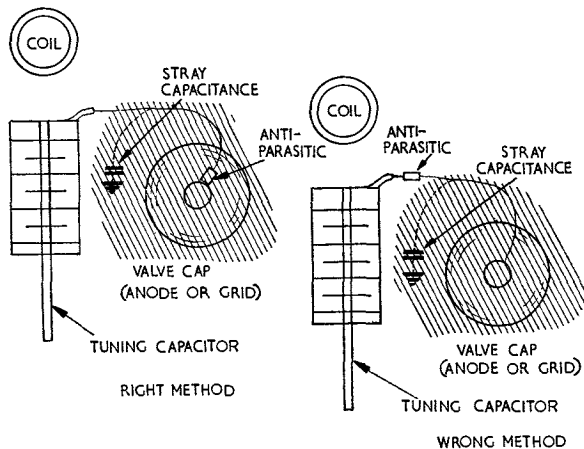Another anti-parasitic device is the use of short-circuited resistors. The value of resistor is only a few ohms, but at v.h.f. the self capacitance and inductance of the resistor and the shorting loop form a rejector circuit to stop parasitic currents.

Typical General Purpose Transmitter

Fig. 16 shows the block diagram of a typical low-power ground transmitter suitable for transmitting c.w., r.t. and m.c.w. signals: approximate power levels are indicated.
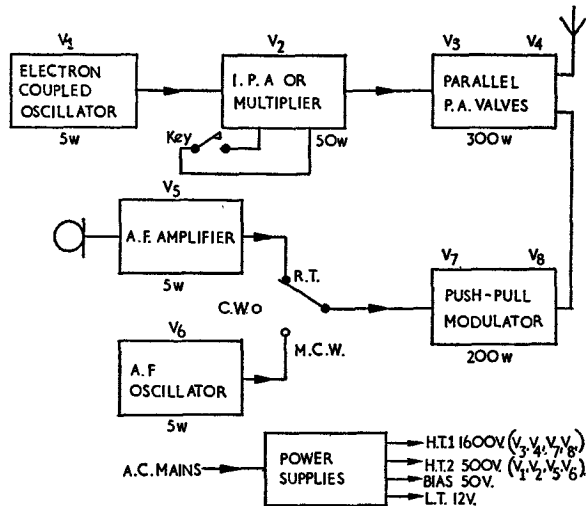
FIG. 16. GENERAL PURPOSE TRANSMITTER

The circuits of all the stages except the modulator stage have already been covered in these notes. Even the modulator stage is not really new since it is basically a push-pull a.f. amplifier. Fig. 17 shows the front layout of such a transmitter for use on the ground.

The basic transmitter comprises four main units, usually mounted vertically with the heavy power unit at the bottom and the output unit at the top. All units are well ventilated and often include cooling fans.

Notice that the h.t. is switched separately and that mains input, l.t. and h.t. circuits have independent fuses.

Monitor Circuit

It is sometimes necessary to listen to the output from a transmitter to check on the quality of the transmission. All that is needed is a simple tuned circuit with a diode detector, as shown in Fig. 18.

The monitor would be used near the transmitter and an aerial is not necessary. For most purposes it would be adequate to plug a pair of headphones into the monitor to listen to the transmitted signal.

FIG. 17. TYPICAL GROUND TRANSMITTER LAYOUT

FIG. 18. SIMPLE MONITOR CIRCUIT

CHAPTER 4

VHF AND UHF TRANSMITTERS
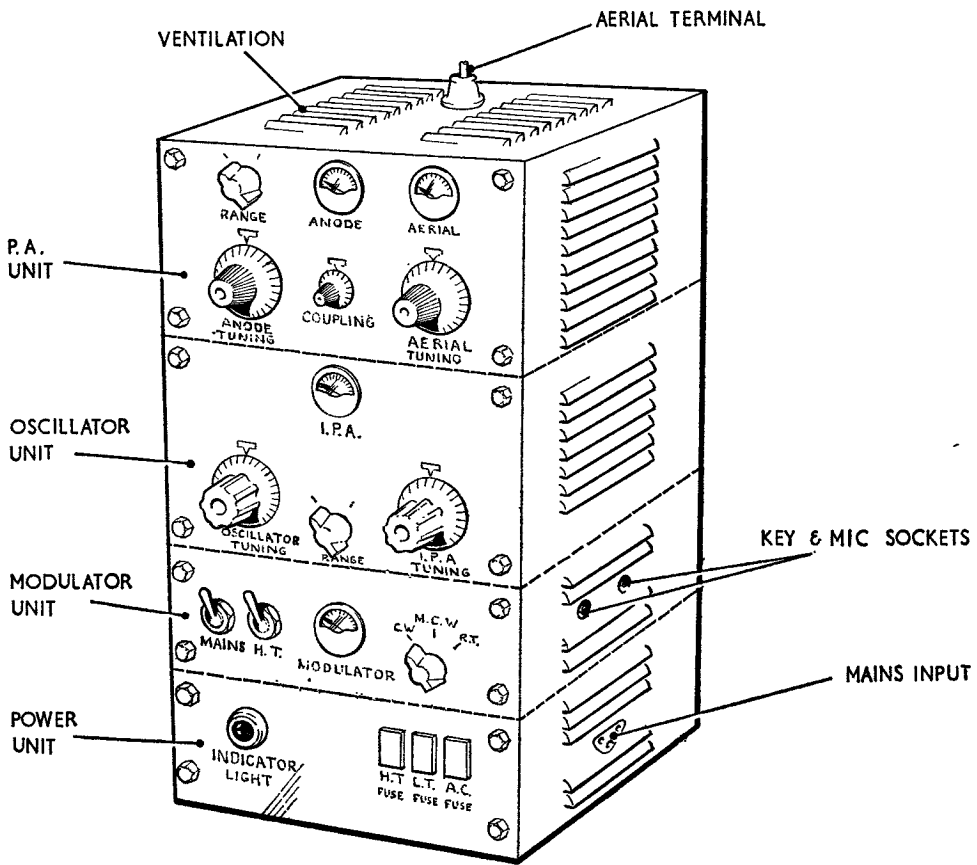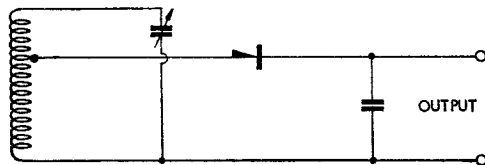
Introduction

So far in this section, transmitters which are suitable for operation in the l.f., m.f. and h.f. bands have been considered. These frequency bands were the first to be exploited and satisfactory communication was, and still is, obtained on them. However, with the increase in signal traffic during the last war, these frequency bands became overcrowded and strict control over the allocation of frequencies and the bandwidth of transmissions was imposed. Meanwhile radio engineers were investigating the possibilities of using what are now known as the v.h.f. (30 Mc/s to 300 Mc/s) and the u.h.f. (300 Mc/s to 3,000 Mc/s) bands. Many problems arose in the design of components and circuits which would be effective at these frequencies but eventually these problems were solved and nowadays v.h.f. and u.h.f. communication is commonplace in both Service and commercial applications.

The difficulties encountered arose from a number of basic causes, the main ones being:—

a. Stray capacitance.

b. Effective inductance of even short leads.

c. Unsuitability of conventional valves, including transit time effects.

d. Physical size of tuned circuit components.

e. Frequency drift.

These difficulties and the ways they were resolved will be considered in this chapter. It should be borne in mind, however, that the basic form of v.h.f. and u.h.f. transmitters is the same as described in Chapters 1 and 3.

Stray Capacitance

Capacitance exists between any two conductors separated by an insulator. Hence, inside a valve there must always be capacitance between the various metal electrodes. The anode/grid capacitance can be reduced by using a screen grid but it can never be eliminated. There is capacitance between each wire in a transmitter and the chassis, and of course between the wires themselves. The value of this capacitance depends on the size of the conductors and the distance between them. This value is normally very small, perhaps only a few pico-farads, but as the frequency of the energy in the wires and valves increases, so the *reactance* of the small capacitance decreases. For example, a capacitance of 1 pF would have a reactance of:

$$16 \text{ k at } 10 \text{ Mc/s}$$
$$1 \cdot 6 \text{ k at } 100 \text{ Mc/s}$$
$$\text{and } 160 \text{ ohms at } 1,000 \text{ Mc/s}.$$

This example shows that capacitance which may be ignored at 10 Mc/s must certainly be considered at 1,000 Mc/s. Suppose that C_{ak} of a transmitter valve is 1 pF and that the output voltage of the valve is 800V r.m.s. If the valve is operating at a frequency of 10 Mc/s C_{ak} will pass a current of 50 mA, which is probably negligible in comparison with the load current. At 1,000 Mc/s however, the current through C_{ak} would be 5 amps and the valve would be effectively short circuited by C_{ak}. This is illustrated in Fig. 1.

Thus in v.h.f. and u.h.f. transmitters it is most important to keep stray capacitances to a minimum. This is achieved by specially designed valves and components and by carefully
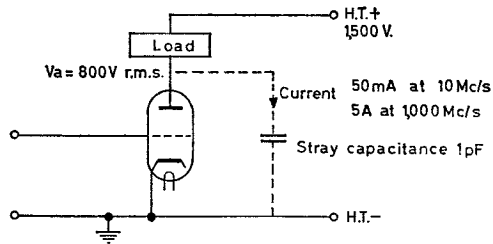
FIG. I. EFFECT OF STRAY CAPACITANCES

planned layout of components and wiring; the connecting leads are kept as short as possible, and positioned so that coupling between circuits (particularly output and input circuits) is at a minimum.   Any displacement or re-arrangement of wiring during servicing must be avoided.

Lead Inductance

A straight piece of wire a few inches in length has an inductance of roughly $0 \cdot 1$ μH, and its reactance would be:

<div align="center">

6 ohms at 10 Mc/s

60 ohms at 100 Mc/s

and 600 ohms at 1,000 Mc/s.

</div>

If this length of wire is used to connect two stages of a u.h.f. transmitter as shown in Fig. 2, and if 1 amp r.f. current flows along this wire from A to G, then the voltage drop between A and G will be:

<div align="center">

6V at 10 Mc/s

60V at 100 Mc/s

600V at 1,000 Mc/s.

</div>
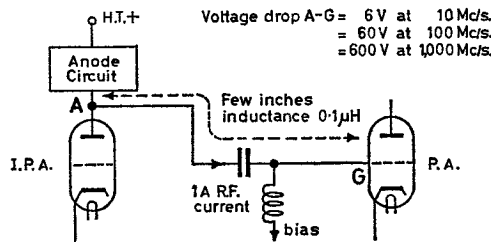
FIG. 2. EFFECT OF LEAD INDUCTANCE

From this example it is obvious that in order to avoid considerable losses the connections between stages in v.h.f. and u.h.f. transmitters must be kept as short as possible.

Link Coupling

In some v.h.f. and u.h.f. transmitters the drive unit and p.a. unit are mounted on separate chassis and short connecting leads are not possible.   In such cases the two units are connected by a length of transmission line coupled as shown in Fig. 3.
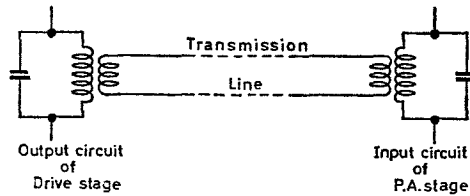
FIG. 3. LINK COUPLING

By correct matching at both ends, maximum power is transferred from the drive stage to the p.a. stage and losses are kept to a minimum.

Cathode Lead Inductance

A further problem, which arises when normal valves are used at v.h.f., is due to the inductance of the cathode lead. This lead extends from the actual cathode inside the valve to the external connection which in some types of circuit may be an earth connection (see Fig. 4).
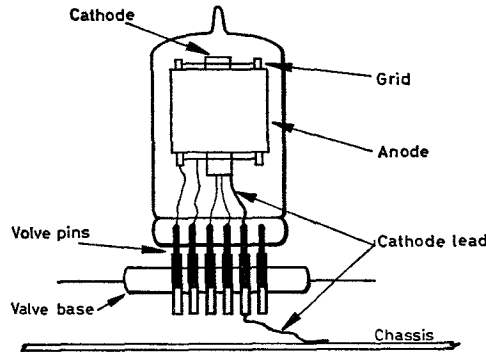
FIG. 4. CATHODE LEAD

The impedance of the cathode lead is quite appreciable at v.h.f. and if this lead is connected as shown in the circuit of Fig. 5, the impedance is in both the input and output circuits of the valve.

FIG. 5. CATHODE LEAD INDUCTANCE

38

The anode current of the valve flows through the cathode lead inductance across which a voltage is set up. This voltage is effectively between grid and cathode in series with, and opposing, the input voltage. Thus at v.h.f. and u.h.f. the effective gain of such a stage is reduced.

One method of reducing this adverse effect is shown in Fig. 6. The valve is constructed with two (or more) cathode leads, one of which is connected into the input circuit and the other

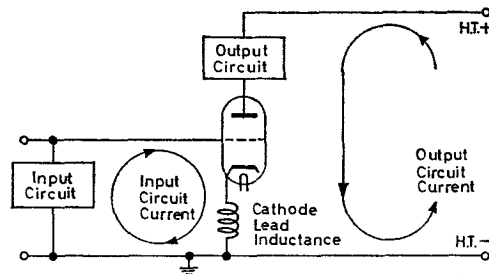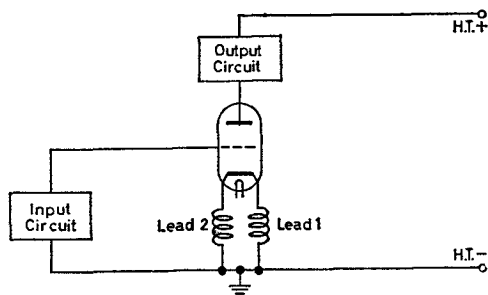FIG. 6. MULTIPLE CATHODE LEADS

into the output circuit. The inductances of the leads are therefore in parallel so that the total inductance is reduced.

Another solution is to use a grounded-grid triode circuit when the cathode lead inductance can form part of the tuned cathode circuit.

Transit Time Effect

In the valve actions so far considered it has been assumed that a change in grid voltage results in an immediate change in anode current. This assumption is quite permissible when considering frequencies below v.h.f., but at frequencies above 300 Mc/s the assumption is no longer valid. When the grid goes more positive, electrons leaving the cathode take a certain time to cross the inter-electrode space and reach the anode. This time is called transit time. Above 300 Mc/s the time of one cycle of the grid waveform is comparable to the transit time of the electrons, and a cloud of electrons attracted towards the anode by the positive-going half cycle of grid voltage may not pass the grid before the negative-going half cycle repels them.

When this happens three effects become apparent:—

a. The input impedance of the valve is reduced and more power is required to supply a given input voltage.

b. The amplification factor of the valve is reduced.

c. The normal phase relationship between grid and anode potentials is disturbed and, in an oscillator circuit, oscillations are difficult to maintain.

UHF Valves

Valves for use at u.h.f. must be designed to reduce the effects mentioned in the previous paragraphs. The valves are usually small with the electrodes spaced closely together, so reducing transit time. The electrons can be speeded up by using higher voltages, but arcing between electrodes must be avoided.

The electrode leads must be kept as short as possible and in some designs the leads form part of the tuned circuit.

Two types of valve used at v.h.f. and u.h.f. are illustrated in Fig. 7.
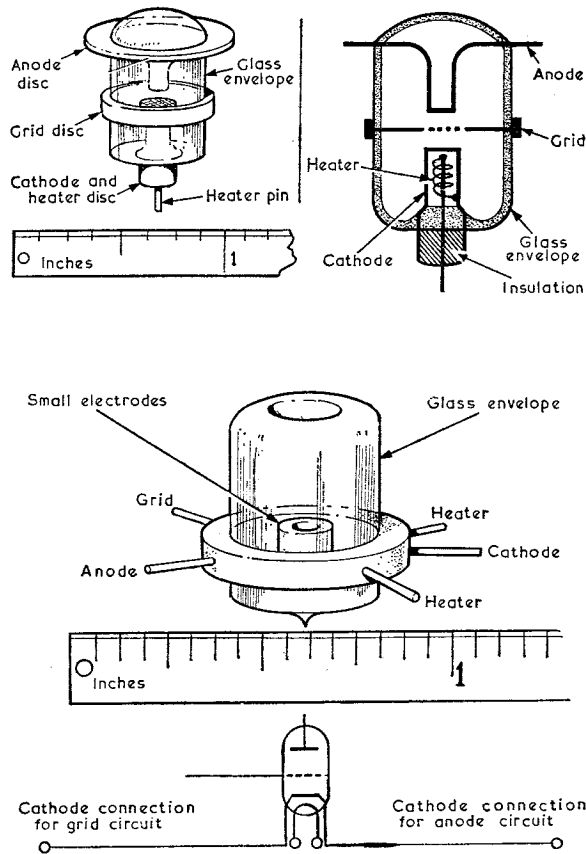
FIG. 7.   VHF AND UHF VALVES

Miller Effect

Feedback from anode to grid via the C_{ag} of a valve, known as Miller effect, can adversely affect an amplifier circuit.   It is most significant with triode valves where the C_{ag} is large.   At v.h.f. and u.h.f. the reactance of C_{ag} is low and feedback increases, tending to make an amplifier oscillate.   The grounded-grid triode circuit, already described, is one method of overcoming Miller effect but a greater input power is required.   Because of this the power gain at frequencies above 300 Mc/s may be as low as 4 or 5.

In order to obtain high power gains at v.h.f., special types of beam tetrode have been developed.   By forming the electrons into well defined beams the tetrode reduces the grid driving power required for a given output.   The power gain of such valves may be as high as 100. Furthermore, by careful alignment of grid and screen wires it is possible to keep the anode/grid capacitance very low indeed.

Tuned Circuits

At low frequencies tuned circuits can be constructed quite easily from coils and capacitors. The frequency at which the tuned circuit will resonate depends on the total values of inductance and capacitance, including "strays".   For example, if a circuit of resonant frequency 10 Mc/s

40

is required, a capacitance of 5 pF and an inductance of 50 μH could be used. The 5 pF could comprise a fixed capacitor of 4 pF and stray capacitances of 1 pF (Fig. 8). The 50 μH coil would consist of about 10 to 20 turns of wire and would be easy to construct.

FIG. 8. TUNED CIRCUIT WITH STRAY CAPACITANCE

For a resonant frequency of 100 Mc/s, with the same capacitance, the inductor would be 0·5 μH and for 1,000 Mc/s it would have to be 0·005 μH. Quite apart from the effect that the inductance of the connecting wires will have, these values are not feasible. A straight piece of wire 1 inch long has an inductance greater than 0·006 μH.

The upper frequency limit at which conventional inductors and capacitors can be used to form a tuned circuit is about 150 Mc/s. Above this frequency special techniques have to be used. These involve the use of short-circuited sections of transmission line (known as lecher bars) as the tuned circuit (Fig. 9).
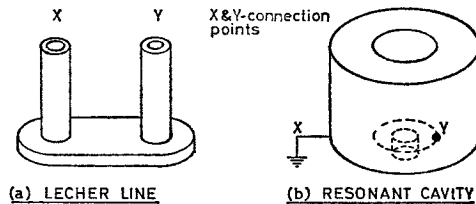

FIG. 9. VHF AND UHF TUNED CIRCUITS

The length of the lecher bar governs the resonant frequency and stray capacitance forms a part of the circuit. This form of tuned circuit is suitable for frequencies up to about 500 Mc/s. Between 500 and 3,000 Mc/s concentric line oscillators, as discussed in basic theory, are used in conjunction with disc-seal valves. At higher frequencies than 3,000 Mc/s, resonant cavities form the tuned circuits.

Parasitic Oscillations

As mentioned in Chapter 3, parasitic oscillations are caused by stray capacitance and inductance forming an unwanted tuned circuit. In v.h.f. and u.h.f. equipments parasitic oscillations can quite easily occur, especially in the frequency multiplier stages, unless precautions are taken to avoid them. These precautions consist of carefully planned layout of chassis components and wiring, and the use of anti-parasitic resistors where they are necessary.

Frequency Stability

In any transmitter it is important to keep the radiated frequency as stable as possible in order to avoid interference with other transmissions and to make as many channels as possible available in a given frequency band. This is especially so in v.h.f. and u.h.f. transmitters, since a frequency stability of 1 part in 10,000 at 300 Mc/s means a possible frequency drift of 30 Kc/s. Thus to avoid interference between two transmissions they would have to be spaced 60 kc/s apart.

41

In a v.h.f. transmitter using a frequency multiplication factor of 18, a drift in the master oscillator frequency of 100 c/s results in a drift of 1·8 kc/s in the radiated frequency. This means that high quality crystals and other components must be used, with possibly a crystal oven, in order to maintain reasonable frequency stability.

A UHF Airborne Transmitter

The block diagram of Fig. 10 shows a typical u.h.f. r.t. transmitter suitable for ground-to-air and air-to-air communication.
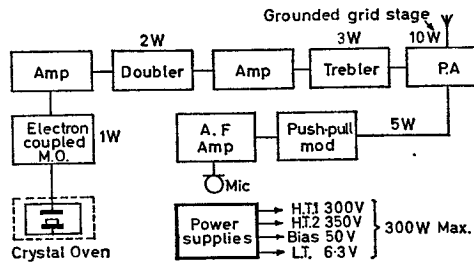
FIG. 10. BLOCK DIAGRAM OF A UHF AIRBORNE TRANSMITTER

The m.o. is crystal controlled and the crystal is housed in an oven to increase frequency stability. The output frequency range is about 200 to 400 Mc/s. This is obtained by multiplying the fundamental crystal frequency 18 times at the lower end of the band and 24 times at the top end: crystal frequencies required are therefore between 11 and 18 Mc/s. The anode circuit of the m.o. is tuned to either the third or fourth harmonic of the crystal. Because the power contained in the harmonics is small, amplifiers are placed between each multiplier stage.

The communication range of such a transmitter would depend upon the height of the aircraft and would be about 120 miles at 10,000 feet.

The external details are illustrated in Fig. 11: the size of the case is about 2ft. by 2ft. by 1ft. and the total weight is about 150 lbs.

FIG. 11. TYPICAL UHF AIRBORNE TRANSMITTER

CHAPTER 5

MULTICHANNEL TRANSMITTERS

Introduction

The pilot of a modern high speed aircraft needs to communicate with many ground stations operating on different frequencies. Airborne v.h.f. and u.h.f. equipments provide for this need by en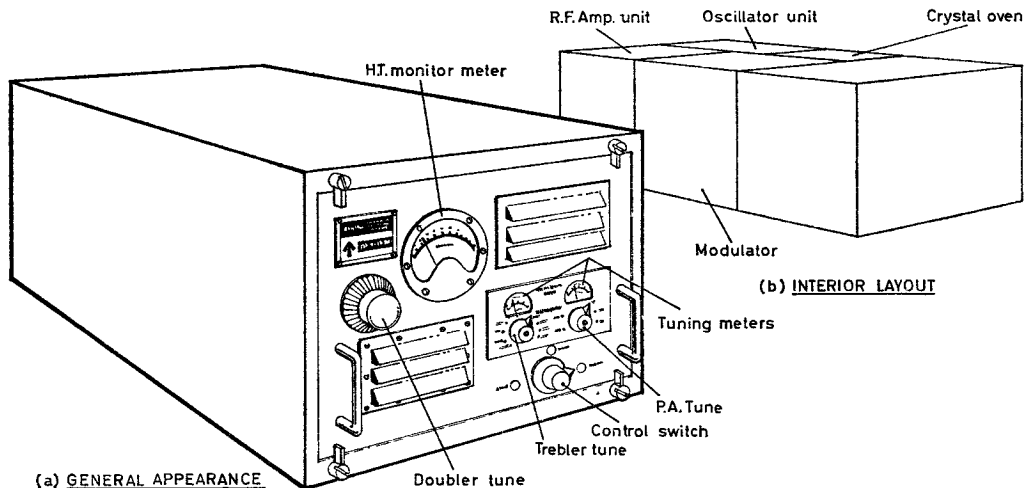abling the pilot to select any one of a number of pre-set frequencies merely by pressing a button or turning a switch. Such a transmitter is called a *multichannel* transmitter since a number of communication channels can be pre-set and any one selected quickly.

Crystal Control

With multichannel v.h.f. and u.h.f. equipments it is important that when a change of frequency takes place the new frequency is correct. For this reason the m.o. of the transmitter is usually crystal controlled. There are two ways in which crystal control of a multichannel transmitter can be achieved.

a. Direct control of the m.o. frequency by means of a crystal: this means using one crystal for each channel and, in a transmitter designed to cover a frequency band of 50 Mc/s with 100 kc/s channel separation, 500 crystals must be available.

b. Indirect control of the m.o. frequency by comparing it with a frequency obtained from a number of crystals. For example a 100 kc/s crystal can be used with a 6 Mc/s crystal to check 6·1 Mc/s. It can also be used with a 7 Mc/s crystal to check 7·1 Mc/s. In this way each crystal can be used more than once and, in a transmitter covering a frequency band of 175 Mc/s with 100 kc/s channel separation, 32 crystals will provide sufficient combinations to control all 1,750 channels.

Direct Crystal Control of MO Frequency

With this system a push-button or rotary switch is used to control a motor. The movement of the motor shaft selects the required crystal and rotates the tuning controls of the frequency multiplier and p.a. stages to the required positions. One method of doing this has already been discussed in Chapter 2, and Fig. 1 shows an alternative method.

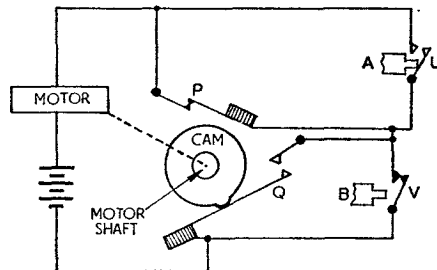FIG. I. POSITION SELECTOR TUNING SYSTEM

To simplify the explanation the diagram shows only two switch positions, but in practice there can be as many as 20. When the motor is energised it rotates a cam, which opens contacts P and Q in turn. If button A is pressed it opens switch U and closes switch V; the motor circuit is completed via P and V and the motor shaft rotates the cam until contact P is opened, so

breaking the supply to the motor. It is a simple matter to have the motor shaft operate other switches which connect the required crystal to the oscillator circuit.

Further cams on the motor shaft can be arranged to operate push-rods coupled to the tuning controls of the frequency multiplier and p.a. stages. A system of set-screws and lock-units can be incorporated so that, when the transmitter is initially tuned on the test-bench to each channel frequency, the cams and push-rods can be adjusted to enable the settings to be reproduced automatically.

With this system, up to 20 channels can be controlled simply by connecting extra push-buttons and extra cam contacts in series with the ones shown in Fig. 1.

Indirect Crystal Control of MO Frequency

In this method of control the transmitter uses a variable-frequency m.o. The output frequency of the m.o. is compared with selected crystal frequencies by means of mixer circuits followed by filters. The m.o. frequency is then corrected until it is at exactly the required frequency. The principle is best illustrated by means of a numerical example.

Assume that the frequency band covered by the m.o. is 12 to 17 Mc/s. The m.o. tuning control is turned by an electric motor so that when no channel is selected the control swings the m.o. frequency backwards and forwards through the band. When a particular frequency in the band is required, a frequency comparison circuit is used to indicate when the tuning control is in the correct position. A basic frequency comparison circuit is shown in Fig. 2.
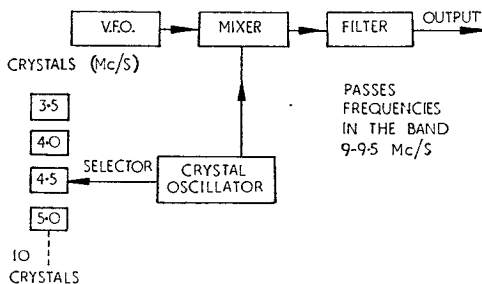
FIG. 2. BASIC FREQUENCY COMPARISON CIRCUIT

If the m.o. is required to oscillate on a frequency between 13·5 and 14 Mc/s the 4·5 Mc/s crystal is selected by the push button. As the m.o. sweeps through 13·5 to 14 Mc/s, the output from the crystal oscillator will mix with the m.o. output to give a difference frequency of between 9·0 and 9·5 Mc/s. The filter following the mixer stage is designed to pass frequencies between 9·0 and 9·5 Mc/s and will reject all other frequencies. Thus an output from the filter is obtained only when the m.o. is oscillating at a frequency between 13·5 and 14 Mc/s. This output can be used to control the oscillator tuning motor so that it rocks the oscillator frequency between 13.5 and 14 Mc/s.

Thus, using this comparison circuit, the m.o. can be set to a frequency within a 0·5 Mc/s interval. By using a further set of crystals with a second crystal oscillator, a mixer and a filter with a bandwidth of 0·05 Mc/s it is possible to set the m.o. to a frequency within a 0·05 Mc/s interval within the original 0·5 Mc/s interval. It can be seen that any desired accuracy of oscillator setting can be obtained by using a sufficient number of sets of crystals with the necessary filter and mixer.

Table 1 illustrates the principle of indirect crystal control. In this table it is assumed that the second set of crystals, the mixer and filter selects the interval 13·6 to 13·65 Mc/s.

```
Mc/s                              Mc/s
13·00 ⌐                           13·00 ⌐
      |                           13·05 |
      |                           13·10 |
      |                           13·15 |
      |     NO OUTPUT             13·20 |
      |     FROM FIRST            13·25 |     NO  OUTPUT
      |     FILTER               13·30 |     FROM SECOND
      |                           13·35 |     FILTER.
      |                           13·40 |
      |                           13·45 |
13·50 ⌐                           13·50 |
                                  13·55 ⌐
                                  13·60 |     THIS INTERVAL
             THIS INTERVAL        13·65 |     SELECTED BY
             SELECTED BY          13·70 |     SECOND FILTER
             FIRST FILTER         13·75 |
             AND ASSOCIATED       13·80 |
             CIRCUITS.            13·85 |     NO  OUTPUT
                                  13·90 |     FROM SECOND
                                  13·95 |     FILTER.
14·00 ⌐                           14·00 |
      |     NO OUTPUT             14·05 |
      |     FROM FIRST            14·10 |
      |     FILTER.               14·15 ⌐
```
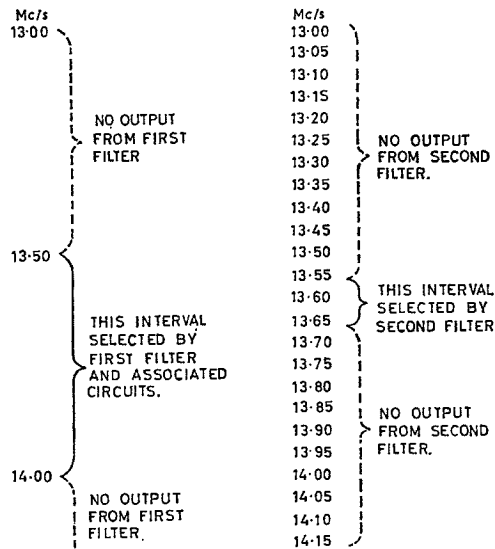
TABLE I. PRINCIPLE OF INDIRECT FREQUENCY CONTROL

If, due to changes in temperature, or for any other reason, the m.o. frequency tends to drift away from the correct frequency as determined by the selected crystals of the frequency control system, the m.o. tuning motor is automatically caused to rotate in such a direction that the tendency for the frequency to drift is counteracted. Thus the accuracy of the m.o. frequency is determined by the frequency accuracy of the crystals and is in no way affected by random variations in the m.o. circuit.

In a practical multichannel equipment the m.o. is followed by frequency multiplier stages which provide a final frequency in the v.h.f. or u.h.f. band.

TRANSMITTER KEYING

Introduction

The basic transmitter considered in Chapter 1 would transmit a constant frequency, constant amplitude r.f. wave. Such a wave would convey no information to a receiver; all the receiver could do would be to estimate the frequency and power of the transmitter.

To convey information it is necessary to alter the transmitter output in some way. One method is to switch the transmitter output on and off to transmit messages in Morse code. This switching process is known as *keying* the transmitter.

One way of interrupting the transmitted output would be to switch the master oscillator on and off. This method is not normally employed since it may cause the output frequency to drift. The m.o. is therefore allowed to oscillate continuously without interruption and keying is carried out in one of the p.a. stages following the m.o. Sometimes two or more stages are keyed simultaneously. When ON/OFF keying is employed with a simple transmitter the signal sent is referred to as continuous wave (c.w.). This indicates that if the key is kept in the ON position the transmitter will send a continuous wave of constant amplitude and frequency.

Simple Keying Circuits

The p.a. stage of a transmitter can be keyed by one of two basic methods or by a combination of both:—

a. The anode current can be switched on and off directly by means of a switch in the h.t. circuit.

b. The anode current can be switched on and off indirectly by switching a suitable bias to an electrode between cathode and anode. This electrode may be control grid, screen grid or suppressor grid.

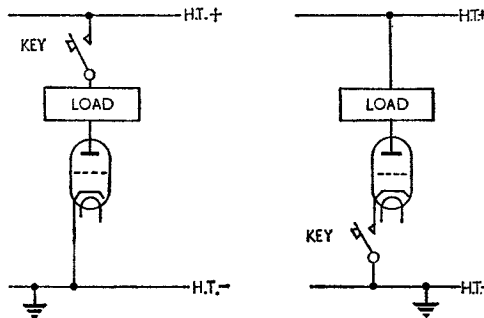The simple keying circuits of Figs. 1 to 3 are really self explanatory but certain features merit attention.

FIG. I. KEYING THE HT SUPPLY

HT Keying

The key interrupts the h.t. supply to the p.a. stage. Since the h.t. supply might be quite considerable even in medium power transmitters, the switching is usually done by a relay operated by a keyed, low-voltage d.c. supply.

46

If the relay contacts are connected into the cathode circuit of the valve the cathode circuit and the grid circuit are both interrupted when the key is raised. This is a fairly rapid method of keying which is used in low-power transmitters.
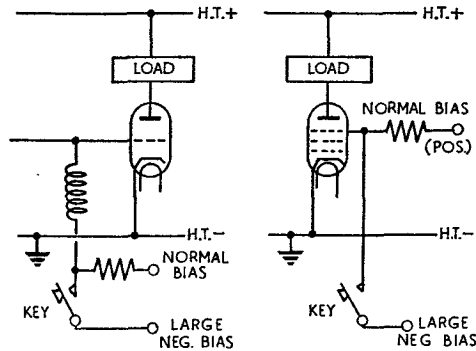

FIG. 2. BIAS KEYING

Bias Keying

With this method of keying (Fig. 2) the key is in a relatively low power circuit but bias voltages of 1 kV or more may be used. The large voltage is needed to break the transmission because it must be sufficient to cut off the valve even when the drive voltages are applied: a separate power supply unit is often required to provide this keying bias.

With suppressor grid bias keying, the anode current is cut off from the anode circuit and therefore it must flow to the screen circuit. There is quite a low limit to the current which can be accepted by a screen grid, otherwise it gets too hot, and therefore this arrangement can be used only in low power p.a. stages.
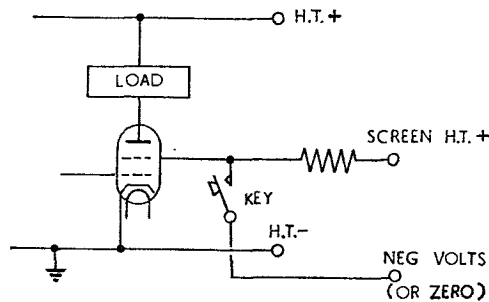

FIG. 3. SCREEN KEYING

Screen Keying

The anode current can be cut off by making the screen grid negative or in some p.a. stages simply by making the screen voltage zero (Fig. 3). Since the screen current is always much lower than the anode current, the screen circuit is much easier to key. In some transmitters the anode *and* screen of the keyed p.a. stage are switched simultaneously by using a relay with contacts in the anode and screen circuits.

47 *(A.L. 1, June, 1964)*

Absorber Valve Keying

In this arrangement, when the p.a. valve current is cut off, a separate valve is switched on to "absorb" the d.c. power which was taken by the p.a. This has the advantage of maintaining a practically constant load on the power supply which aids power supply regulation (a subject dealt with in Chapter 10), but it is of course uneconomical. The advantage of absorber valve keying lies in the high degree of m.o. stability obtained by good power regulation. Fig. 4 shows the system in outline.
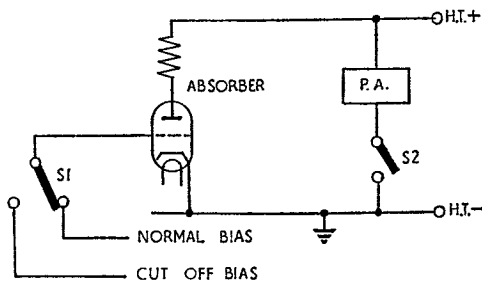

FIG. 4. ABSORBER VALVE KEYING

Switches S1 and S2 operate together by means of keyed relays. The absorber valve is so biased that when the key is open the absorber takes the same current as that normally taken by the amplifier when the key is closed. With the key closed the absorber valve is cut off.

Key Filters

In all keying arrangements the act of keying involves the interruption of current flow. Therefore sparking inevitably occurs and may result in radio receiver interference. To avoid this a simple LC filter is included in the keying circuit and thus the interference is suppressed. Fig. 5 shows a key filter with typical values for L and C.
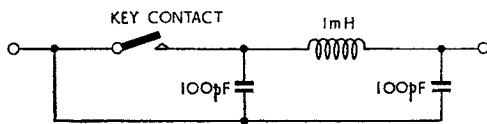

FIG. 5. KEY FILTER

Another type of filter circuit is often provided to eliminate the sharp corners of the "natural" keying waveform. This filter reduces the bandwidth of the transmitter signal thereby eliminating interference in the form of clicks which would otherwise be experienced in nearby receivers even if they are tuned to different frequencies.

Fig. 6 shows such a filter (known as a key click filter) in a cathode keying circuit; radiated waveforms for the morse code letter A, with and without the filter are also shown.

When the key is closed, valve current rises slowly through L as the back e.m.f. dies away. When the key is opened the current dies away slowly as C charges. Thus the action of the filter is to slow down the rise and fall of circuit current and so the output waveform has rounded corners. This means there are no sudden changes of amplitude and therefore less risk of causing interference.
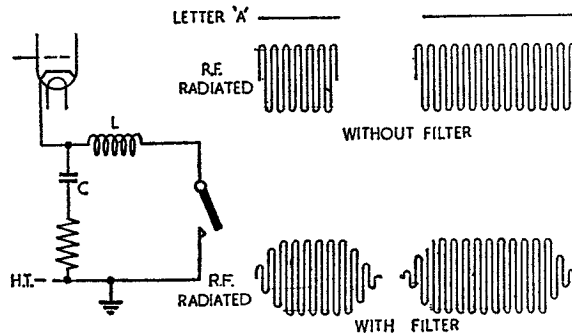
FIG. 6. KEYING WAVEFORMS

A disadvantage of a key click filter is that it places an upper limit on the keying speed. The effect of the filter is to increase the time taken by the transmitter to reach full power when the key is closed and to return to zero when the key is opened. Fig. 7, which shows the output from a monitor used to detect the transmitter output, illustrates the effect of the key click filter at different keying speeds.
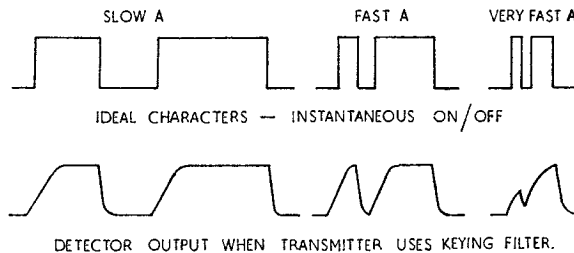
FIG. 7. EFFECT OF KEYING FILTER AT DIFFERENT KEYING SPEEDS

Practical Keying Arrangements

In most cases it is not convenient to key the transmitter directly. With ground installations the transmitter is normally sited near the aerial and the operators are positioned in a building several hundred yards away. Also, several circuits in the transmitter may have to be keyed and this is best done with multi-contact relays. Fig. 8 shows the basic arrangement for a d.c. remote relay circuit.
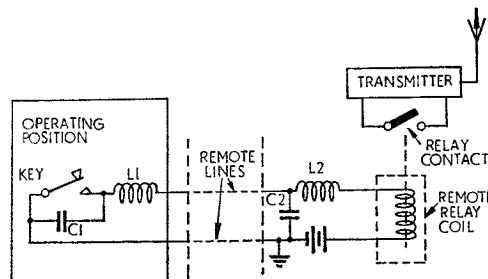
FIG. 8. REMOTE RELAY KEYING

When the key is up the relay is de-energised and the keying relay contacts are open. When the key is down the relay circuit is completed via the remote lines and the keying contacts are closed.

$L_1 C_1$ and $L_2 C_2$ are r.f. filters to isolate (in terms of r.f.) each end of the line. Note that any key click filter will be contained in the transmitter and connected across the *relay contacts*, not across the key itself.

Listening-through (Break-in)

The use of a multi-contact relay to key the transmitter enables the key to perform several functions simultaneously.

In many transmitter-receiver installations the same aerial is used for both transmitter and receiver. To avoid damage to the receiver circuits which may result from applying the transmitter output directly to the receiver input, the keying is arranged to switch the aerial from one to the other. A simple keying relay system is shown in Fig. 9.
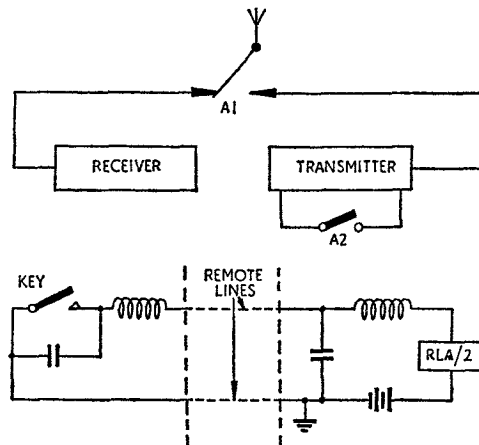

FIG. 9. LISTENING-THROUGH CIRCUIT

The relay designation RLA/2 indicates that the relay has two contacts, A1 and A2. With the key up the relay is de-energised and contact A1 connects the aerial to the receiver: contact A2 breaks the keying circuit of the transmitter. With the key down the relay is energised and contact A1 moves over to connect the aerial to the transmitter output: contact A2 keys the transmitter.

Thus during the interval between the dots and dashes of the Morse code, the receiver is fully operative and it is possible to listen for other signals through the gaps in transmission. In a long transmission this allows the distant receiving station to interrupt the transmission and request a repeat of any parts of the message which may have been missed.

Sidetone

In some transmitters, provision is made for the operator to listen to the morse symbols he is transmitting. To provide this, an a.f. oscillator is keyed with the main transmitter. Extra keying contacts switch the operator's headset from the receiver output to the output of the a.f. oscillator when the key is pressed. This facility is known as "sidetone".

TRANSMITTER AMPLITUDE MODULATION

Introduction

The basic principles of modulation have been outlined and modulation was defined as the process of varying the carrier wave in accordance with the modulating signal. The r.f. energy generated by an oscillator and amplified by p.a. stages *cannot by itself* convey intelligence; to be of any use as a means of communication it must be varied in some way. That is, it must be *modulated*. To modulate the carrier wave, any one of three characteristics of the carrier can be varied; the amplitude, the frequency or the phase. In this chapter the practical effects and methods of *amplitude* modulation will be considered.

MCW and Sidebands

The two important features of an amplitude modulated wave are:—

a. The amplitude of the r.f. carrier varies at the audio modulating frequency.

b. The transmitted "signal" contains sidebands.

Thus if a carrier of 1 Mc/s is modulated by a single a.f. of 10 kc/s, the result will be the production of two new r.f. outputs, (1 Mc/s + 10 kc/s) and (1 Mc/s — 10 kc/s) in addition to the 1 Mc/s carrier. Fig. 1 shows the idea in outline.
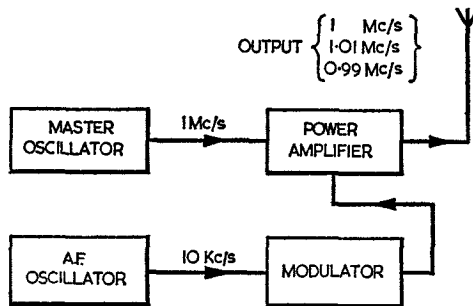

FIG. 1. MODULATION

The transmitted signal really consists of three frequencies, the carrier and two side frequencies as shown in Fig. 2. The frequency (1 Mc/s + 10 kc/s) is termed the upper sideband (u.s.b.) and the frequency (1 Mc/s — 10 kc/s) is termed the lower sideband (l.s.b.).
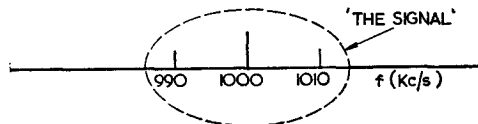

FIG. 2. MCW SIGNAL FREQUENCIES

In order to handle such a signal without introducing distortion, the transmitter p.a. stage and the receiver tuned circuits must have sufficient bandwidth to respond equally to all the

"signal" frequencies. Notice that since the three frequencies are r.f., *individually* they convey nothing: each is merely an unmodulated r.f. oscillation. This fact is made clear in Fig. 3 which compares the two forms of presentation of an amplitude modulated carrier.
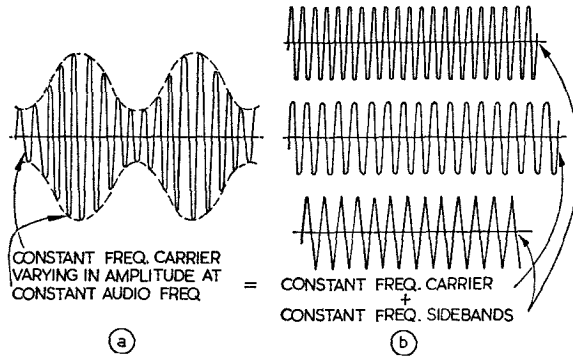

CONSTANT FREQ. CARRIER VARYING IN AMPLITUDE AT CONSTANT AUDIO FREQ = CONSTANT FREQ. CARRIER + CONSTANT FREQ. SIDEBANDS

(a)　(b)

FIG. 3.　MODULATION WAVEFORMS

The three constant frequency, constant amplitude waveforms of Fig. 3b combine to form the "signal". Sometimes the sidebands and carrier are in phase; sometimes they are out of phase; and so the amplitude of the resultant waveform changes. The composite signal waveform is shown in Fig. 3a.

RT Sidebands

The only difference between MCW and speech (or any complex sound) modulation is that there are no *fixed* sidebands in speech modulation. Speech consists of a range of frequencies continuously varying and so there is a range of sideband frequencies, also continuously varying. In practice the range of speech frequencies required to transmit intelligible speech is taken to be about 300 to 3,000 c/s and so the resultant sidebands range from (carrier minus 3,000 c/s) to (carrier plus 3,000 c/s) i.e. the bandwidth occupied by the signal is:—

$$(f_c + 3,000) - (f_c - 3,000)$$
$$= 6,000 \text{ c/s}$$
$$= 2 \times \text{maximum audio frequency.}$$

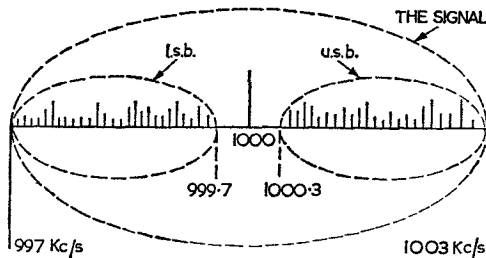This is shown in Fig. 4, using a carrier frequency of 1 Mc/s.

FIG. 4.　RT SIGNAL FREQUENCIES

In this case there are no longer just the three frequencies. The "signal" is in fact a band of frequencies covering 6 kc/s centred on 1 Mc/s.

Sidebands Produced When a Transmitter is Keyed

When a c.w. transmitter is keyed the bandwidth of the signal transmitted depends upon the shape of the waveform produced. If the keying circuits are capable of switching the transmitter on and off very rapidly, a wide range of sideband frequencies is produced and key click interference will be heard in receivers which are tuned to frequencies well away from the transmitter frequency. Therefore, as explained in Chapter 6, the pulse waveform is given a rounded shape similar to an a.f. waveform so as to limit the extent of the sidebands and avoid interference.

Modulation Depth

The process of amplitude modulation alters the amplitude of the carrier wave in accordance with the audio signal. As the audio signal varies in strength (for example, as speech becomes louder or softer), so the degree to which the carrier is affected varies.

The influence of the modulation on the carrier is expressed in terms of *modulation depth* which is also called *modulation factor*. The modulation depth is usually denoted by the symbol m and, for modulation by a single sinusoidal audio frequency wave, is defined as indicated in Fig. 5.
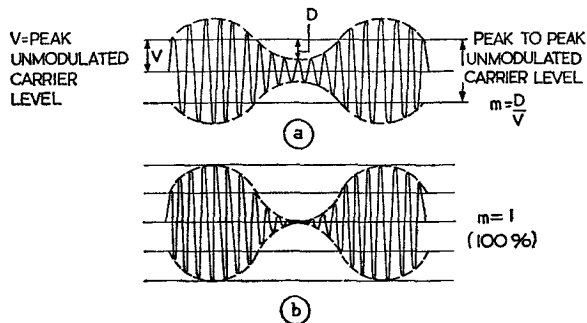

FIG. 5. MODULATION DEPTH

Usually modulation depth is expressed as a percentage but it can also be written as a decimal. For example m = 0·75 means that the modulation depth is 75%.

The maximum depth of modulation which can be achieved without introducing distortion is 100%. In this case the modulated wave goes down to zero amplitude on the negative peaks of the modulation. A depth of 100% corresponds to the "loudest sound" (in the microphone) which can be transmitted.

Power Distribution

In the case of a simple m.c.w. signal, i.e. carrier plus l.s.b. and u.s.b., the total radiated power is made up of three parts:—

  a. Carrier power.

  b. u.s.b. power.

  c. l.s.b. power.

If the amplitude of the unmodulated carrier is V volts, then for a modulation depth of 100% the modulation will cause the transmission to swing between 0 and 2V volts. The 2V volts will

occur when the carrier and both sidebands are momentarily in phase: the 0 volts will occur when the two sidebands are in opposition to the carrier. Thus the amplitude of each sideband is $\frac{V}{2}$ volts $\left(\frac{V}{2} + \frac{V}{2} + V = 2V\right)$ at 100% modulation.

Since power is proportional to V^2 (or to I^2) the radiated power for 100% modulation (Fig. 5b) is given by

$$\text{Total Power} = \left[V^2 + \left(\frac{V}{2}\right)^2 + \left(\frac{V}{2}\right)^2\right] \div R_L$$

$$= V^2 (1 + \tfrac{1}{2}) \div R_L.$$

In this expression V is the r.m.s. carrier voltage and R_L is the load resistance in which the modulated signal is developed.

If the carrier is unmodulated, the power output will be simply:—

$$\text{Carrier Power} = V^2 \div R_L.$$

Thus each sideband adds power equal to a quarter of the original power, and together they constitute the fraction $\frac{1}{2}$ of the total $1\frac{1}{2}$: i.e. $\frac{1}{3}$ of the total output power with 100% modulation. This sideband power is supplied by the transmitter h.t. and therefore more current is drawn from the power supply when the carrier is modulated.

It should be noted that the signal readability depends on the power contained in the sidebands, and not on the carrier power. Because of this, it is desirable to have as large a percentage modulation as possible. For modulation percentages less than 100 (modulation factor less than 1), the fraction of power in the sidebands is small. For example, if the modulation depth is 30% (modulation factor $0 \cdot 3$), then:—

$$\text{Total Power} = \left[V^2 + \left(\frac{0 \cdot 3V}{2}\right)^2 + \left(\frac{0 \cdot 3V}{2}\right)^2\right] \div R_L$$

$$= V^2 \left(1 + \frac{0 \cdot 09}{2}\right) \div R_L$$

$$= V^2 (1 + 0 \cdot 045) \div R_L.$$

In this case less than 5% of the total power is contained in the sidebands.

Although it may seem easy always to use 100% modulation this is possible only in the case of m.c.w. In the case of r.t., the amplitude of the modulating signal varies continuously and it is possible to specify only the average modulation percentage. Thirty per cent is in fact a high *average* r.t. modulation level, although in some Service communication systems special measures are adopted to raise this level and thus increase the readability of the signal.

High and Low-Level Modulation

Modulation is said to be high or low level depending on whether it is applied in the transmitter at a point of high or low r.f. power. This is illustrated in Fig. 6. If the modulation is applied to an early stage (low-level modulation) the following p.a. stages must be operated in class A or B in order to avoid distorting the modulated r.f. input to the stage. Class B amplification of a modulated wave often employs push-pull amplifiers but a single valve amplifier could be used since the oscillatory action of the tuned circuit reproduces the correct waveform. An a.f. class B amplifier on the other hand, since its load would not be a tuned circuit, must employ push-pull connected valves.
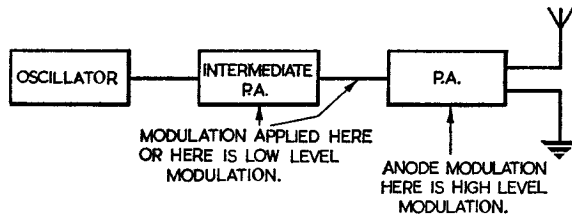
FIG. 6. LOW- AND HIGH-LEVEL MODULATION

For a given total output power low-level modulation would need very little a.f. power but a large p.a. stage. These stages, because they are operated in class B, would be inefficient but to offset this disadvantage the a.f. stages would be small and relatively light since there is no need for heavy high power a.f. transformers.

High-level modulation, where the modulation is applied at the final p.a. stage, enables efficient class C amplifiers to be used in the r.f. stages but requires bulky high power inefficient a.f. amplifiers.

Modulation Requirements

The previous calculations emphasise the importance of the modulator in terms of transmission efficiency. In fact it is a waste of time (and energy) to generate a powerful r.f. carrier unless it can be adequately modulated.

The additional power in the modulated signal of a high-level modulated transmitter is supplied by the modulator. Thus the modulator must be a large power amplifier. Since the r.f. p.a. is the "load" into which the modulator works, the impedances of the r.f. p.a. and the modulator must be matched for maximum efficiency. These points are emphasised in Fig. 7 which shows the modulator as an a.f. power amplifier matched into the r.f. power amplifier.
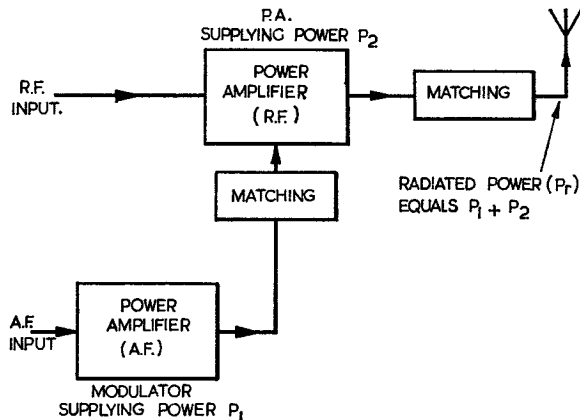

FIG. 7. AF MODULATOR AND RF POWER AMPLIFIER

Since the original r.f. and a.f. inputs both provide very low power, the radiated power must be provided by the p.a. and modulator stages i.e. $P_r = P_1 + P_2$. Matching between modulator and p.a. is therefore just as important as matching between p.a. and aerial.

The matching arrangement is usually part of the amplifier circuit. In the case of the modulator, which is an a.f. circuit, matching is easily achieved by an a.f. transformer. Since the p.a. is an r.f. circuit the matching arrangement may be less obvious.

Notice that since modulation provides part of the total power output, the reading in an aerial current meter will indicate modulation. An estimate of the depth of modulation m, can be made from the aerial current readings but unfortunately the method is limited to large modulation factors, since a modulation depth of even 100% will give only about 20% increase in aerial current.

Transmitter Modulators

The basic circuits for modulating a transmitter have already been discussed. They can all be summarized as "circuits designed to control the amplitude of r.f. output by means of a varying voltage on a valve electrode", and are classified as high or low level modulating circuits as already explained.

In the case of a pentode p.a. valve, there are four possible control electrodes, as indicated in Fig. 8.
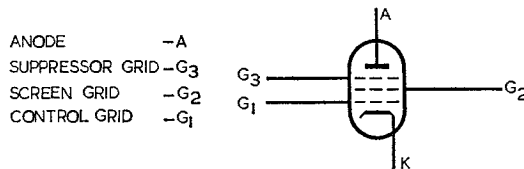

FIG. 8. PENTODE ELECTRODES

Cathode modulation is really a combination of modulating all the other electrodes.

Thus variation of the power supply to any of the electrodes will vary the anode current and can therefore modulate the output. Variation of the h.t. to anode or screen requires appreciable a.f. input power. Variation of the bias voltage to suppressor or control grid requires little a.f. power input.

A modulator must impress the audio signal onto the carrier so as to produce a modulated output, i.e. an output which contains sidebands. Simply adding the a.f. and the r.f. carrier in the same circuit does not produce a modulated output. The two inputs must be properly mixed. (Fig. 9).
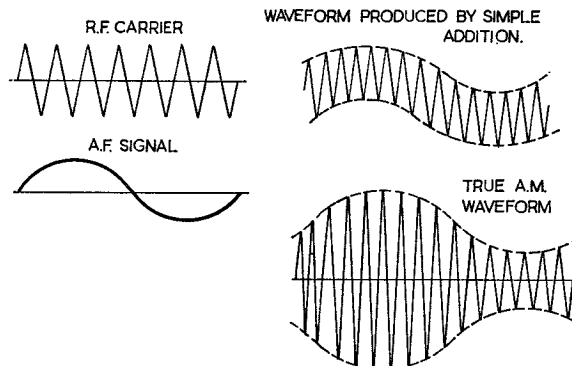

FIG. 9. MODULATION WAVEFORMS

Anode Modulation

High-level anode modulation of the final p.a. stage is often employed in r.t. transmitters. Since high powers (r.f. and a.f.) are involved, matching between the modulator and p.a. stage is very important: it is achieved by using a modulation transformer.
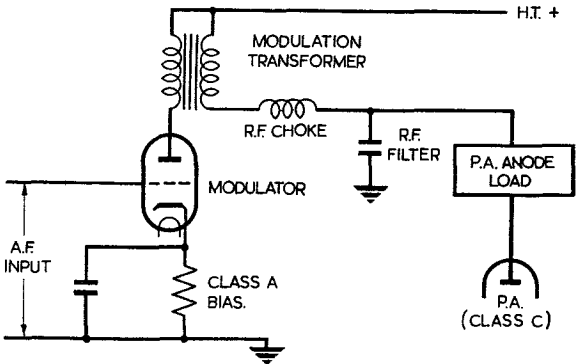

FIG. 10. SIMPLE ANODE MODULATION

The function of the modulator is to vary the h.t. supply to the p.a. stage so that the r.f. output from the p.a. stage varies in accordance with the a.f. input. The a.f. power necessary to do this is about half the magnitude of the power taken by the p.a. If the p.a. stage requires 10W d.c., or less, the simple circuit of Fig. 10 could be used, with the modulator stage working in class A. The p.a. stage would of course be very low power.

In medium and high power transmitters the r.f. power will be high and so will be the modulating power required. For example, a p.a. stage consuming 2 kW d.c. (almost 3 horsepower) will need about 1 kW a.f. power from the modulator. Such a modulator would be a high power a.f. power amplifier and its efficiency would be important. To obtain a reasonable efficiency the modulator would have to work in class B push-pull. Thus medium and high power r.t. transmitters employ class B push-pull modulators with anode modulation.

The p.a. stage may employ beam tetrodes or pentodes: if this is so the modulator output will have to be applied to the screen grids as well as to the anodes, since variations in the anode
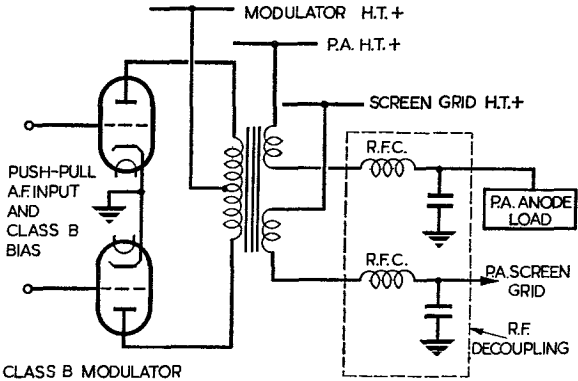

FIG. 11. HIGH POWER ANODE AND SCREEN MODULATION

voltage alone would have little effect on the pulses of anode current. All that is required is an extra winding on the modulator transformer to supply the screen voltage. A typical circuit is shown in Fig. 11. The p.a. stage may consist of a number of valves in parallel and/or push-pull, but this does not affect the a.f. circuit.

Screen Grid Modulation

Screen grid modulation is used only in conjunction with anode modulation as outlined in the paragraph headed "Anode Modulation". If used alone screen grid modulation would not produce a large depth of modulation without introducing severe distortion.

Grid Modulation

Grid modulation requires a very low modulating power compared with anode modulation because it makes use of the gain of the p.a. stage after modulation has been added. Unfortunately, high modulation depths cannot be achieved without appreciable distortion. Grid modulation is therefore not used a great deal.
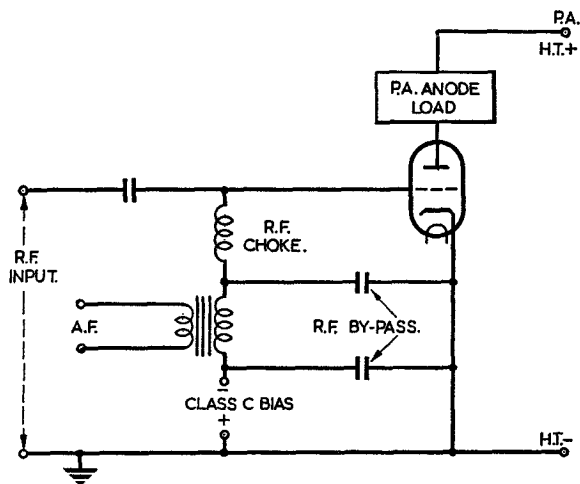

FIG. 12. GRID MODULATION

As shown in Fig. 12, the a.f. input is applied in series with the d.c. bias and the r.f. input to the p.a. stage. The grid voltage is thus the sum of the three inputs d.c., a.f. and r.f., and the valve acts as a class C r.f. amplifier whose "bias" voltage varies at a.f. As a result, the anode current, and therefore the r.f. output, is amplitude modulated.

As already mentioned, grid modulation has the advantage of needing comparatively little a.f. power and therefore the modulator can be small and light. However, the very fact that two varying inputs are applied to the same grid means that the maximum r.f. input is limited by the amplitude of the a.f. input. Fig. 13 shows the grid waveforms for unmodulated and modulated conditions.

In the unmodulated condition the maximum r.f. amplitude is approximately equal to the d.c. bias.

In the modulated condition the maximum r.f. amplitude is approximately equal to the bias minus the a.f. peak amplitude. Thus, with grid modulation, the maximum r.f. input is reduced.
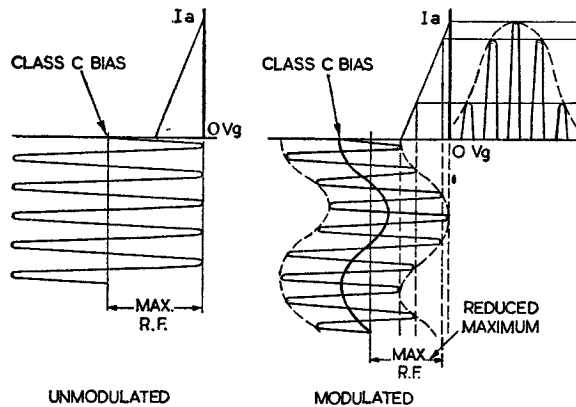
FIG. 13. GRID MODULATION WAVEFORMS

Suppressor Grid Modulation

If the p.a. valve is a pentode, modulation may be accomplished by varying the suppressor voltage.   The suppressor grid is biased negatively to an amount approximately half that which will cause anode current cut-off and the modulation voltage is applied in series with the bias. Fig. 14 shows a typical circuit arrangement.
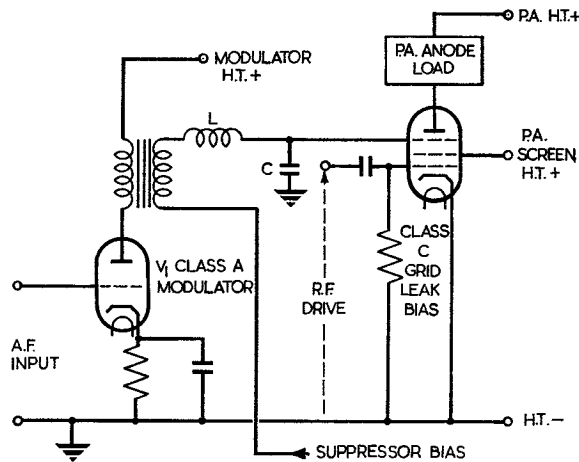

FIG. 14. SUPPRESSOR GRID MODULATION

The a.f. output from V_1 varies the p.a. suppressor grid voltage above and below the bias voltage.   Usually, as with grid modulation, the positive peak of the modulating voltage is just sufficient to reduce the bias to zero.   The r.f. filter LC isolates the two stages as in previous modulator circuits.   The waveforms are shown in Fig. 15.

Suppressor grid modulation is very similar to control grid modulation: in both cases high modulation depths cannot be achieved without appreciable distortion.   For this reason suppressor grid and control grid modulation are not commonly used in simple transmitters.   They
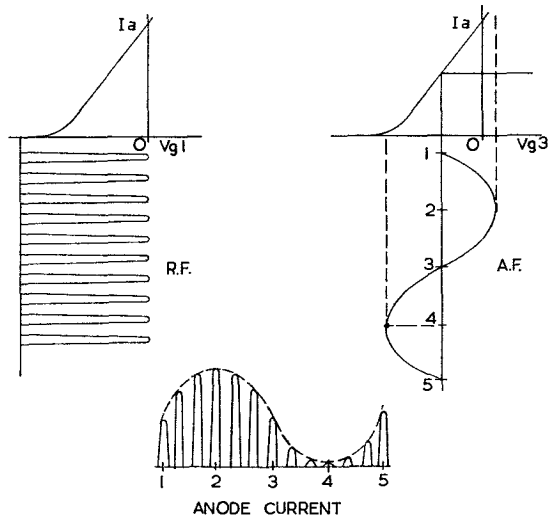
59

FIG. 15. SUPPRESSOR GRID MODULATION WAVEFORMS

are sometimes used in systems where low-level modulation is employed and the initial depth of modulation is not very important. For example, balanced modulators in suppressed carrier transmitters may employ control grid or suppressor grid modulation.

Adjustment of AM Transmitters

To obtain the desired modulation depth in an a.m. transmitter, the a.f. gain of the modulator must be adjusted. Either meters or a c.r.o. can be used as modulation indicators. Fig. 16 shows the basic meter arrangement which consists of a receiver and two moving coil meters. M_1 indicates the mean level of the r.f. carrier and M_2 shows the a.f. amplitude. Since the modulation factor is given by $\dfrac{\text{a.f. amplitude}}{\text{r.f. amplitude}}$,

the percentage modulation is given by $\dfrac{M_2 \text{ Reading}}{M_1 \text{ Reading}} \times 100$.

In the practice the two measurements are made by one meter switched between the two positions.
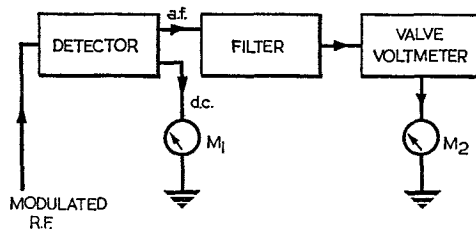

FIG. 16. MODULATION METER

CRO Modulation Indications

When a cathode ray oscilloscope is used to indicate modulation, two methods of display can be employed. The modulated r.f. signal can be fed to the Y plates of the c.r.t. and the time-

base waveform fed to the X plates. The display would then be a simple voltage variation of the transmitter output against time as shown in Fig. 17 and could be used to monitor the modulated transmitter output.
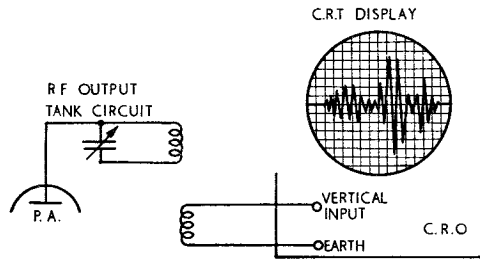

FIG. 17. CRO MONITOR METHOD

A better and more commonly used method is to dispense with the c.r.o. timebase and use in its place the output from the modulator. The c.r.t. then indicates directly the modulated wave in *terms of the modulation*. This is illustrated in Fig. 18 for 100% modulation.

With reference to the five time instants:—

1. When the modulating a.f. wave is at a negative peak the modulated r.f. is zero. Therefore the c.r.t. deflection is a maximum horizontally to the left and the vertical deflection is zero.
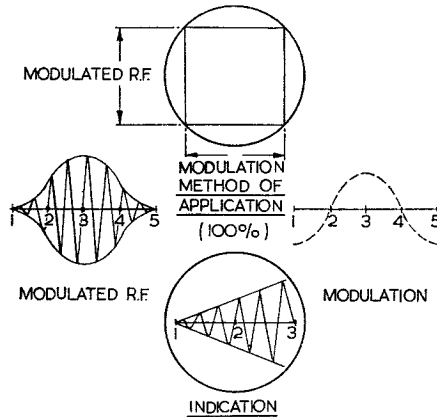

FIG. 18. MODULATION DEPTH DISPLAYED ON CRT

2. When the modulation is at zero the modulated r.f. is at its unmodulated value. Therefore the c.r.t. deflection is zero horizontally (i.e. in the centre of the c.r.t.) and at unity vertically.

3. When the modulation is at a positive peak the modulated r.f. is a maximum. Therefore the c.r.t. deflection is a maximum horizontally (to the *right*), and maximum vertically.

4. Same as 2.

5. Same as 1.

Notice that the c.r.t. really produces only a half pattern (i.e. half a diamond), since the second half merely reverses the process causing the first. Fig. 19 shows some typical modulation patterns.
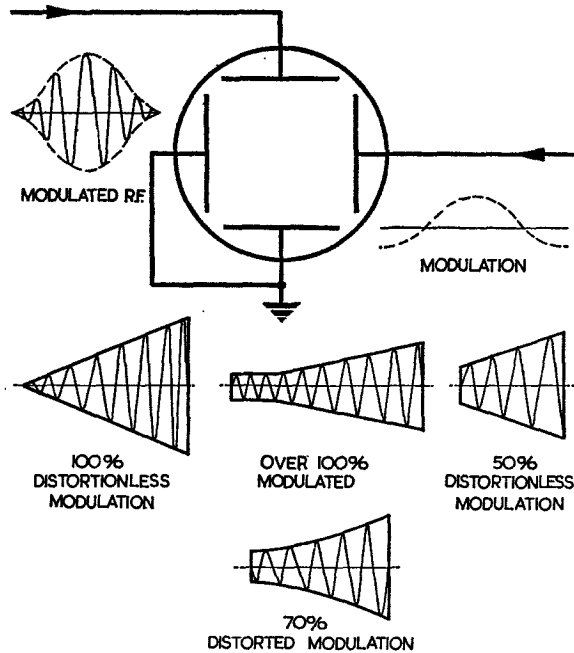
FIG. 19. MODULATION PATTERNS

In practice the carrier frequency will be so high that the separate cycles will not be visible and the picture seen on the c.r.t. will appear to be "filled in" as shown in Fig. 20. There will probably be some phase shift in the circuit in practice and this causes the sloping sides of the pattern to form elipses as shown.

$$M = \frac{B-A}{B+A} \times 100\%$$

MODULATION DEPTH (NO PHASE SHIFT).   OVER 100% MODULATION
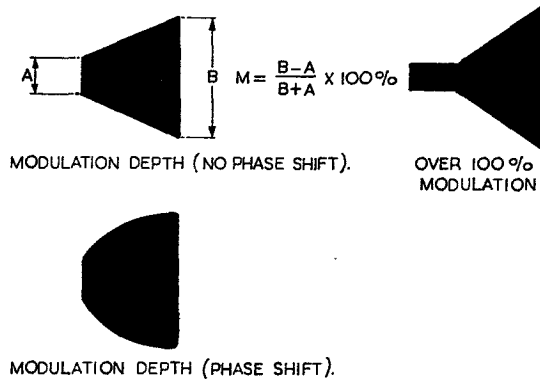
MODULATION DEPTH (PHASE SHIFT).

FIG. 20. PRACTICAL CRT DISPLAYS

Other Modulation Circuits

For some types of transmission it is necessary to radiate a modified form of amplitude modulated wave. For example:—

a. *Suppressed carrier transmission.* In this case only the sidebands are transmitted.

b. *Single sideband transmission.* Only one sideband is transmitted and the carrier is completely suppressed or transmitted at a reduced power level.

When it is necessary either to suppress the carrier or to reduce its amplitude a *balanced modulator* is used. This is a two valve circuit arranged so that the carrier output from one valve cancels the carrier output from the other.

The Balanced Modulator

The circuit of the balanced modulator is shown in Fig. 21. Valves V_1 and V_2 are connected in push-pull and the a.f. input is connected to the grids in the conventional push-pull manner. However, the r.f. input is fed in parallel to the grids.
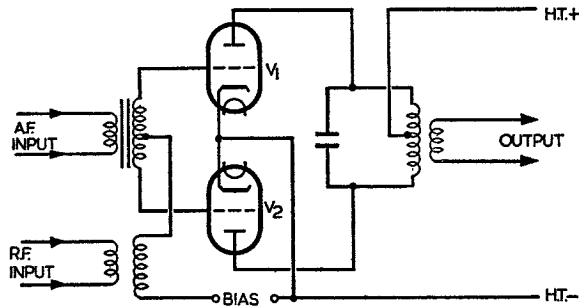
FIG. 21. THE BALANCED MODULATOR

Notice that the output circuit is a r.f. circuit and that therefore none of the a.f. input can appear in the output. Furthermore, since the r.f. input is applied in parallel, the r.f. grid potentials are in phase and therefore, if the valves are identical (balanced), there can be no output
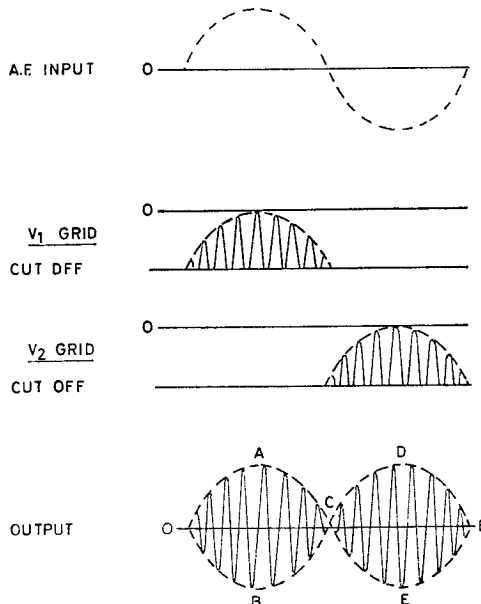
FIG. 22. BALANCED MODULATOR WAVEFORMS

63

at this frequency in the push-pull (anti-phase) anode circuit. Thus both inputs are eliminated and only the frequencies produced by mixing the two, i.e. the upper and lower sidebands, appear in the output.

The input and output waveforms are shown in Fig. 22.

In effect the a.f. input switches the r.f. off and on and modulates the amplitude of the r.f. input. The tuned circuit combines the two anode currents and by its "flywheel" action adds the missing half cycles of modulation. In many balanced modulator circuits semiconductor diodes are used instead of valves.

Notice that the envelope is not the normal amplitude modulated waveform; the curves OACEF and OBCDF are sine curves but the amplitude variation OACDF, is not. This, in fact, is the result of removing the carrier in a normal a.m. waveform.

Speech Clipping

It was mentioned earlier in this chapter that with speech modulating voltages it is difficult to obtain a high average percentage modulation. This is because the speech waveform is very "peaky", the peaks usually occurring on the vowels or explosive consonants. Therefore, to avoid overmodulating on these peaks and causing distortion, the average modulation depth has to be kept low (about 30%), with a corresponding lowering of readability.

"Speech clipping" circuits are used in some r.t. transmitters to overcome this difficulty. These circuits attenuate the louder sounds without making them unintelligible and the new, lower peak is then used to modulate to 100%. In this way, the average depth of modulation can be increased to about 50% without distorting the speech too much and so the readability of the signal is increased.

CHAPTER 8

SINGLE SIDEBAND TRANSMISSION

Introduction

When transmitting r.t. signals it is usual to employ amplitude modulation. This requires a bandwidth equal to twice the highest audio frequency component. The principle is illustrated in Fig. 1. It is convenient to draw a triangle to represent the range of frequencies corresponding to the original speech. For intelligible speech it is necessary to use the 300 to 3,000 c/s a.f. and the original triangle represents this band.
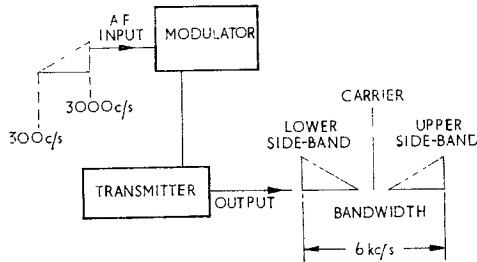

FIG. 1. SIMPLE AM TRANSMITTER

The output from a simple a.m. transmitter consists of a carrier and two sidebands. The sideband frequencies are obtained by adding the audio frequencies to the carrier frequency and by subtracting the audio frequencies from the carrier frequency. If the carrier frequency is 100 kc/s the frequency distribution of the transmitted signal will be as shown in Fig. 2.
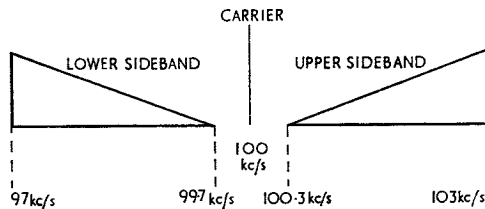
FIG. 2. FREQUENCY OUTPUT FROM A SIMPLE AM TRANSMITTER

Notice that the "sharp end" of the triangle corresponds to the lowest audio frequency; this is the reason why the triangular shape is used: *it does not indicate the amplitude.* In fact the amplitude of the sideband components varies all the time as the speech varies and so it cannot be shown.

Certain difficulties arise when simple amplitude modulation is employed, particularly when long-range communication is required. They are caused by:—

a. Selective fading.

b. Noise.

c. Interference.

Selective Fading

To receive a.m. signals without distortion it is necessary that the relationship between carrier and sidebands be undisturbed. In long-range circuits parts of one or both sidebands may fade out from time to time causing the signal reaching the receiver to be unintelligible. This type of fading is called *selective fading* and is due to the fact that the carrier and sidebands are attenuated by different amounts in the atmosphere. It is partly because of this type of fading that aircraft have not in the past used long-range h.f. radio telephony. This effect is illustrated in Fig. 3. Note that certain parts of the sidebands fade, while other parts remain unaffected; hence the term selective fading.
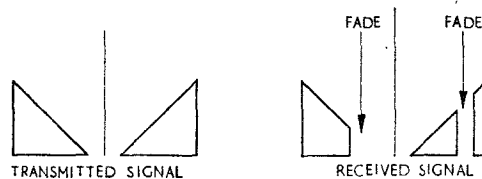

FIG. 3. SELECTIVE FADING

Noise

A receiver, in addition to receiving and amplifying the required signal, will also receive and amplify unwanted noise. The noise may originate in a tropical storm over the Sahara or in an electric drill next door: it may even come from outer space. Because of the diverse character of the sources of noise it is fairly evenly distributed over all frequencies and the amount of noise picked up by a receiver depends on the bandwidth of the receiver. With simple a.m. systems, although the audio bandwidth is only 3 kc/s it is necessary to receive a bandwidth of 6 kc/s because of the two sidebands. This means that with simple amplitude modulation the receiver must accept twice as much noise as is strictly necessary, i.e. the noise included in a bandwidth of 6 kc/s instead of 3 kc/s.

Interference

Interference can arise if the sidebands of one transmission overlap into the sidebands of another transmission. To minimise this effect transmitter frequencies are spaced far enough apart to allow for both sidebands. In the h.f. broadcast band this spacing is 9 kc/s. If only one sideband were transmitted instead of two the risk of interference would be reduced and more channels could be accommodated.

Single-sideband Systems

In s.s.b. systems one sideband only is transmitted; the carrier wave is sent at a low power level, transmitted intermittently or is not sent at all. At the receiver the carrier is separated, amplified and recombined with the sideband before the detector stage. The low-level carrier is transmitted because when the carrier is combined with the sideband in the receiver it is essential that the carrier frequency is accurate within a few cycles per second. Because the transmitter frequency drifts slightly, the receiver can keep in step only by using the actual transmitter carrier frequency.

Recent developments in component stability have made it possible to reproduce the carrier in the receiver with the necessary degree of accuracy (to within 10 c/s). This means that the sideband can be transmitted without a carrier.

The Nature of a SSB Signal

The signal from a single sideband (s.s.b.) transmitter differs from a simple a.m. signal in two important respects:—

a. Only one sideband is transmitted.

b. The mean carrier power is very much reduced or the carrier is not sent at all.

CARRIER

LS.B.   U.S.B.   CARRIER   U.S.B.

SIMPLE A.M. SIGNAL
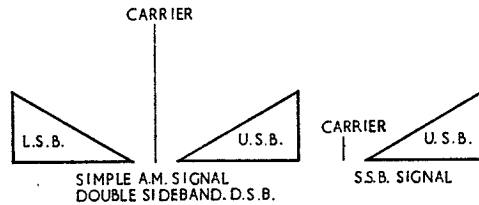DOUBLE SIDEBAND. D.S.B.   S.S.B. SIGNAL

FIG. 4. DSB AND SSB SIGNALS

In Fig. 4 the s.s.b. signal is shown as including the upper sideband; it may in fact consist of either the upper or lower sideband.   For convenience throughout this chapter it will be assumed that the upper sideband is used, but it must be remembered that the lower sideband can be used instead.

There are three forms of s.s.b. transmission:—

a. *Pilot carrier.*   The carrier is sent continuously at a power level much lower than would be necessary for double sideband (d.s.b.) transmission, when full carrier power and both sidebands are transmitted.

b. *Controlled carrier.*   The carrier is not sent continuously but only during breaks in speech, i.e. between words and sentences, whenever the speaker pauses.   A special circuit (on the lines of an automatic gain control circuit) is used to ensure that when speech signals are present the carrier is cut off.   When there is a break in speech and the carrier is transmitted it is sent at full strength.

c. *Suppressed carrier.*   No carrier is transmitted.   Either the upper or lower sideband is selected, its frequency level raised, and it is transmitted by itself.   This is the most efficient method of s.s.b. transmission since all the transmitted power is contained in the one sideband.   It is necessary, however, to re-insert the carrier at the receiver and a very high degree of frequency stability is therefore necessary.

It is of course possible to combine any of the above systems.   Thus a pilot carrier can be transmitted with two sidebands or two sidebands can be transmitted without a carrier.   In both these systems the two sidebands would carry different intelligence and so two messages could be sent simultaneously by the one transmitter, thus doubling the number of communication channels.   Since the intelligence contained in the two sidebands is independent of each other, this system is called independent sideband (i.s.b.) transmission.

Advantages of SSB Over DSB

The advantages of s.s.b. transmission over d.s.b. transmission can be summarised as follows:

a. *Saving in bandwidth.*   The bandwidth of a s.s.b. signal is half the bandwidth of a d.s.b. signal.   This can be used either to reduce interference between channels or to provide more channels within a given frequency range: for example, a 90 kc/s frequency range will contain 10 channels with 9 kc/s spacing or 18 channels with 5 kc/s spacing.

b. Saving in carrier power. The s.s.b. carrier power is relatively low: this fact can be used in one of two ways:—

 i. To develop a given sideband power the s.s.b. transmitter can be made much smaller and lighter than the d.s.b. transmitter.

 ii. Using the same transmitter the sideband power output will be much greater with s.s.b. than d.s.b. transmission. It is, of course, the sidebands which contain the information being transmitted. Note that although only a weak pilot carrier is transmitted, it must be amplified in the receiver before any intelligence can be obtained from the signal.

c. Reduction in noise. Because the s.s.b. signal occupies half the bandwidth of the d.s.b. signal, the s.s.b. receiver need accept only half the normal bandwidth. It will therefore accept only half the normal noise.

d. Selective fading. S.s.b. signals are subject to selective fading in the same way as d.s.b. signals but the final effect is much less severe. This is because when part of a d.s.b. signal is lost the sidebands become unbalanced and severe distortion occurs in the detector stage. Thus distortion usually makes the received signal unintelligible. On the other hand, selective fading simply "cuts" part of the s.s.b. signal: there is no "balance" to destroy and therefore less distortion.

Methods of Producing a SSB Signal

In order to produce a s.s.b. signal it is necessary to employ a balanced modulator, the circuit of which was described in Chapter 7. The output from the balanced modulator consists of sidebands only: there is no carrier present in the output. The difference between the output from a simple amplitude modulation circuit and that from a balanced modulator is illustrated in Fig. 5.
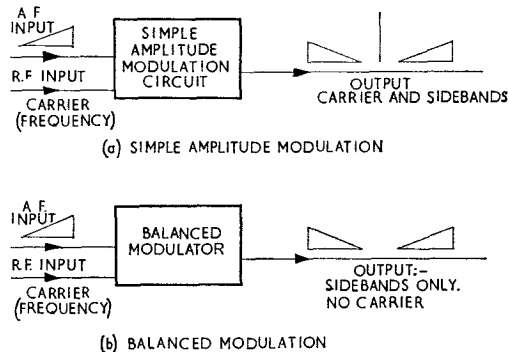

FIG. 5. COMPARISON OF SIMPLE AND BALANCED MODULATION

There are three ways of producing a s.s.b. transmission; only the two methods of importance to the RAF will be described. The third method (it is actually called "the third method") is a combination of the other two techniques.

The "Filter" Method

This method has been used for some time in long-range ground-to-ground multi-channel teleprinter circuits. Basically the systems consists of a balanced modulator followed by a filter which allows one sideband to pass through but stops the other one (see Fig. 6).

In practice the process must be repeated in a further stage because of the difficulty of making filters with sufficient selectivity. The frequency separation between the upper and lower sidebands is very small and they can be separated by filters only if the mean frequency is low. In the teleprinter system the lowest audio frequency is 420 c/s so that the upper and lower sidebands
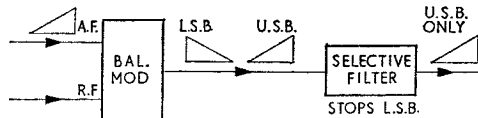
FIG. 6. SIMPLE SSB SYSTEM

are separated by only 840 c/s. The first filtering is carried out at a mean frequency of 100 kc/s. This gives an upper sideband starting at 100·42 kc/s and to this is added a pilot carrier at 100 kc/s. These first stages are shown in Fig. 7.
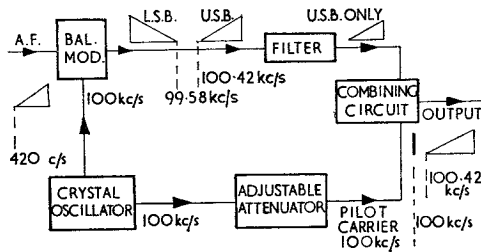
FIG. 7. FIRST STAGES OF A SSB TRANSMITTER (FILTER TYPE)

The above process is repeated using another balanced modulator, a 3 Mc/s oscillator and a second filter. The second balanced modulator produces upper and lower sidebands spaced by 200 kc/s; they are therefore easy to separate. The output from this filter consists of a pilot carrier on 3·1 Mc/s and an upper sideband beginning at 3·10042 Mc/s. This process is illustrated in Fig. 8
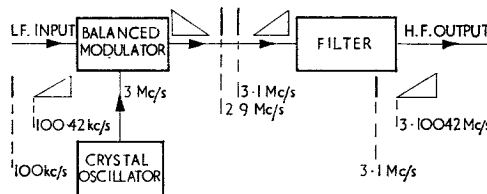
FIG. 8. INTERMEDIATE STAGES OF A SSB TRANSMITTER (FILTER TYPE)

The 3·1 Mc/s signal is then fed into a frequency converter which changes the frequency to the required final (transmitted) frequency. Power amplifiers then develop the necessary power.

Suppressed Carrier Filter Transmitter

A development of the filter method of s.s.b. transmission is used in the suppressed carrier system which is now being used for air-to-ground and air-to-air r.t. and c.w. communication. Improvements in the design of filters have enabled the sideband separation to be carried out in one stage. The filter used is a very narrow band filter known as a *mechanical filter*. It is a mechanically resonant device which converts an electrical input into very stable mechanical vibrations which are then converted back into electrical oscillations.

A block diagram of a suppressed carrier transmitter employing mechanical filters is shown in Fig. 9.
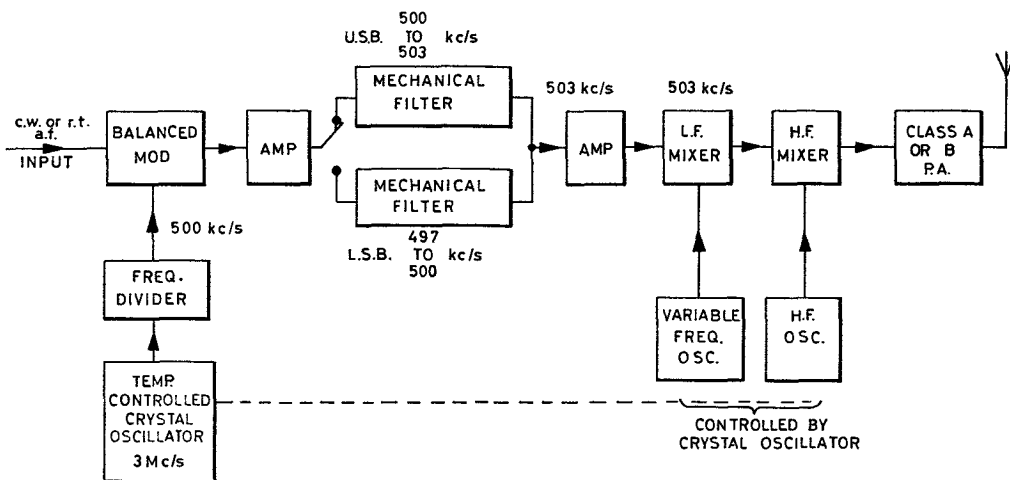

FIG. 9. SUPPRESSED CARRIER TRANSMITTER

The audio input and an input at 500 kc/s, obtained from a very stable temperature controlled crystal oscillator, are applied to a balanced modulator. The output consists of the upper sideband (500 to 503 kc/s) and the lower sideband (497 to 500 kc/s). Either of these can be selected by one of the two mechanical filters and after amplification the selected sideband is fed through two frequency converter stages. The output from the final converter stage is at the required frequency for transmission. This voltage is fed to the p.a. stage where it is amplified to give the required power output. The p.a. stage must be operated in class A or B since any distortion introduced by this stage would appear in the output of the associated receiver.

The transmitted signal consists only of the selected sideband. No carrier is transmitted but the original 500 kc/s carrier, which was eliminated by the balanced modulator, must be re-inserted at the receiver.

The "Out-phasing" Method

This method also is suitable for use as a long-range h.f. radio telephone in ground-to-air links and will shortly be introduced into Service aircraft. It has certain advantages over the previous method in that filters are not required and it is easier to change from upper to lower sideband working. The disadvantage is that the unwanted sideband is never completely eliminated and so it can cause interference.

The method relies on the production of two pairs of sidebands such that the lower sideband of one pair cancels the lower sideband of the other pair. The two upper sidebands then re-inforce one another. The principle is illustrated in Fig. 10.
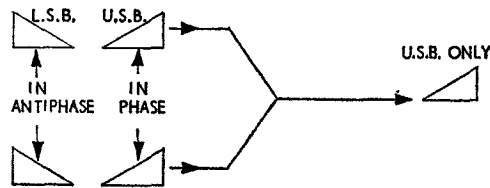
FIG. 10. THE "OUT-PHASING" METHOD

The difficulty, of course, is to produce the sidebands with the required phase relationships. Since the sidebands consist of a number of frequencies it is not possible to obtain exactly the right relationship but an error of one or two degrees is acceptable. As before, the wanted sideband is produced at a set frequency and then a frequency converter changes the set frequency to the required transmission frequency. A block diagram showing the system is given in Fig. 11.
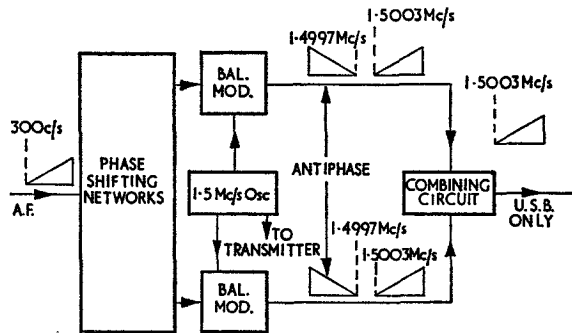

FIG. 11. BLOCK DIAGRAM OF "OUT-PHASING" METHOD

The phase shifting networks consist of complicated bridge arrangements of capacitors and resistors, known as *dome* circuits. The 1·5 Mc/s oscillator is a stable, fixed frequency oscillator with three outputs.

The remainder of the transmitter consists of frequency converter and power amplifier stages together with a carrier control circuit operated by the audio signal. When there is no a.f. signal the carrier is transmitted at full strength (Fig. 12).

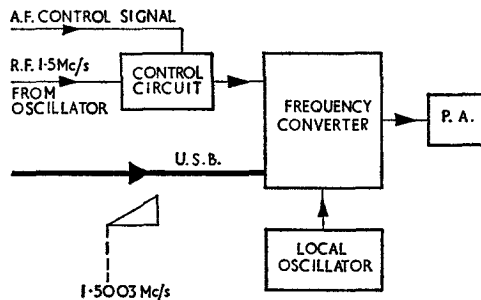FIG. 12. FREQUENCY CONVERTER

With this method of producing a s.s.b. signal it is possible to change from upper to lower sideband working very simply by switching the connections to the phase shifting network. The filter method requires a further set of filters to provide the alternative sideband. In practice such a change from one sideband to the other provides a convenient way of avoiding interference.

71

CHAPTER 9

FM AND FSK TRANSMISSION

Introduction

Previously it was stated that to impart intelligence to a carrier wave, any one of three of its characteristics could be varied: its amplitude, its frequency or its phase. Amplitude modulation has been dealt with and this chapter is devoted to various aspects of frequency and phase modulation.

Features of Frequency Modulation

The main features of frequency modulated transmission are summarized as follows:—

a. *Bandwidth.* In order to accommodate as many stations as possible the l.f., m.f. and h.f. bands are divided into "channels". These channels are frequency bands 9 kc/s wide and each transmitter is allocated to one of these channels. FM transmitters require a bandwidth of about 200 kc/s and so one f.m. transmitter would occupy more than 20 "channels" in the l.f., m.f. or h.f. bands. For this reason f.m. transmissions are confined to the v.h.f. band and above.

b. *Fidelity.* The use of a wide bandwidth means that f.m. is suitable for "hi-fi", i.e. complex sounds such as music are reproduced better by using a f.m. system than by using an a.m. system.

c. *Range.* Since f.m. systems are confined to frequencies in the v.h.f. band and above, the range is limited to "line of sight". The range, then, will depend upon the heights of the transmitter and receiver aerials, and a fairly low power output (about 50W) is all that is required.

d. *Noise.* Noise voltages usually cause the amplitude of the carrier wave to vary. A f.m. receiver can be designed so that it does not respond to amplitude variations, thus greatly increasing the signal-to-noise ratio.

Fundamental Principles

In frequency modulation the modulating signal varies the frequency of the carrier. The amount by which the carrier frequency varies above or below its mean value is termed the *frequency deviation.* Actually, there is no limit to the amount of frequency deviation which can be given to a carrier but, since a large deviation requires a large bandwidth, it is necessary to limit the deviation to some reasonable maximum value. In practice, broadcast transmitters employ a maximum deviation of 75 kc/s.

It is important to remember that with a f.m. transmission the frequency deviation is proportional to the *amplitude* of the modulating signal: the *rate* at which the carrier frequency deviates depends upon the *frequency* of the modulating signal. These two points are illustrated in Fig. 1 where a square wave modulating signal is considered.

Notice that since f_{a_2} and f_{a_3} are of the same amplitude the resultant deviation of the carrier is the same in both cases, i.e. f_c varies between $f_c \pm f_{d_2}$.

Methods of Frequency Modulation

The basic ways in which the output of a transmitter can be frequency modulated were considered in Part 1B. The most commonly used of these methods is that employing the reactance valve.
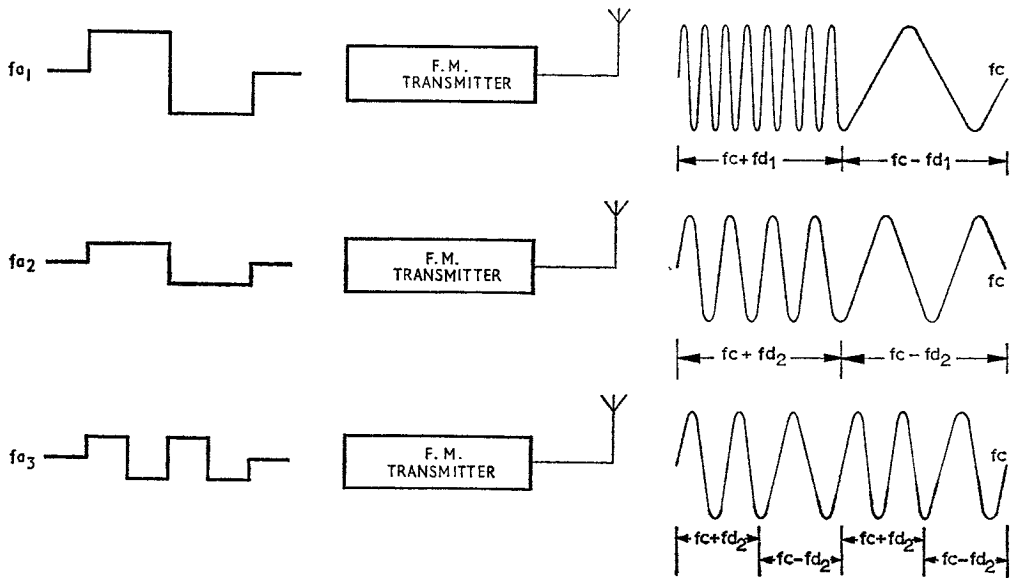
FIG. I. EFFECT OF A CHANGE IN AMPLITUDE AND FREQUENCY OF MODULATING SIGNAL

In this circuit a valve passes a current which leads or lags 90° on its anode voltage. Thus the valve is effectively an inductive or capacitive reactance, the value of which is easily varied by varying the mutual conductance of the valve. Fig. 2 shows the basic arrangements and the equivalent reactance circuits; for simplicity the screen and suppressor circuits are omitted.
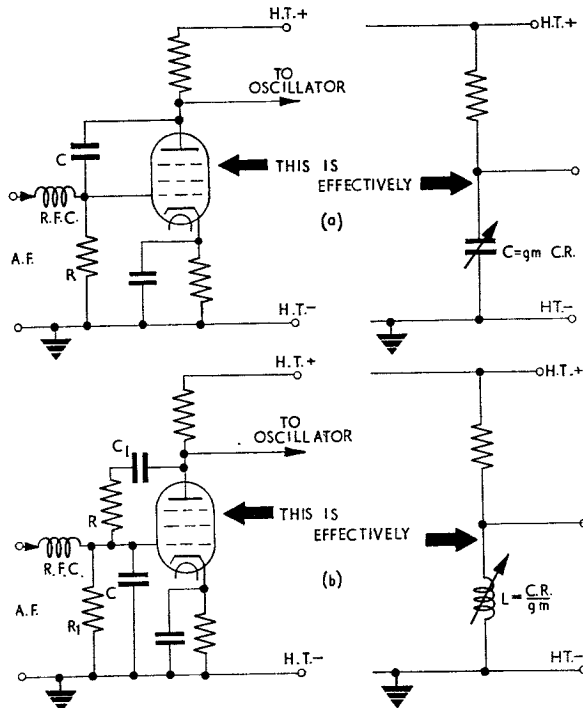

FIG. 2. REACTANCE VALVES

73

Notice that the additional components C_1 and R_1 in the circuit of Fig. 2b play no part in determining the effective reactance. C_1 merely blocks the d.c. h.t. from the grid circuit, while R_1 provides the usual d.c. grid return. C_1 is very much larger than C and R_1 is very much larger than R.

There are other circuit arrangements which will provide the necessary 90° phase shift, such as a combination of inductance and resistance, but the one shown in Fig. 2a is most often used.

Phase Modulation

Not all f.m. transmitters employ frequency modulation as derived directly by a reactance valve modulator. Phase modulation can also be effected by varying the phase of the carrier so that the phase change is proportional to the amplitude of the modulating signal. The phase modulation can then be converted quite easily into frequency modulation.

Suppose that the amplitude of the modulating signal is increasing such that it produces a continuous change of carrier phase. Fig. 3 shows such a signal causing a phase change of 90° for every cycle of the carrier.
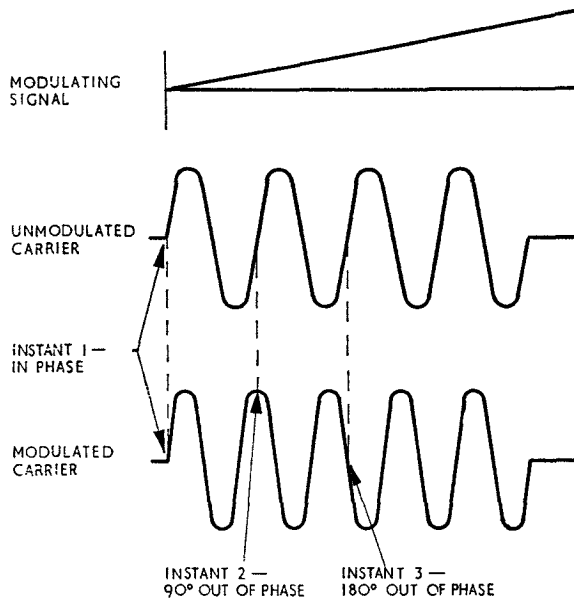
FIG. 3. PHASE MODULATION

The effect of four changes of 90° is to advance the carrier phase by 360°, i.e. f_c increases by one cycle. This shows that modulating the phase of a carrier results in frequency modulation and in fact the two systems are interchangeable. Some f.m. transmitters employ a process which is really phase modulation. The advantage of this method of modulation is that the transmitter oscillator can be crystal controlled and can thus provide a stable centre frequency.

Outline of the Phase Modulator

A simple circuit which will change the phase of a constant frequency r.f. input voltage is shown in Fig. 4. If an alternating input voltage is applied across C and R then, by varying the value of R, the output voltage taken across the resistor will vary in phase between about 0° when R is high and nearly 90° when R is low.
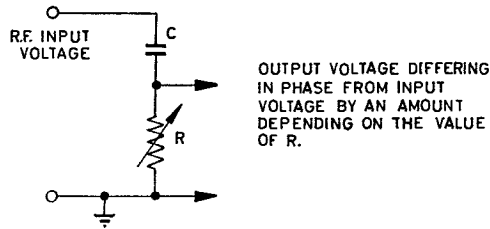
FIG. 4.   SIMPLE PHASE SHIFT NETWORK

If the resistor is replaced by a triode valve as shown in Fig. 5 and a modulating a.f. is applied to the grid then as the grid voltage increases so the valve current increases and the effective resistance of the valve is reduced.   Thus the phase of the r.f. output will change.   When the grid voltage decreases the anode current decreases and the effective resistance of the valve increases.   In this way, the output voltage taken from the anode will be phase modulated in sympathy' with the modulating a.f. voltage.
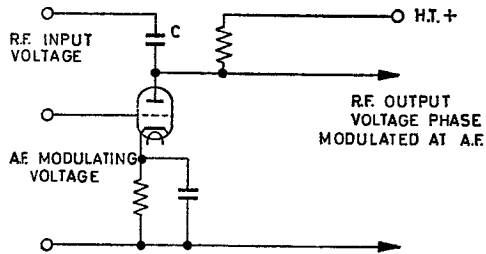

FIG. 5.   PHASE SHIFT CIRCUIT USING A TRIODE VALVE

With the circuit as shown in Fig. 5, the variations in a.f. amplitude will vary the phase of the r.f. input but also a variation in audio frequency will vary the rate of phase change and therefore the equivalent frequency deviation.   The carrier frequency deviation of a f.m. wave must be proportional only to the *amplitude* of the a.f. modulating wave and must not be affected by the *frequency* of the modulating voltage.   The *rate* at which the carrier frequency deviates is governed by the modulating frequency.
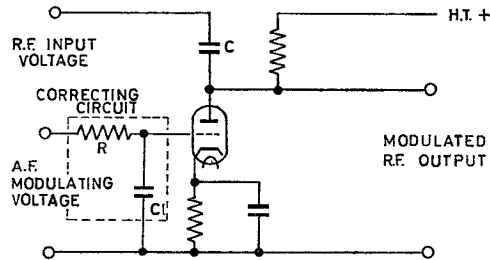

FIG. 6.   PHASE MODULATOR WITH CORRECTING CIRCUIT

To eliminate the effect of increasing the frequency deviation as the frequency of the modulating voltage increases a correcting circuit is necessary.   This circuit must ensure that as the modulating frequency goes up, so the amplitude of the modulating voltage comes down.   A CR circuit

similar to a treble-cut circuit as used in a tone control network will achieve this result if it is placed in the grid circuit of the modulator valve as shown in Fig. 6. This basic circuit could be used to modulate a f.m. transmitter where the modulating waveform is not very sharp.

Frequency Multipliers in FM Transmitters

The reactance valve modulator will produce only small frequency deviations and a practical f.m. transmitter must raise the frequency deviation by using frequency multipliers.

Fig. 7 shows a typical arrangement of a f.m. transmitter employing a reactance valve modu-
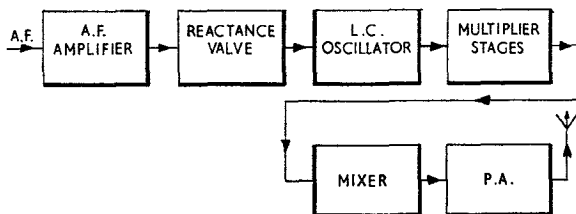
FIG. 7. BASIC FM TRANSMITTER

lator. The output from the oscillator stage may be 2 Mc/s $\pm$ 650 c/s and it is necessary to convert this output by means of multiplier and mixer stages to a v.h.f. output of say 90 Mc/s with a frequency deviation of about 75 kc/s.
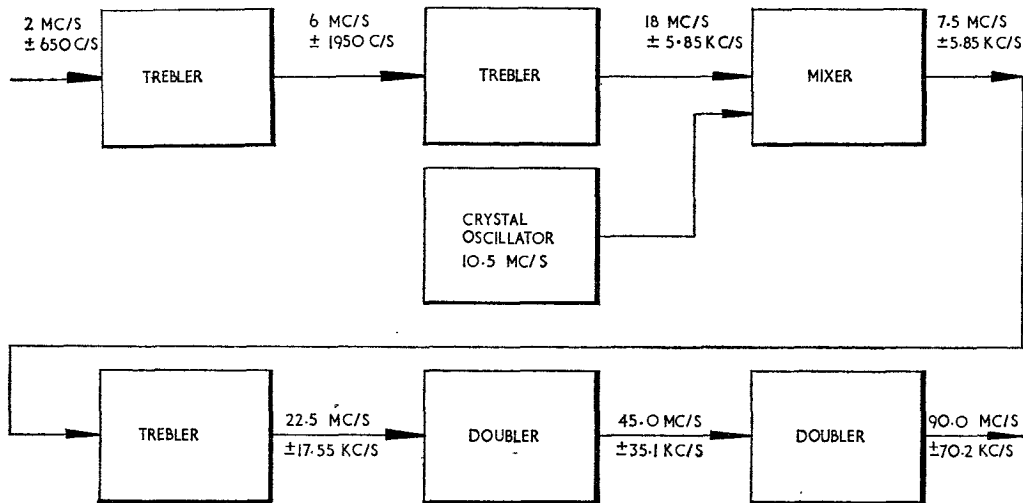
FIG. 8. DETAILS OF MIXER AND MULTIPLIER STAGES

Fig. 8 shows a breakdown of the multiplier and mixer stages of Fig. 7. To change the frequency deviation from 650 c/s to about 75 kc/s requires frequency multiplication by about 115. Frequency multipliers are normally doublers and treblers and using these it is possible to attain multiplication by 108 ($3 \times 3 \times 3 \times 2 \times 2 = 108$). This would give a frequency deviation of 70·2 kc/s, which is a practical deviation. If the carrier frequency were multiplied by 108 it would become 216 Mc/s which is too high. Thus, as shown in Fig. 8, a mixer stage

is used to reduce the carrier frequency to 90 Mc/s. The mixer stage does not affect the frequency deviation.

Crystal-stabilized FM Transmitter

LC oscillators are normally subject to frequency drift and since f.m. transmitters operate in the v.h.f. band, a stable frequency output is essential. Therefore, a system using automatic frequency control of a conventional LC oscillator, as shown in Fig. 9, is often used.
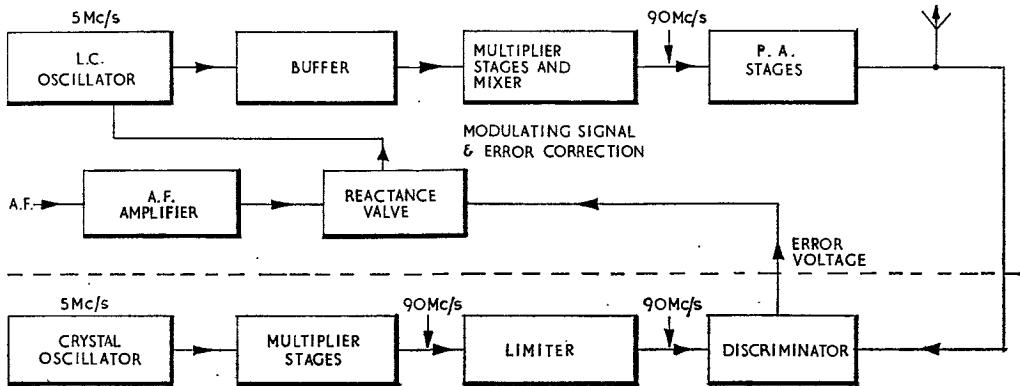

FIG. 9. CRYSTAL STABILIZED DIRECT FM TRANSMITTER

This form of crystal-stabilized transmitter uses a LC oscillator with a reactance valve modulator plus a crystal controlled reference oscillator. It thus combines the advantages of direct f.m. and the stability of a crystal oscillator.

If the p.a. frequency output is exactly 90 Mc/s there is no d.c. error voltage output from the discriminator and the stages below the dotted line are non-effective. If there is an error the fixed bias on the reactance valve is varied to correct the output carrier frequency.

Frequency Shift Keying

Keying a transmitter by switching the carrier wave on and off was discussed in Chapter 6. A different method of keying employed on some teleprinter networks uses the principle of frequency modulation and is known as *frequency shift keying* (f.s.k.). The transmitter is effectively a f.m. transmitter with a very small frequency deviation.

Disadvantage of Simple ON/OFF Keying

With simple on/off keying the transmitter sends a signal (i.e. makes a mark) when the key is closed and sends nothing (i.e. makes a space) when the key is open. If fading occurs on the transmitter frequency the receiver will not receive a mark signal and so will record nothing. This makes the received message unintelligible; it can be overcome to a certain extent by sending an alternative signal on another frequency to denote a space.

Consider a simple stop/go traffic light system. The "stop" could be indicated by a red lamp and the absence of the red light could denote "go". Such a system is a simple form of on/off signalling. It has the disadvantage that if the lamp is obscured by an obstruction, or if the bulb fails, an observer is likely to think it means "go", i.e. he receives the wrong message.

If, however, two colours are used, one to denote "go" and the other to denote "stop", the observer will know that if he sees anything at all it is the correct message. The absence of both colours simply means "no message".

When a transmission using on/off keying is not received the receiver does not know whether to record "space" or "nothing received". If, on the other hand, a transmitter uses two frequencies, one for space and one for mark, it will not matter if reception on one of the frequencies fades from time to time.

In the illustration of Fig. 10 it is assumed that a simple on/off transmitter is sending the word "radio" on a frequency f_1, and that the signal fades twice.
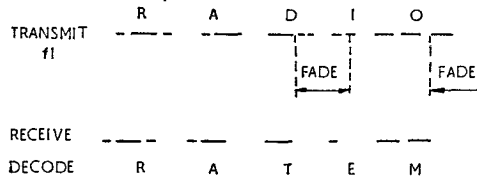
FIG. 10. SIMPLE ON/OFF KEYING

If the transmitter had been using a second frequency f_2 to denote the gaps between the characters (the spaces) then even if f_1 faded completely it is unlikely that f_2 would fade as well at the same time. The message received on f_2 would be as shown in Fig. 11.
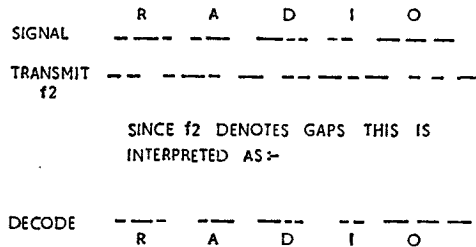
FIG. 11. SIGNALLING USING A "SPACE" FREQUENCY

With this system it is reasonable to assume some fading on both f_1 and f_2 throughout the transmission of a message but so long as f_1 and f_2 do not fade *together* the message would be received clearly. This is the principle of frequency shift keying.

Fundamentals of FSK

The difference between on/off keying and f.s.k. is illustrated in Fig. 12. Instead of interrupting a continuous train of r.f. energy (Fig. 12a), f.s.k. allows the aerial to radiate continuously

FIG. 12. COMPARISON OF ON/OFF KEYING AND FSK

78

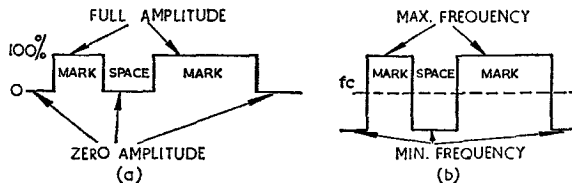but the *frequency* radiated is varied depending on whether a mark or a space is being transmitted (Fig. 12b). In other words, the transmission is frequency modulated. The frequency separation between the mark and the space frequencies is standardised at 850 c/s for the h.f. band. Thus a space impulse is radiated on the nominal frequency (f_c) plus 425 c/s and a mark impulse on a frequency of f_c minus 425 c/s. Since this is f.m. using a small frequency deviation, f.s.k. has all the advantages of f.m. without using too wide a bandwidth. The system can therefore be used in the h.f. band.

Basic FSK Transmitter

A block diagram of a basic f.s.k. transmitter is shown in Fig. 13. The impulses corresponding to mark and space pulses are fed into a keying valve circuit which converts them into positive and negative bias voltages suitable for applying to the reactance valve.

The reactance valve varies the frequency of the 200 kc/s LC oscillator which therefore generates a constant amplitude r.f. signal whose frequency is shifted by the reactance valve in
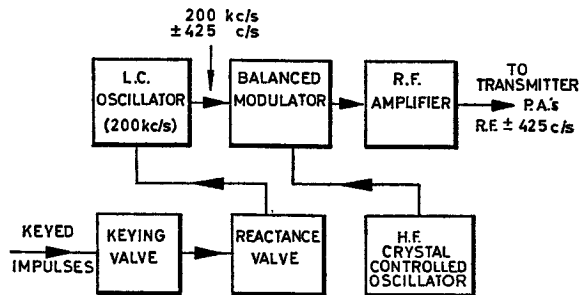

FIG. 13. BASIC FSK TRANSMITTER

accordance with the mark and space impulses. The amount of frequency shift, or deviation, depends on the amplitude of the pulses fed to the reactance valve.

In any f.s.k. transmitter frequency stability is essential. In an ordinary c.w. or r.t. transmitter a frequency drift of a few hundred cycles per second would not be serious, but such a frequency drift in a f.s.k. transmitter cannot be tolerated. Therefore the main h.f. oscillator must be crystal controlled.

The output from the 200 kc/s oscillator is applied to the push-pull input of the balanced modulator and the parallel input is fed with the h.f. crystal-controlled oscillations. The sum frequency of the two inputs is selected by the output tuned-circuit and fed via r.f. amplifiers to the p.a. stages of the transmitter.

POWER SUPPLIES FOR TRANSMITTERS

Introduction

To enable a transmitter to operate it must, of course, be provided with electric power. The anodes and screen grids of the valves require d.c. at several hundred (or thousand volts) and the valve heaters or filaments require a l.t. supply which may be a.c. or, in the case of directly heated valves, d.c. The source of power used depends on the nature of the transmitter: for fixed installations the mains can be used (240V 50 c/s a.c. or three phase 400V 50 c/s a.c.); for small portable transmitters batteries are suitable: and for airborne transmitters electrical power derived from the aircraft engines can be employed.

This chapter deals with the transmitter power units which convert the original power into h.t. and l.t. suitable for operating the transmitter.

Power Unit Requirements

If a transmitter is required to have a r.f. power output of 500 watts, the power unit must be capable of supplying much more than 500 watts. The most efficient class C power amplifier has an efficiency less than 80% and, of course, power is also required for heating the valve filaments, for screen grids and for bias supplies. In addition power is also needed to drive such auxiliary machinery as cooling fans and blower motors.

Thus the input d.c. power is much larger than the output r.f. power and in large installations the power supply unit is very bulky.
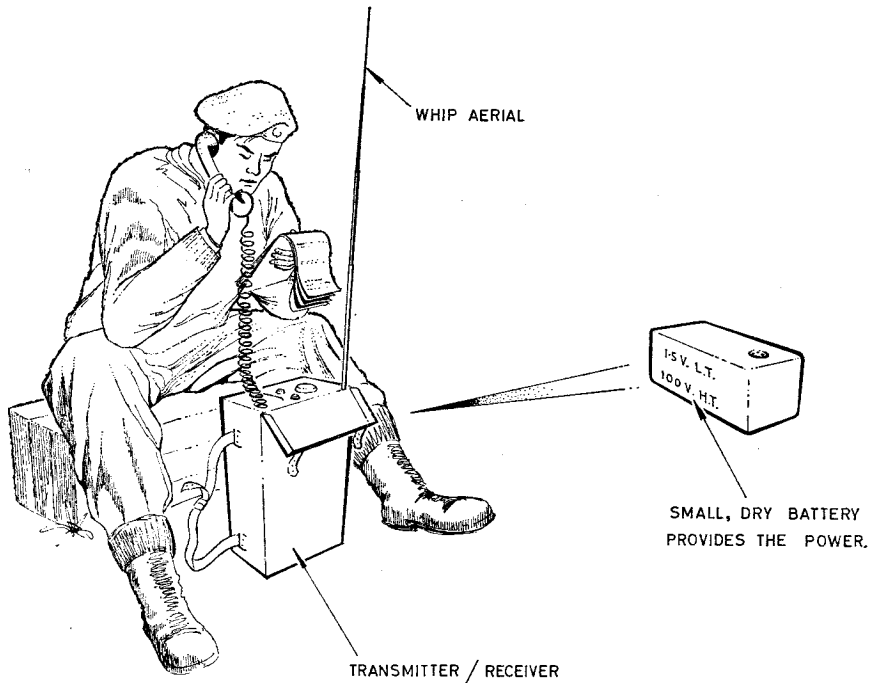

WHIP AERIAL

1·5 V. L.T.
100 V. H.T.

SMALL, DRY BATTERY
PROVIDES THE POWER.

TRANSMITTER / RECEIVER

FIG. 1. PORTABLE TRANSMITTER-RECEIVER

Mobile Transmitters

The type of power unit used in mobile transmitters will depend to a great extent on what is required of the transmitter. A large proportion of the total weight of a transmitter is taken up by the power unit and the greater the r.f. power output, the heavier must be the power unit. For example, a "walkie-talkie" set used by paratroopers has a power output of less than one watt and a range of about 5 to 10 miles (Fig. 1). It requires very little input power which can be obtained from dry batteries; these provide 1·5 or 2V l.t. and about 100V h.t. Transistorized sets with their low power consumption and low working voltages require a single 6 or 12V battery of about one-tenth the size required by a valve walkie-talkie.

Mobile sets designed to work over a range of 50 to 100 miles would have a r.f. power of about 50 watts. Therefore these sets are much larger and must be mounted in lorries and provided with a gasoline-electric generating set (Fig. 2).
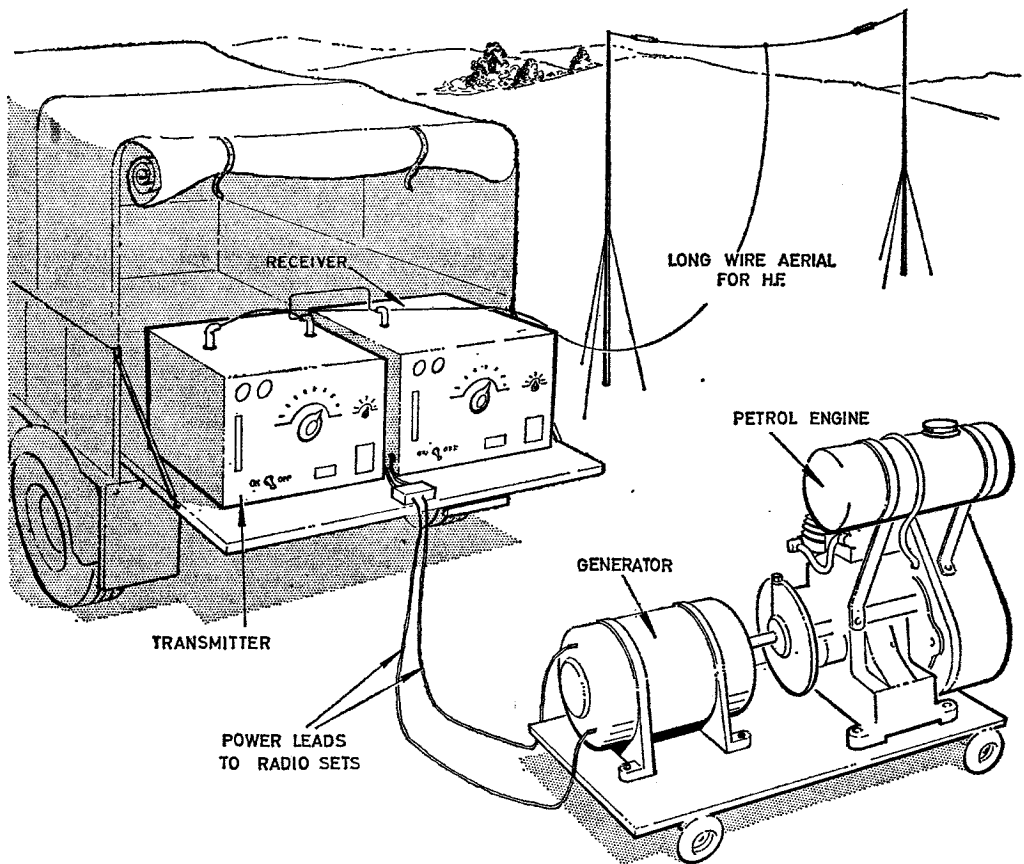

FIG. 2. MOBILE HF TRANSMITTER-RECEIVER

Larger generator sets driven by diesel engines and capable of delivering many kilowatts are used to supply self-contained mobile radio stations comprising several transmitters.

Aircraft Installations

The radio equipment carried in an aircraft can be supplied with power in much the same way. A d.c. generator is driven by an electric motor supplied from the aircraft's main 24V d.c. supply.

An alternative system, which reduces weight and also noise caused by brushes, employs a.c. generators. The a.c. is rectified by the transmitter power unit to provide d.c. power; further, where various voltage levels are required they can be obtained easily by using transformers with the correct turns ratio.

Fixed Installations

Permanent installations usually draw power from the public mains supply or they may have a local power station which produces all the necessary power. The size of the power unit required for each transmitter will depend on the r.f. power output of the transmitter.

High-power Supplies

For high power requirements it is more efficient to use a three-phase supply than a single phase one. Since the ripple frequency in a three phase rectified output is higher than in a single phase rectified output, the physical size of filter chokes and capacitors is much smaller. This is important where very high voltages and powers are concerned.

Three-phase Rectifiers

Full-wave and half-wave rectifier circuits are both used on three-phase systems. Fig. 3 shows a three-phase half-wave rectifier circuit and its action can easily be followed if the circuit path of each valve is traced out in turn.

Notice that the output ripple frequency is three times the supply frequency.
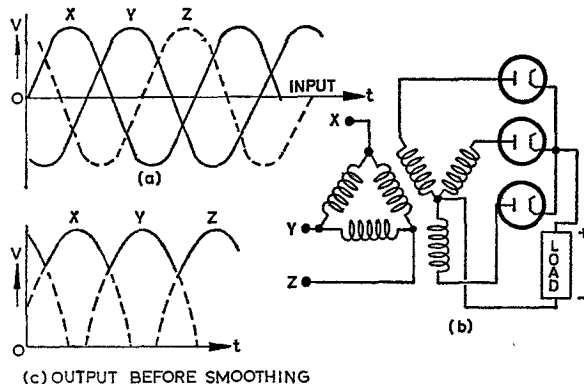
FIG. 3. THREE-PHASE HALF-WAVE RECTIFIER CIRCUIT

Three-phase Full-wave Rectifier

A full-wave rectifier circuit which could be used with a three-phase input is shown in Fig. 4.

With reference to the waveforms of Fig. 3(a) it can be seen that for a time while line X is positive, both lines Y and Z are negative. During this time current flows through V_1 to the positive side of the load, through the load, and then back through V_5 and V_6 to points Y and Z.
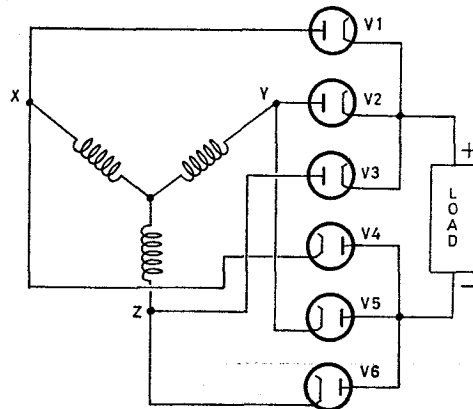
FIG. 4. THREE-PHASE FULL-WAVE RECTIFIER CIRCUIT

Types of High-power Rectifier Valve

Gas-filled valves are often used for high-power rectification. These "soft diodes", as they are called, are filled with mercury vapour or one of the inert gases. A big advantage of soft valves when used as rectifiers is that they have a low conducting resistance, a useful property when voltage regulation is important.
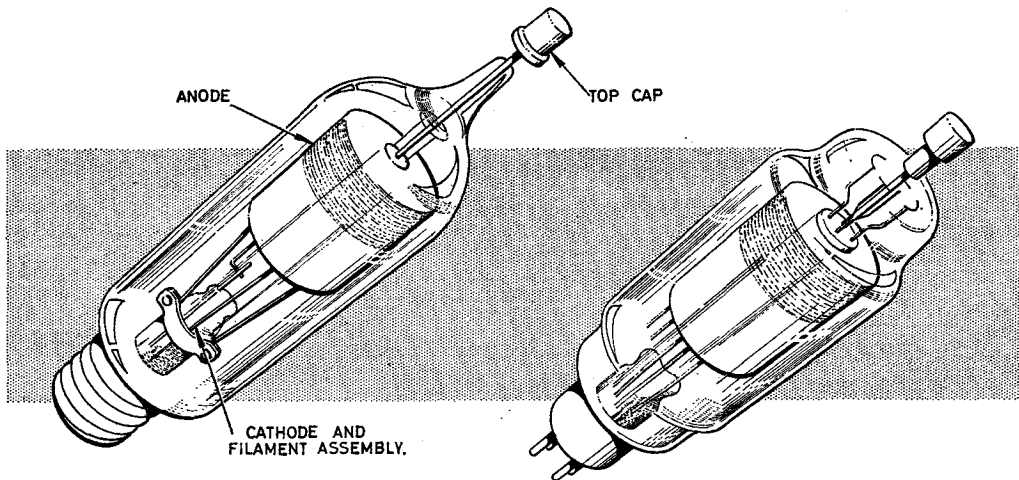
FIG. 5. GAS-FILLED DIODES

Handling Mercury-vapour Valves

The mercury-vapour inside the valve envelope is obtained by a small bead of mercury vapourizing as the valve heats up. The valves must be handled carefully so as to avoid splashing the mercury over the electrodes. Mercury-vapour valves must always be allowed to heat up before the h.t. is applied to the anode. If the h.t. supply is switched on before the cathode has reached the correct temperature, the cathode may be destroyed.

(A.L. 1, June, 1964)

Special control circuits are used to ensure automatically that the valves have reached the correct working temperature before the h.t. is switched to the anodes. Small semi-conductor rectifiers, e.g. silicon controlled rectifiers, are capable of handling heavy currents and high voltages and are replacing gas-filled and mercury rectifiers.

Thermostats

A thermostat is a device widely used in automatic delay circuits and where it is required to control temperature.

The basic construction of a thermostat is illustrated in Fig. 6. Two small strips of different metals are joined together and anchored at one end.
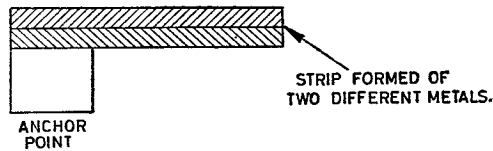
STRIP FORMED OF
TWO DIFFERENT METALS.

ANCHOR
POINT

FIG. 6. BASIC THERMOSTAT

When this "bi-metal" strip is heated the two metals expand, but one metal (the bottom one in Fig. 7a) will expand more than the other. The result is that the complete strip bends, or warps; the higher the temperature the more the bend. When the bend has reached its upper limit a set of electric contacts is closed and the required circuit is thus completed (see Fig. 7).

CONTACTS TOUCHING

TO CIRCUIT

(a) THERMOSTAT HOT

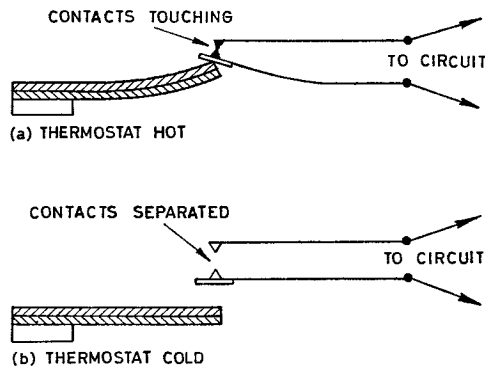CONTACTS SEPARATED

TO CIRCUIT

(b) THERMOSTAT COLD

FIG. 7. THERMOSTAT OPERATION

When the thermostat is used in a delay circuit the important factor is the time taken to heat the thermostat to the operating point, not the actual temperature required to do this.

Bi-metal strips can be made to operate at various temperatures so that the delay between switching the heating element on, and the instant at which the contacts are closed can be as required.

Thermostats used in this way are usually mounted in glass envelopes and look like small valves. The base pins carry the connections to the heater element and to the contacts. They are usually called *thermal delay switches.*
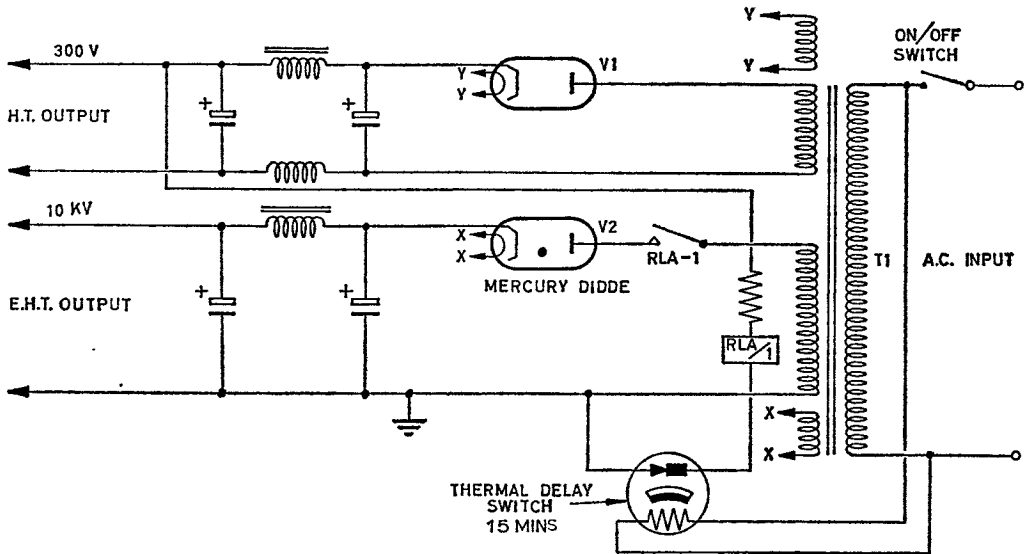
Safety Control Circuits

FIG. 8.   SIMPLE DELAY CIRCUIT

A simple circuit using a thermal delay switch to prevent the h.t. being applied to a mercury vapour rectifier valve before it reaches its operating temperature is shown in Fig. 8. V_1 is a normal hard valve half-wave rectifier which produces a d.c. h.t. voltage of 300V, and V_2 is a mercury-vapour half-wave rectifier producing an e.h.t. d.c. voltage of 10 kilovolts.   When the ON/OFF switch is closed, power is applied to the filament windings of T_1 and both valve heaters start warming up as does the heater element of the thermal delay switch. V_1 then produces a 300V d.c. output, but the a.c. h.t. is not applied to V_2 anode since contact RLA-1 of relay RLA is open.   The winding of RLA is connected to the 300V d.c. positive line through a current limiting resistor and through the contacts of the thermal delay switch.   Thus the a.c. h.t. is not applied to the mercury-vapour valve until the delay of 15 minutes is completed.   The e.h.t. output then becomes available.

Mercury-arc Rectifiers

When a rectifier is required to provide a very large current in the order of 10 to 1,000 amps a mercury-arc rectifier valve is used.   A pool of mercury in a steel container forms the cathode which, with one or more carbon anodes, is enclosed in a glass envelope (Fig. 9).   A striker anode is mounted close to the pool of mercury and a p.d. exists between this anode and the pool.   A striker electro-magnet, when energised, moves the striker anode into the pool thus short-circuiting the electro-magnet supply (Fig. 10).   The striker arm is spring-biased and as it comes out of the pool an arc is formed which generates sufficient heat to vapourize some of the mercury.   A current is thus established between the cathode and the main anodes.   Positive ions produced by the ionised mercury vapour strike the cathode and provide sufficient heat to maintain ionisation. A white-hot cathode spot is formed on the surface of the mercury and heavy currents can be drawn.   Once the main arc is struck the striker anode voltage is removed.

The main arc between anode and cathode exists only when a substantial load current is being drawn.   If the load current were significantly reduced, the arc would cease and would
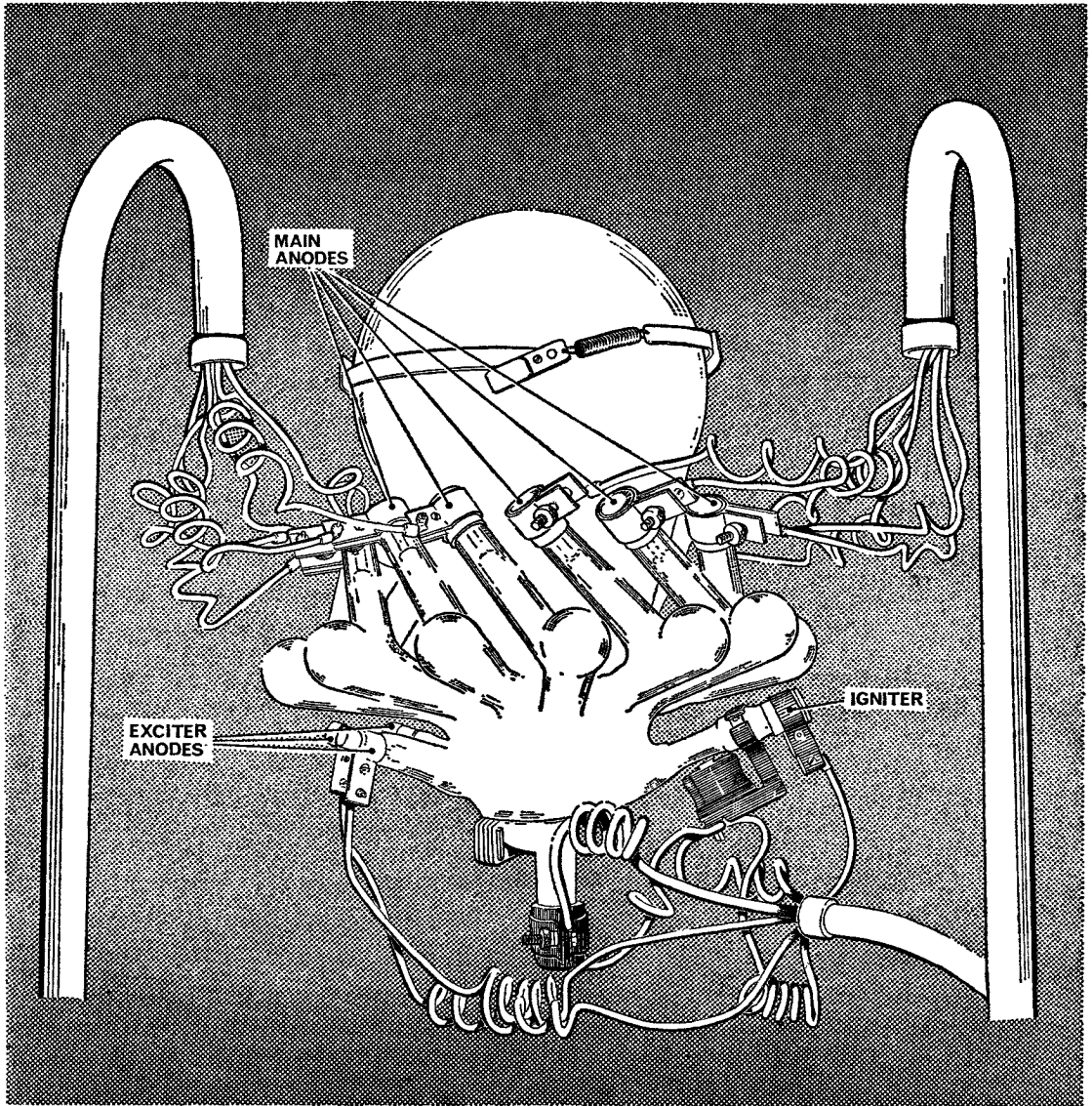
FIG. 9. MERCURY-ARC RECTIFIER VALVE

have to be struck again. To overcome this an exciter anode is inserted between the striker anode and the main anodes. The exciter anode is provided with a separate steady supply so that once a cathode spot has been formed by the striker anode it will be maintained irrespective of d.c. load changes.
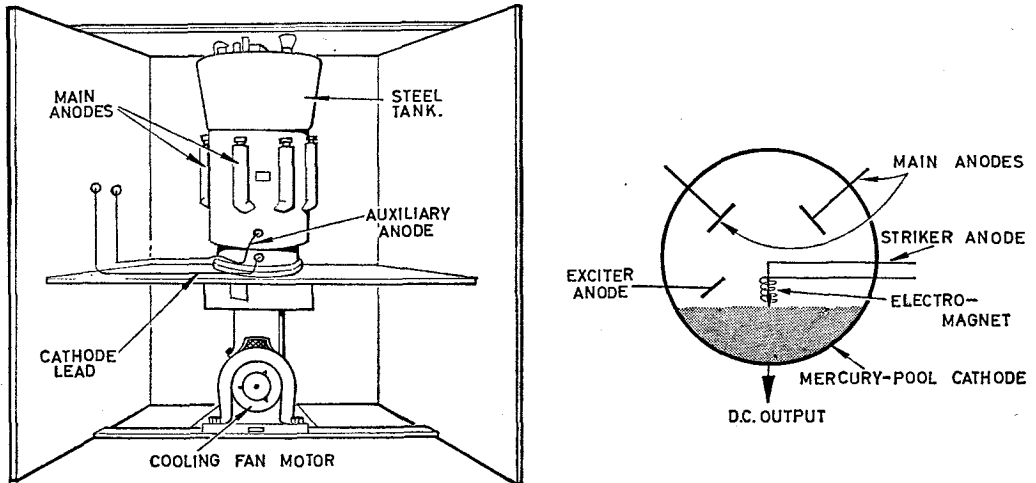
FIG. 10. A MERCURY-ARC RECTIFIER

Stabilized Power Supplies

A characteristic of a simple power supply is that as more current is drawn from it, so the output voltage falls. For example, a power unit which provides 240V at 10 mA may give only 200V if 50 mA is drawn from it. This is due to the fact that every power supply unit has an effective internal resistance. The load current must flow through this resistance, so the voltage drop across it increases as the load current increases. This voltage drop reduces the output voltage. In the above example the effective internal resistance of the power unit would be 1,000 ohms. Obviously the higher the internal resistance the greater will be the internal voltage drop.

In some electrical circuits the variation in power unit output voltage with load current is not important. In communication equipments, however, changes in supply voltage can have adverse effects; the radiated frequency of a transmitter can drift and a receiver can drift off tune.

In a transmitter, the load current varies tremendously between the key down and key up conditions and voltage stabilizers are used to keep the supply voltage steady despite these variations in load current. Two types of voltage stabilizer will be considered: the type which controls the a.c. voltage input to the power unit, and the type which controls the d.c. voltage output of the power unit.

DC Voltage Stabilization

A simple method of d.c. voltage stabilization involving the use of a gas-filled stabilizer valve was described in Part 1. The circuit is given in Fig. 11.
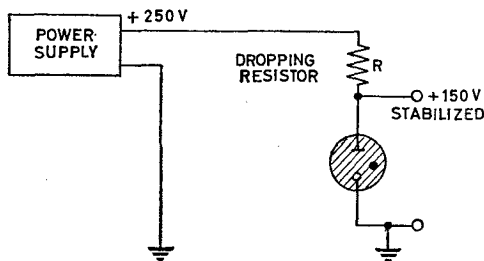

FIG. 11. SIMPLE STABILIZER CIRCUIT

87

The power unit produces a greater voltage than that required; both the load current and the current through the stabilizing valve flow through the dropping resistor (R). If the load current increases, the output voltage will tend to fall because the voltage dropped across R will increase. But the current flowing through the stabilizing valve also falls and then there is less voltage dropped across R and the output voltage rises again to 150V. Because of this action the valve current varies just enough to maintain a steady 150V output.

In effect, the valve works as an automatically variable resistor but it can operate only within a certain current range. For transmitter purposes the permissible current variation is insufficient and the output voltage is not steady enough. For the circuit of Fig. 11 the possible voltage variation may be about 4V in 150V, or 2·7%. This is not good enough for transmitter supply regulation which requires something nearer 0·5% variation.

Series Hard Valve Stabilizer

With this system a triode or beam tetrode valve is used in place of the dropping resistor. The series hard valve V_1, acts as an automatically variable resistor as did the soft valve of Fig. 11 but it can be made much more sensitive to changes in output voltage and can therefore maintain a more steady output. Fig. 12 shows the circuit.
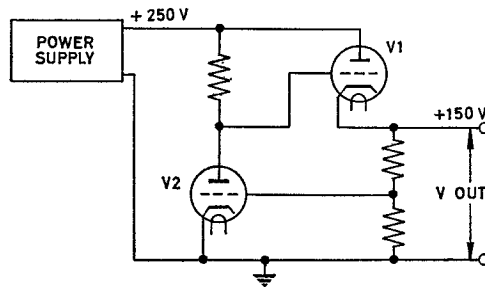

FIG. 12. SIMPLE HARD VALVE STABILIZER CIRCUIT

The voltage dropped across V_1 depends upon the grid voltage. If the grid voltage is very negative with respect to the cathode, the valve will pass only a small current and so acts as a high series resistance. If the grid is only slightly negative with respect to the cathode the valve will pass a large current and will appear as a small series resistance. V_2 is an amplifier which controls the grid voltage of V_1.

The action is as follows: suppose the output voltage tends to rise, then the grid voltage of V_2 will rise. This makes the anode voltage of V_2 fall; thus the grid voltage of V_1 falls and the effective resistance of V_1 increases. More voltage is dropped across V_1 and the output voltage falls back to 150V. The big advantage of this circuit is that it requires only a very small change in the output voltage to affect the grid voltage of V_1 because V_2 is an amplifier.

Output Voltage Adjustment

Another advantage of this circuit is that the output voltage can be set to a required value by altering the grid voltage of V_2 which then alters the grid voltage of V_1. This can be done by connecting the grid of V_2 to a variable resistor instead of to a fixed point on the resistor chain across the output terminals (Fig. 13).
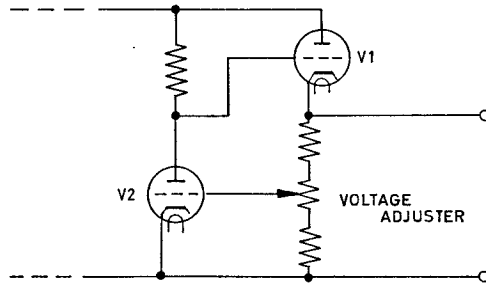
FIG. 13.   VOLTAGE ADJUSTMENT

Practical Hard Valve Stabilizer

In the basic circuits of Figs. 12 and 13, V_2 is shown with a positive grid bias.   This is unsuitable for an amplifier so in the practical circuit of Fig. 14, V_2 is provided with a positive bias on the cathode so that the grid, although still positive with respect to earth, is negative with respect to cathode.   The cathode voltage is stabilized by a simple soft-valve stabilizer.   The series stabilizer valve is often drawn on its side as in Fig. 14.
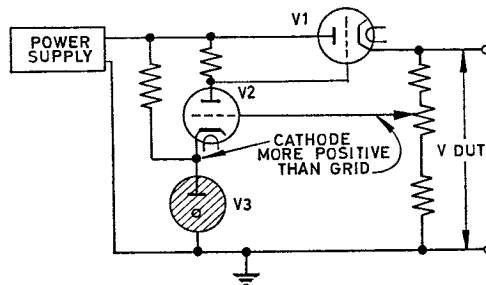
FIG. 14.   PRACTICAL VOLTAGE STABILIZING CIRCUIT

AC Stabilization

If a valve rectifier supplies a varying current to the load, then it must draw a varying current from the a.c. supply.   The same problem of voltage regulation then occurs in the a.c. generator.

One way of stabilizing a.c. supplies is by using an induction regulator, which is simply a variable transformer.   Fig. 15 shows the arrangement.
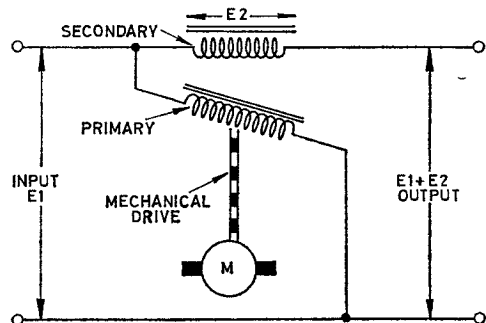
FIG. 15.   INDUCTION REGULATOR

89

The a.c. input supply is connected across the primary winding of a transformer. The primary coil can be turned about its centre point by a motor so that the coupling between it and the secondary winding is variable. When the primary is at right-angles to the secondary the voltage E_2 across the secondary is zero so that the a.c. output voltage is simply E_1, the input voltage. When the primary is turned into line with the secondary, E_2 gradually increases and the output voltage, which is the sum of E_1 and E_2, rises.

To make the voltage control automatic, the motor driving the primary coil must automatically turn it into line with the secondary if the output voltage tends to fall, and turn it out of line if the output voltage tends to rise. One way of doing this is shown in Fig. 16.
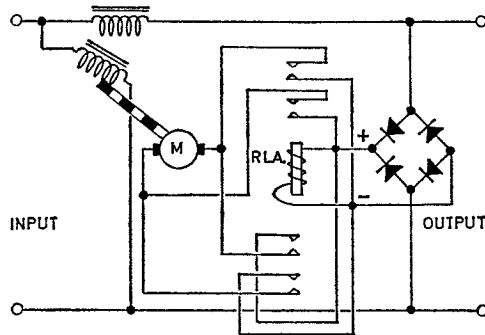

FIG. 16. INDUCTION REGULATOR WITH CONTROL CIRCUIT

A bridge metal rectifier provides a direct current, the value of which depends on the a.c. voltage output. This direct current supplies the motor through contacts operated by a relay RLA which is also energised by the direct current.

RLA is a special type of relay called a "floating" or "balanced" relay. The iron core over which the relay coil is wound is free to move up or down and is held in a central position by a coiled spring. If the current through the relay coil increases, the increased magnetic field attracts the iron core upwards. As it does so, the top set of contacts are made and the motor is energised. If the current through the relay coil decreases, the magnetic field decreases and the spring draws the iron core downwards. The bottom set of contacts then makes and the motor supply, and hence the motor rotation, is reversed.

Thus, if the a.c. voltage output (and hence the direct current) rises, the motor turns one way. If the a.c. voltage output falls the motor turns the other way. This action tends to stabilize the a.c. voltage output.

Magnetic Amplifier Voltage Stabilizer

A disadvantage of the induction regulator is that it employs moving parts which wear out and have to be replaced from time to time. The magnetic amplifier, on the other hand, can be used as a voltage stabilizer without the use of moving parts.
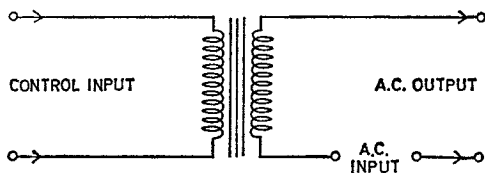

FIG. 17. PRINCIPLE OF THE MAGNETIC AMPLIFIER

Basically, this device is a variable transformer with its secondary winding in series with the a.c. supply to be stabilized. Since alternating current flows in the secondary circuit when a load is applied, the secondary coil possesses reactance. Thus a voltage will be developed across the secondary the value of which will depend upon the load current. The output voltage will therefore be somewhat less than the input voltage. In this respect the secondary winding acts in a similar manner to the dropping resistor in a simple d.c. stabilizer circuit.

The reactance of the secondary coil and hence the value of the output voltage can be varied by passing a direct current, called the control current, through the primary coil. If a very heavy control current is passed the core becomes saturated and the alternating current in the secondary cannot cause the magnetic flux in the core to change. The inductance of the secondary coil, and so its reactance, are therefore reduced. If the control current is reduced, the core is no longer saturated and the reactance of the secondary increases.

In the magnetic stabilizer, part of the a.c. output is rectified and passed through the primary coil in such a manner that an increase in output voltage causes a decrease in control current. If the a.c. output voltage tends to rise the reactance of the secondary winding rises and so the voltage drop across this winding increases thus bringing the output voltage back to its required level. If the a.c. output voltage tends to fall the control current increases and the voltage drop across the secondary winding decreases so counteracting the original fall in output voltage. In this manner the output voltage is stabilized.

SECTION 3

RECEIVERS

CHAPTER 1

RF AND IF STAGES

Introduction

The essential requirements of a radio receiver as outlined in Part 1 are sensitivity, selectivity and fidelity. In a practical receiver a compromise is struck between the ability to *receive* the signal and the ability to *reproduce it faithfully;* the former depends on selectivity and sensitivity and the latter depends on bandwidth. These fundamental points are summarized in Table 1.

Wanted	Obtained by
Sensitivity	Narrow bandwidth; many amplifier stages
Selectivity	Narrow bandwidth; choice of i.f.
Fidelity	Wide bandwidth

TABLE I. BASIC RECEIVER REQUIREMENTS

In a communication receiver the problem is more acute than in a domestic receiver and since a communication receiver is often remotely operated, the additional problem of *stability* arises. For example, an airborne receiver is pre-tuned and any one of several frequencies is selected by the pilot; the frequency of the tuned circuits selected must remain constant and not drift. The most critical circuit of the communication receiver, as far as stability is concerned, is the local oscillator and so the factors concerning basic oscillator stability, which were dealt with in Section 2, apply equally to receiver oscillators.

The Communication Receiver

Perhaps the most obvious difference between a domestic receiver and a communication receiver is that the latter has a b.f.o. to enable it to receive c.w. signals: there are, of course, many other less obvious differences which make the communication receiver a very specialised piece of equipment.

A communication receiver must be very sensitive so that it can receive signals from great distances: it must cover a wide range of frequencies and be able to select one station from several on nearby frequencies: it must be capable of receiving c.w. signals and its sensitivity and selectivity must be variable over a wide range to enable it to receive signals of variable strength and bandwidth.

A typical communication receiver which incorporates these features is shown in Fig. 1. The "service" switch in this receiver combines b.f.o. and selectivity switching.
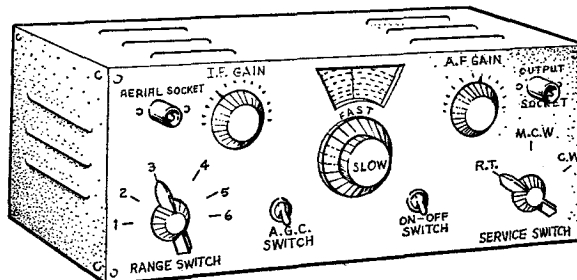

FIG. I. TYPICAL COMMUNICATION RECEIVER

(*A.L.* 2, *July,* 1964)

The total frequency range is divided into six bands, each covered by a separate scale on a circular tuning dial.

It is normal practice to use the outer scale to cover the highest frequency band, thus using the longest tuning range for the greatest frequency range. Typical frequency bands for a communication receiver are given in Table 2.

Band Number	Frequency Range (Mc/s)
1	0·6 to 1·5
2	1·5 to 3
3	3 to 6
4	6 to 10
5	10 to 15
6	15 to 30

TABLE 2. TYPICAL COMMUNICATION RECEIVER FREQUENCY BANDS

Because practical tuning capacitors have a limited variation in value, the frequency range allotted to each band is limited. The capacitance range of available tuning capacitors is approximately 9 to 1, from highest to lowest value. This gives a corresponding frequency variation (since $f \propto \dfrac{1}{\sqrt{C}}$) of about 1 to 3 from lowest to highest value. In domestic broadcast receivers it is usual to arrange that as big a frequency coverage as possible is provided by each scale range, whereas in Service communication receivers as big a scale space as possible is provided for a given frequency range.

Fig. 2 shows the block diagram of a typical communication receiver and indicates which stages mainly determine the receiver characteristics. Also shown are typical signal strengths at each stage.
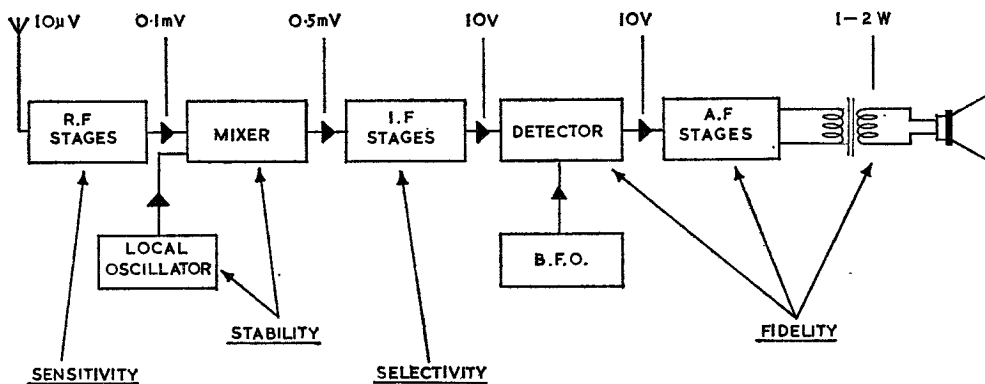

FIG. 2. SUPERHETERODYNE STAGE CHARACTERISTICS

Although a loudspeaker is shown, communication receivers often have an output stage feeding earphones. The output power is then a few milliwatts.

For the purpose of explanation, receivers are best considered in two distinct parts:—

a. Up to the detector stage.

b. From the detector stage onwards.

Using this division the i.f. stages, which are really low frequency r.f. circuits, can be considered together with the r.f. stages and will be discussed in this chapter. The non-resonant a.f. circuits will be dealt with in Chapter 2.

RF Amplifier Circuits

The aerial input signal to the r.f. amplifier stage may be only a few micro-volts and so the stage must be very sensitive. Its job is to pick out a very small signal from a background of other signals and noise. Unless the r.f. stage can lift the wanted signal clear of this background, the rest of the receiver is useless.

In order to achieve this, two things are essential:—

a. Maximum energy must be transferred from the aerial to the r.f. amplifier.

b. Noise produced by the r.f. amplifier valve must be a minimum.

Triode valves do not generate as much noise as pentodes and are therefore preferred for the first stage of the receiver, either in a grounded grid or neutralised circuit. Fig. 3 shows in outline the two basic types of r.f. amplifier circuit, one using a pentode and the other a triode valve. The critical points where noise due to bad contacts can arise are also shown.
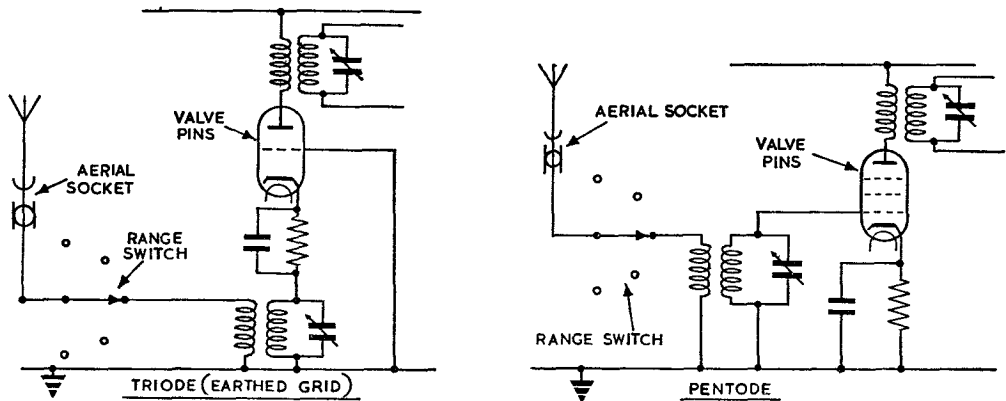

FIG. 3. THE TWO BASIC RF AMPLIFIER CIRCUITS

The main difference between the two circuits is that the grounded grid circuit provides a low input impedance. It is thus most suitable for low impedance resonant aerials such as are used at v.h.f.

An r.f. amplifier circuit using a transistor is shown in Fig. 4. The transistor should have a sufficiently high cut-off frequency. The circuit is a conventional tuned voltage amplifier. The signal is selected in the input tuned circuit and matched into the base by a transformer. The small signal current at the base causes a large variation in collector current and an amplified signal voltage is developed across the tuned circuit load in the collector.

The Cascode Amplifier

The cascode r.f. amplifier combines the high gain of a pentode valve with the low noise of a triode valve. It consists of two triodes in cascade: the first stage is a conventional grounded-

95

cathode circuit and the second stage is a grounded-grid amplifier. A simple circuit is shown in Fig. 5. Negative feedback is necessary in the first stage to avoid possible instability caused by Miller effect. The stage gain of V_1 circuit is low but the input impedance is high. Thus heavy damping of the aerial input circuit is avoided and less signal power is required to develop a voltage across the input tuned circuit than in the case of a simple grounded-grid circuit.
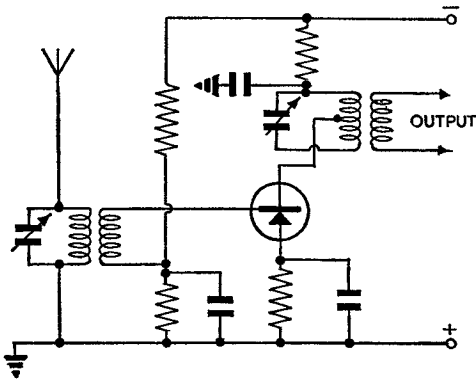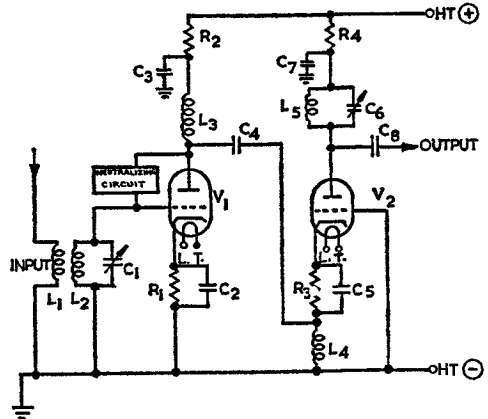
FIG. 4. TRANSISTOR RF AMPLIFIER CIRCUIT

FIG. 5. CASCODE AMPLIFIER

Number of RF Stages

The number of r.f. stages that can be usefully employed in a receiver is limited by problems of instability. Large voltages at high frequencies cause feedback between anode and grid circuits. This is particularly so in multi-band h.f. receivers where many switched wavebands are involved and wiring capacitances are high. In a superhet receiver it is not necessary to use more than two r.f. stages. Most communication receivers use only one and domestic receivers, which can rely on a comparatively strong input signal, usually do not have an r.f. stage at all. In fact an r.f. amplifier has very little effect on the overall receiver response *except in the case of very weak signals*. Fig. 6 shows the effect of an r.f. stage in terms of gain and selectivity.

The main purpose of an r.f. amplifier is to prevent unwanted signals from reaching the rest of the receiver. It suppresses *second channel* or *image* interference and *harmonic* interference.
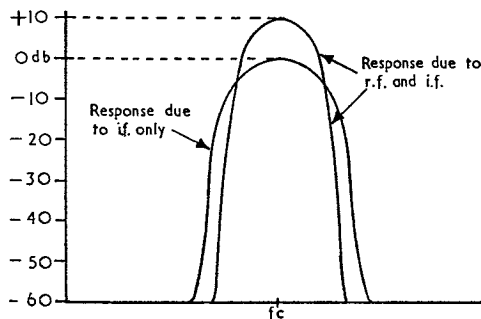
FIG. 6. EFFECT OF RF STAGE

Fig. 7 shows how an r.f. amplifier removes (*a*) an image interference signal, and (*b*) a harmonic interference signal.

IF Rejection

It is also important that the r.f. stage of a receiver rejects any signals at the i.f., particularly if part of the receiver frequency range approaches the i.f. This form of interference can be so severe in practice that it is usually avoided by ensuring that the receiver does not tune over the region of the i.f. For example, a domestic receiver with an i.f. of 465 kc/s would not tune over the range 300 to 500 kc/s.

An alternative method of rejecting a signal at the i.f. is to use an *i.f. wave trap*. Two possible circuits are shown in Fig. 8.

L_1 and C_1 resonate at the i.f. and are screened to avoid stray pick-up. In both cases L_1 and C_1 are pre-set to provide minimum receiver output at the unwanted frequency.

(a)   SUPPRESSION OF IMAGE INTERFERENCE

(b)   SUPPRESSION OF HARMONIC INTERFERENCE

FIG. 7.   INTERFERENCE SUPPRESSION BY USE OF A RF AMPLIFIER

The Mixer Circuit

There are two types of mixer circuit: additive and multiplicative. In the additive mixer the input signal and the local oscillator signal are added to form a common input to *one* electrode. With multiplicative mixing the two inputs are fed to *separate* electrodes and thus have independent control of the valve current. This results in the product of the two inputs being present in the output; hence the term multiplicative.

The efficiency of both types of mixer is given in terms of output *at the i.f.* for a given input *at the r.f.* The expression:

$$\frac{\text{anode current at i.f.}}{\text{input signal voltage at r.f.}}$$

is known as the *conversion conductance* (g_c) of the mixer.

Fig. 9 shows the two basic valve mixer circuits, which are similar to grid and suppressor modulation circuits. As with a modulator circuit the valve of the additive mixer must be biased so as to operate on the curve of the valve characteristic in order to produce the required i.f.

The output from the mixer circuit will be at many different frequencies, including the sum and difference frequencies of the two inputs. Since the superhet principle is based on reducing the input frequency, it is always the difference frequency which is selected as the output i.f.
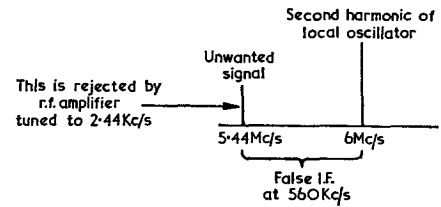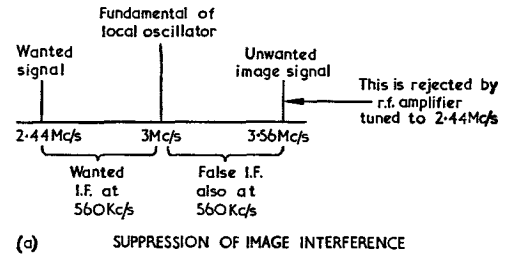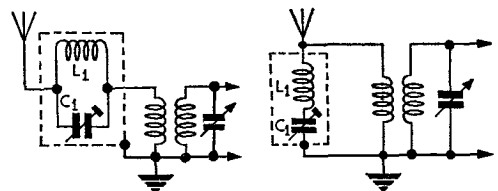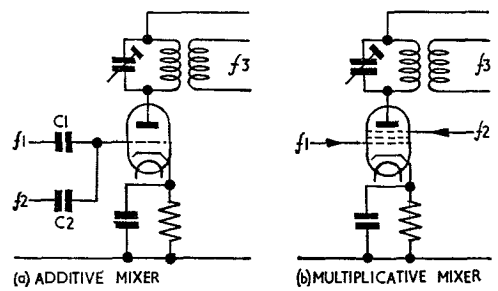

FIG. 8.   IF WAVE TRAPS

FIG. 9.   BASIC MIXER CIRCUITS

97

Choice of Mixer

The type of mixer circuit used in a particular receiver depends on two main factors: the frequency of the incoming signal and the frequency of the required i.f. The higher the signal frequency the more troublesome become inter-electrode capacitances and valve noise. The frequency of the i.f. determines the amount of frequency separation between the signal and local oscillator and therefore the amount of interaction between these two input circuits. Another factor which determines the amount of interaction is the coupling that exists between the two circuits. Fig. 10 shows the equivalent circuits for the circuits of Fig. 9 and illustrates this inter-coupling.
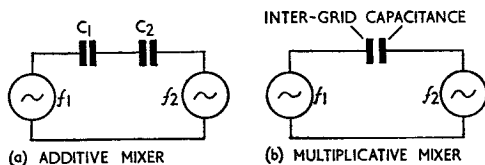

FIG. 10. COUPLING BETWEEN MIXER INPUTS

For minimum interaction, either the two input frequencies must be far apart, necessitating a high i.f., or the coupling must be small, requiring small coupling capacitors.

At l.f., m.f. and h.f. special multiplicative mixer valves such as the heptode, hexode and triode-hexode are used. These reduce the inter-grid coupling to a few picofarads and enable a selective, high-gain *low*-frequency i.f. amplifier to be used. They also allow the combination of oscillator and mixer in one valve envelope making a single valve frequency changer.

At v.h.f. and u.h.f. even a few picofarads represents considerable coupling and the degree of interaction is then reduced by using a much higher i.f. At these frequencies valve efficiency is fairly low and it is usual to employ additive mixers and a high i.f. Additionally a buffer stage is often used to isolate the oscillator and signal circuits. This buffer often takes the form of a frequency multiplier stage as shown in Fig. 11.
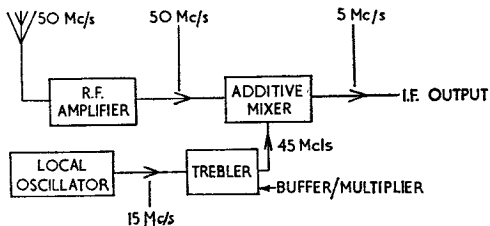
FIG. 11. FREQUENCY CHANGING AT VHF

Balanced Mixers

At s.h.f. conventional valves cannot be used in the mixer stage and additive mixing using silicon crystal diodes, and a high i.f. (about 45 Mc/s) is employed. The local oscillator valve is a velocity modulated device known as a klystron and unfortunately it introduces into the mixer stage a high level of noise. This is most undesirable as it reduces the signal-to-noise ratio.

Most of the noise in the local oscillator signal can be cancelled out in the mixer stage if a balanced mixer circuit is used. This circuit employs two crystal mixers connected so that the aerial signal is in phase at each crystal and the local oscillator voltages are in anti-phase at each

crystal. With this arrangement the components of local oscillator noise in the i.f. output from the mixers cancel, while the signal components add together. (See Fig. 12.)

Thus in a balanced mixer circuit the noise generated by the local oscillator is not present in the signal fed to the i.f. amplifiers. This results in an increased signal-to-noise ratio.
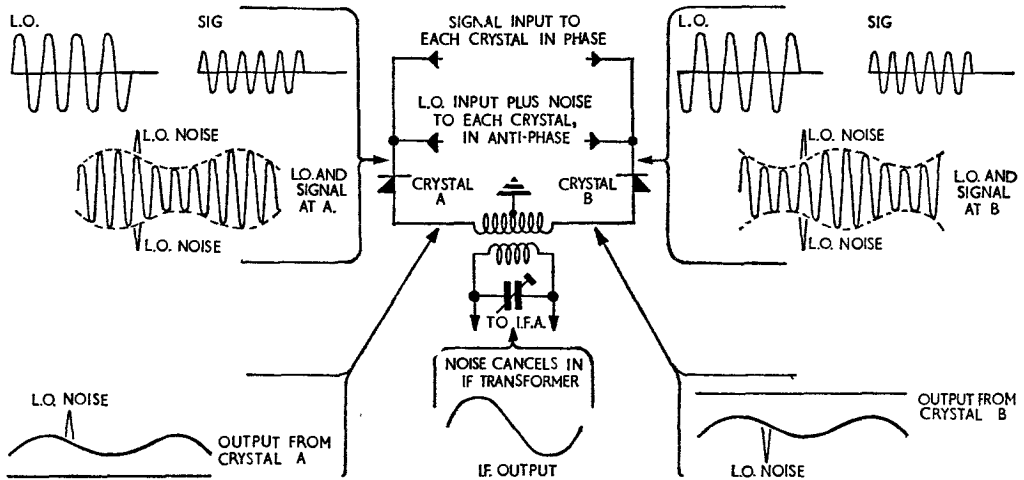
FIG. 12. BALANCED MIXING

Transistor Frequency Changer

A basic circuit of a frequency changer stage using a transistor is shown in Fig. 13. It is a self-oscillating additive type frequency changer.

A Meissner oscillator is used with feedback from the collector (L_3) to the emitter (L_2) via the tuned circuit which controls the local oscillator frequency.

The input signal from the r.f. stage is coupled through C_1 to the base of the transistor where it mixes with the local oscillation to produce the i.f. The tuned circuit (I.F.T.1) in the collector selects the i.f. and passes it to the first i.f. amplifier stage.

Bias is obtained from the resistance network provided by R_1, R_2 and R_3 with their associated capacitors. The collector is de-coupled by $R_4 C_2$.

Local Oscillator

The local oscillator circuit can be any one of the basic r.f. oscillator circuits. The important factor to be considered in the choice of circuit is that it must be stable over the required frequency range.

Choice of Local Oscillator Frequency

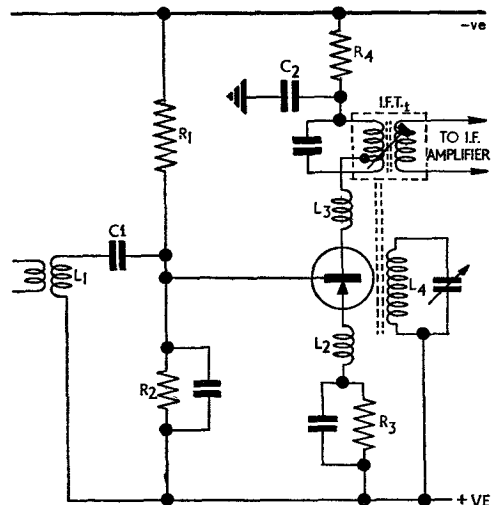Another important problem is whether the local oscillator frequency should be set above

FIG. 13. BASIC TRANSISTOR FREQUENCY CHANGER

99

or below that of the signal. For example, to receive a signal of 1 Mc/s and produce an i.f. of 465 kc/s, the oscillator could be set to either 535 kc/s or to 1,465 kc/s.

The tuning capacitors of the local oscillator and signal tuned circuits in a superhet are normally ganged. Thus their tuning ratio, i.e. the ratio of maximum to minimum frequency they will tune to, must be considered.

Fig. 14 shows the initial stages of a receiver, one band of which is required to cover from 1 to 3 Mc/s. It is assumed that the signal circuits have the required tuning ratio of 3:1.
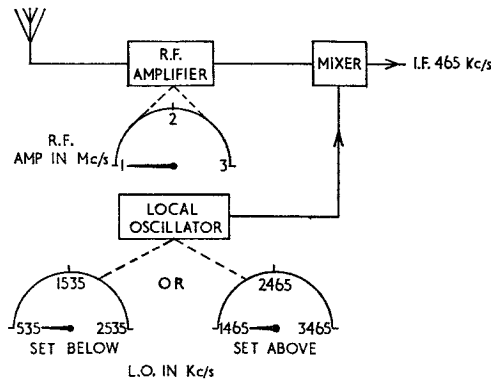

FIG. 14. CHOICE OF LOCAL OSCILLATOR FREQUENCY

With the oscillator set below the signal frequency the oscillator tuning capacitor must be able to tune from 535 kc/s to 2,535 kc/s, a tuning ratio of approximately 5:1. Thus the tuning capacitor should have a maximum to minimum capacitance ratio of 25:1. This is quite impracticable.

If the oscillator is set above the signal frequency its tuning capacitor must cover the frequency range 1,465 to 3,465 kc/s, a tuning ratio of 2·3:1. This requires a tuning capacitance ratio of approximately 5:1, which is easily obtainable. For this reason the local oscillator is usually set above the signal frequency in l.f., m.f. and h.f. receivers. In v.h.f., u.h.f. and s.h.f. receivers the tuning ratio is much lower and the local oscillator frequency can be set below the signal frequency, so giving increased frequency stability.

Padding and Tracking

For correct operation the signal and oscillator circuits should remain at the same frequency difference (equal to the i.f.) as the ganged capacitors are swept through the tuning range.

As the two variable capacitors used are normally the same, some compensation is necessary to allow for the different frequency ranges. Usually a pre-set *padder* capacitor is included in *series* with the oscillator tuning capacitor as shown in Fig. 15.
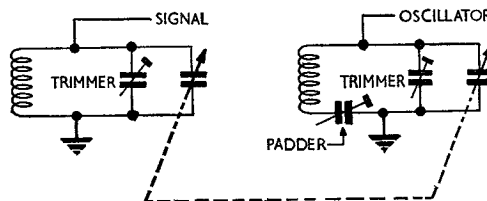

FIG. 15. PADDERS AND TRIMMERS

100

The padder is adjusted so that the difference frequency between signal and local oscillator is exactly right at a set point near the lower end of the frequency scale (see Fig. 16).

It is also necessary to make further adjustments with *trimmer* capacitors connected in *parallel* with each ganged tuning capacitor. These trimmers are usually adjusted so as to make the difference frequency exactly right at a set point near the *top end* of the frequency scale. Each of the r.f. tuned circuits will contain a trimmer in order to compensate for the various values of stray capacitances present in the different circuits.

Most service receiver A.P's list the exact points at which the difference frequency must be correct. Adjustment at the centre point of the tuning range is usually done by means of the inductance pre-set control.

The procedure for arranging the best compromise is called *tracking*. Fig. 16 shows a typical tracking error curve, and indicates the components which must be adjusted at the three alignment frequencies.
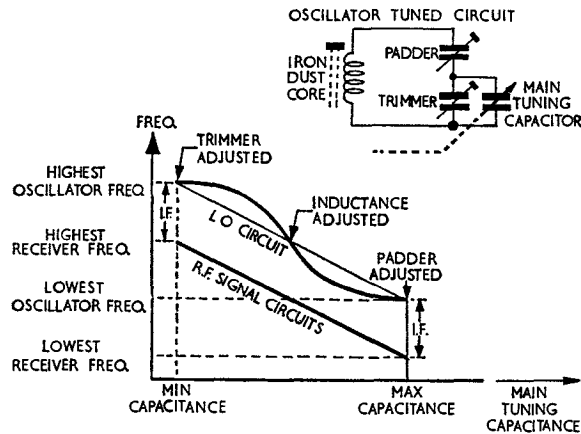
FIG. 16. SUPERHET TRACKING

The IF Stages

An i.f. amplifier is basically a r.f. amplifier operating at a *low fixed* radio frequency. Because the frequency is fixed the stage is always operating under optimum conditions and so the sensitivity, selectivity and stability are much higher than is the case in a r.f. amplifier. In a communication receiver there are usually several i.f. amplifiers in cascade. These stages form the i.f. strip. It is in this part of the receiver that most of the amplification is achieved.

A typical valve i.f. amplifier and one using a transistor are shown in Fig. 17. The valve circuit is quite conventional and is discussed in detail in Part 1. The transistor circuit uses a transistor with a suitable cut-off frequency. Stabilizing bias is provided by a potential divider network. The signal at the i.f. from the previous stage is applied to the base of the transistor. The high output impedance of the previous transistor is matched to the low input impedance of the next via a step-down transformer. Amplified voltages at the i.f. are developed across the collector tuned circuit and are fed to the next stage. Because of the fairly high capacitances which exist between collector-base and base-emitter, neutralization is provided with C. This capacitor can be of a fixed value, or since the feedback may vary if the transistor is replaced during servicing, it can be a pre-set variable capacitor.

Some of the features incorporated in an i.f. amplifier will now be considered.
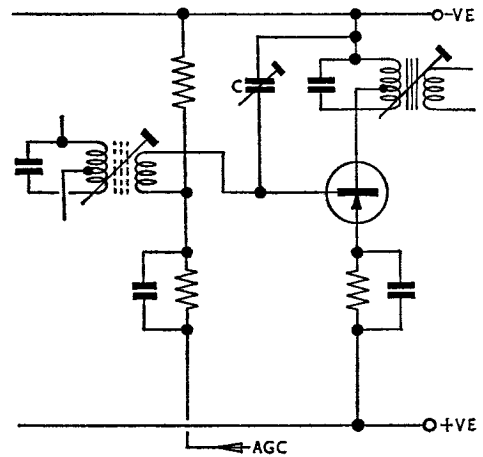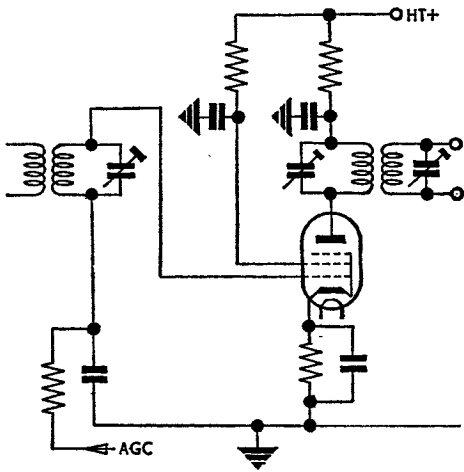
FIG. 17. BASIC IF AMPLIFIER STAGES

Band-pass Coupling

Since all the i.f. tuned circuits are pre-set their magnification factor (i.e. 'Q') can be quite high and because of this they would have a narrow bandwidth. Thus although the tuned circuits of the r.f. stage will pass the required sidebands, the i.f. stages must use *band-pass coupling*. This usually takes the form of a double-tuned transformer in which coupling is just sufficient to give the required overall frequency response.

Fig. 18 shows the difference in circuit and in frequency response for r.f. and i.f. stages.

The dotted curve of Fig. 18 shows the i.f. response when inadequate coupling, such as that provided by a single tuned circuit, is used. The degree of coupling is determined by the spacing between the tuned circuits and this is set during manufacture of the transformer. Fig. 19 shows the constructional details of a typical i.f. coil unit.

Notice that although the unit is referred to as a transformer, it is really two separate r.f. tuned circuits and does not look very much like the usual transformer. In some cases the coils are not even mounted on the same former, but re merely placed near each other.
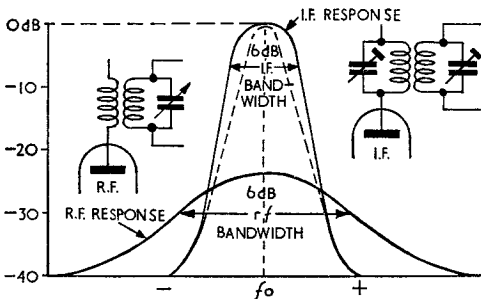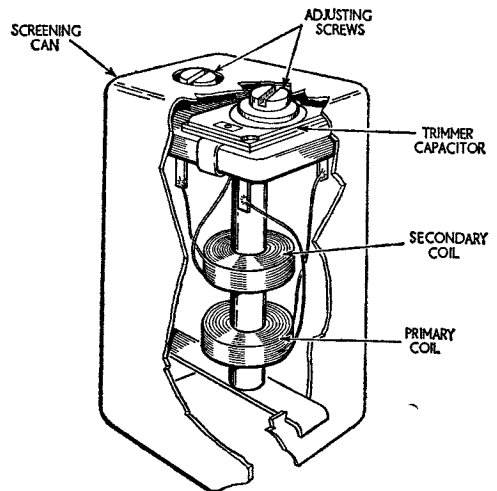

FIG. 18. COMPARISON OF RF AND IF STAGES

FIG. 19. CONSTRUCTIONAL DETAILS OF A TYPICAL IF TRANSFORMER

102

As with other r.f. circuits, tuning is sometimes achieved by varying the inductance of the coil. For a low i.f., iron dust cores are used since iron increases inductance and reduces coil size. For a high r.f. the losses in an iron dust core would be excessive and so brass or copper cores are used. These cores reduce the inductance and thus raise the frequency of the tuned circuit.

Choice of IF

The frequency chosen for the i.f. is determined partly by the signal frequency and partly by the bandwidth of the signal being received. The advantages of a superhet are based on the fact that amplification is carried out at a relatively low frequency, but if amplification at a higher frequency can be satisfactorily achieved, several advantages are apparent.

The choice of the i.f. is a compromise. The larger the value of the i.f. then the further away is the image channel interference and so this interference is easily removed by the r.f. stage. If the i.f. is made higher, adjacent channel interference becomes difficult to eliminate.

Thus to obtain good image and adjacent channel rejection plus adequate bandwidth the i.f. should be both high and low. This paradox is solved by using a double superhet receiver.

The Double Superhet

A block diagram of the initial stages of a double superhet receiver is shown in Fig. 20. The incoming signal is first changed to a *high i.f.*, which is lower than the signal frequency but high enough to provide adequate image rejection. After amplification at this frequency a second mixer is used to produce a final *low i.f.* Then follows the normal sequence of stages.
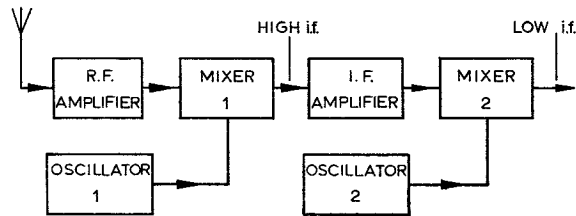
FIG. 20. THE DOUBLE SUPERHET

Notice that since the i.f. input to the second mixer has a fixed centre frequency, the second oscillator can be pre-set. It does not have to be ganged to the tuning capacitors of the r.f. and first oscillator stages. For example, if the first i.f. is 2 Mc/s and the second i.f. is 200 kc/s, the second oscillator could be pre-set to 2·2 or 1·8 Mc/s.

Typical signal and intermediate frequencies for various types of input signal are shown in Table 3.

Type of Signal	Signal frequency	IF
Broadcast (a.m.)	1 Mc/s	465 kc/s
Broadcast (f.m.)	90 Mc/s	10·7 Mc/s
Broadcast (t.v.)	150 Mc/s	34 Mc/s
Communication	10 Mc/s	600 kc/s
Communication	300 Mc/s	{ 15 Mc/s 1st IF { 2 Mc/s 2nd IF

TABLE 3.   TYPICAL IF'S

103

The high i.f's used for f.m. and t.v. broadcasting are necessary because of the wide band-width required for these services.

Variable Selectivity

In order to cater for reception of both wide-band and narrow-band signals the selectivity of a communication receiver is often made variable. Normally the i.f. circuits are controlled, usually by varying the coupling between the circuits. There are two common methods of doing this; either the *degree* of coupling can be varied or the *type* of coupling can be changed. Fig. 21 shows a circuit using variable mutual inductive coupling with typical i.f. response curves.
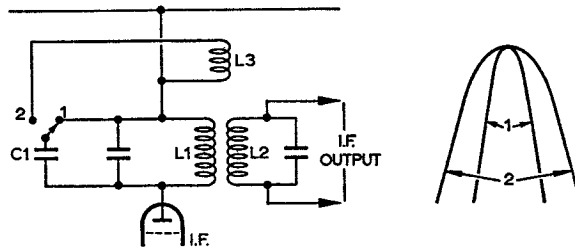
FIG. 21. VARIABLE SELECTIVITY

When in position 1 the selectivity switch connects C_1 across the i.f. tuned primary making it resonant at the i.f. Coupling is then by mutual inductance between the two coils L_1 and L_2.

When in position 2 the selectivity switch connects C_1 to the tertiary winding L_3. This increases the mutual inductive coupling and so broadens the bandwidth.

Crystal IF Filter

A quartz crystal is in effect a high 'Q' resonant circuit. It has a high degree of selectivity and therefore a very narrow bandwidth, of the order of a few hundred cycles per second. It can be used as a coupling component in an i.f. amplifier to provide a very narrow bandwidth and so increase the selectivity of a communication receiver. Fig. 22 shows a suitable circuit and its equivalent circuit.

FIG. 22. BASIC CRYSTAL COUPLING
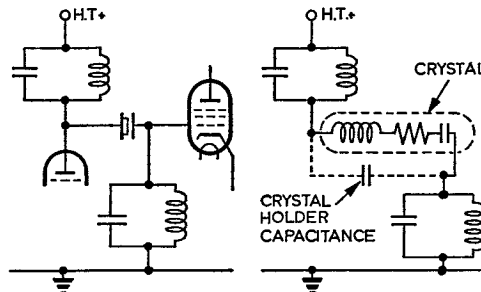
At frequencies within the narrow bandwidth of the crystal, maximum coupling between the two stages occurs, since the crystal is then a low impedance path. Outside this narrow frequency range the crystal impedance is high and coupling is negligible.

In order to overcome the effect of the crystal holder capacitance (see Fig. 23). neutralization is employed in a practical circuit.

Fig. 24 shows a suitable circuit arrangement. The tuned circuit capacitance is centre-tapped to earth and the neutralizing capacitor C_n is adjusted so as to just neutralize the crystal holder capacitance. This moves the dip of the response curve at X (Fig. 23) outside the critical frequency band.
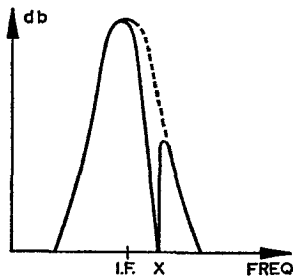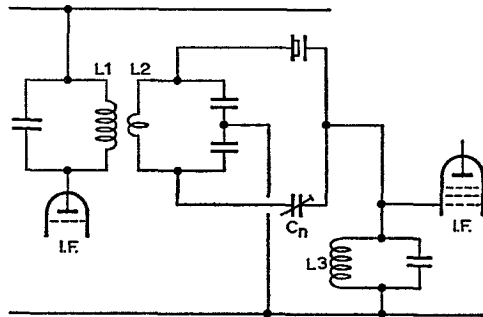
FIG. 23. EFFECT OF PLATE CAPACITANCE

FIG. 24. PRACTICAL IF CRYSTAL FILTER

Variable Sensitivity

It is necessary to have a variable control over the sensitivity of a communication receiver so that weak and strong input signals can be received equally well. To achieve this the gain of the i.f. amplifiers is usually made variable. Thus whereas the domestic receiver merely provides control of volume by varying the a.f. voltage. the communication receiver provides a separate control of both i.f. and a.f. gain. In effect this allows separate control of both sensitivity and sound output, the former depending on signal strength and the latter on sound output required.

Vari-mu valves are used in the i.f. stages and the grid bias is varied either by a potentiometer, which provides manual gain control, or by a.g.c. voltage. Typical bias values are from a minimum of $-3V$ to a maximum of $-30V$.

In some communication receivers gain control of both the r.f. and i.f. stages is necessary, and the bias on both these stages is therefore varied. However, in order to maintain maximum receiver sensitivity, the bias network is so arranged that r.f. gain is not reduced until after an appreciable reduction has been made in i.f. gain. In the simplest case the r.f. stages are fed with only a fraction (say one quarter) of the full control bias. Typical values of bias voltage and gain reduction are shown in Fig. 25.
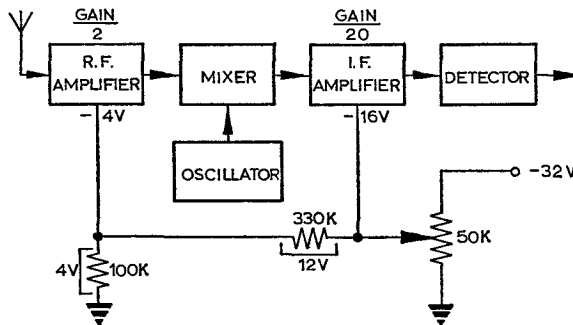
FIG. 25. RF AND IF GAIN CONTROL

Notice that it is usual *not* to vary bias to the mixer stage because it contributes very little to the overall gain. Furthermore, if the gain of this stage were varied it would affect the frequency stability of the oscillator, since there is inevitably some coupling between mixer and oscillator.

If a.g.c. is used the a.g.c. voltage merely replaces the -32 volt bias supply and potentiometer of Fig. 25.

CHAPTER 2
AF AND DETECTOR STAGES

Introduction

The remaining stages of a communication receiver after the i.f. strip, are the detector (sometimes called the demodulator) and the output stage. The detector extracts the a.f. component of the i.f. input signal and applies it to the a.f. stage where it is amplified and matched into the output loudspeaker or telephones.

Various other functions are also accomplished in this section of the receiver. These include the production of a.g.c. voltage and, if required, of beat frequency oscillations. Tuning indication, noise limiting and receiver muting circuits are also provided if the function of the receiver requires them. This chapter will cover these requirements.

Audio Frequency Response

One of the main differences between a.f. stages of a communication receiver and those of a domestic receiver is the reduction in a.f. bandwidth of the former. For good reproduction in a domestic receiver a fairly large a.f. bandwidth is required (e.g. 50 to 8,000 c/s). The efficiency of the detector and the a.f. amplifier can be considerably increased if their frequency response is narrowed. A communication receiver has a typical a.f. response of 300 to 3,000 c/s as shown in Fig. 1.
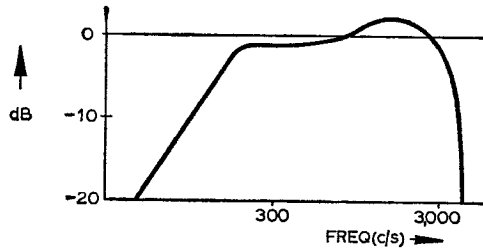
FIG. I. RT RESPONSE OF A COMMUNICATION RECEIVER

This curve represents the response for reception of r.t. signals. For code transmission of m.c.w. or c.w. the bandwidth can be even further reduced with a consequent increase in selectivity and sensitivity.

The Detector

As in most superhets the detector stage of a communication receiver is usually a diode. Provided the input to the diode is of the order of volts it is a good simple detector and introduces very little distortion. Fig. 2 shows the two basic diode detector circuits, the series and shunt arrangements.
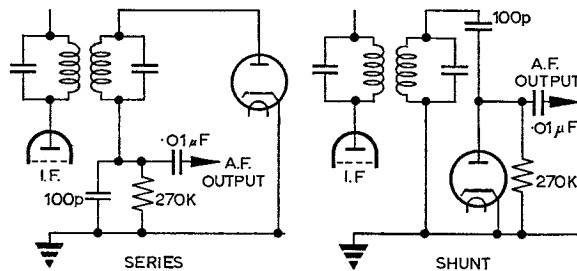
FIG. 2. DIODE DETECTORS

107

The shunt-fed circuit has the advantage over the series fed circuit that there is no d.c. path through the detector from the input tuned circuit. However, its disadvantage is that damping of the preceding stage is greater, since the valve is in parallel with the load, making the input resistance low. This results in a reduction in sensitivity and selectivity of the last i.f. amplifier stage.

Notice that both the circuits of Fig. 2 have the load resistor and capacitor in the anode circuit. This results in a negative-going output, and also allows the cathode to be earthed. This is useful when the detector stage is part of a multi-purpose valve.

It should be noted that in both circuits the diode load resistor is shunted by the coupling capacitor and the input impedance of the audio stage. Thus under dynamic conditions the effective value of detector load impedance will vary, introducing distortion. This effect can be reduced by placing a cathode follower, which has a high input impedance, between detector and audio stages.

Semiconductor germanium crystal diodes have been developed for use as a final detector in place of the valve diode. They can handle input voltages of up to 10 volts and can withstand an inverse voltage of more than 50 volts. They do not require a heater supply and have a low forward resistance (about 150 ohms at 1 volt) with a back resistance the same as that of a valve diode (about 100,000 ohms at 1 volt). The inter-electrode capacitance is much smaller than in a thermionic diode and the physical size is about that of a small resistor.

This type of semiconductor diode cannot be used at low signal levels since the noise factor is high; the maximum operating temperature is about 75°C.

Fig. 3 shows a basic superhet arrangement of combined detector and a.f. stage. A double-diode-triode valve is used, the detector load consisting of R_L and C_L. The i.f. filter consists of R_f and C_f, C_c is the normal a.f. coupling capacitor and RV the a.f. gain (or volume) control.
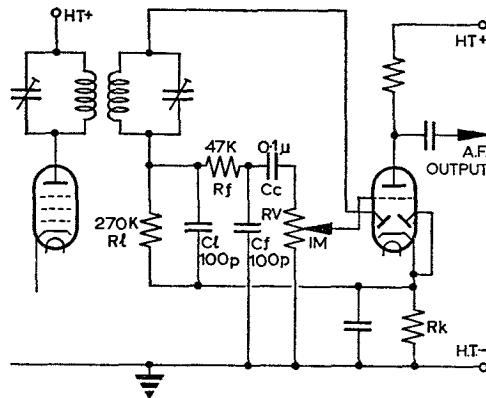
FIG. 3. DIODE DETECTOR AND FILTER

RV is taken to the earth line while R_L goes to the cathode. Thus the detector diode is not affected by the bias developed across R_k which only determines the triode operation. The anode of the unused diode is connected to cathode.

The waveforms associated with the signal during detection are given in Fig. 4. Individual i.f. cycles cannot be shown to scale and in fact appear as a blur on the oscilloscope.

CW Reception

In some communication receivers provision is made for reception of c.w. signals. This involves injecting the output from a beat frequency oscillator (b.f.o.) into either the last i.f.

ACROSS TUNED CIRCUIT   ACROSS Rℓ Cℓ   ACROSS Cf Rv

FIG. 4.   DETECTOR WAVEFORMS

amplifier stage or into the detector stage.   The frequency of the b.f.o. must differ from the i.f. by an audio frequency.   For example, if an i.f. of 465 kc/s is used the b.f.o. would be set to 466 or 464 kc/s.   This frequency would then heterodyne with the i.f. and produce an audio note of 1 kc/s in the headset.

Two possible circuits are shown in Fig. 5.   A Hartley oscillator is often used as the b.f.o. and the frequency can be varied slightly by means of the pre-set capacitor C_2.   This gives a

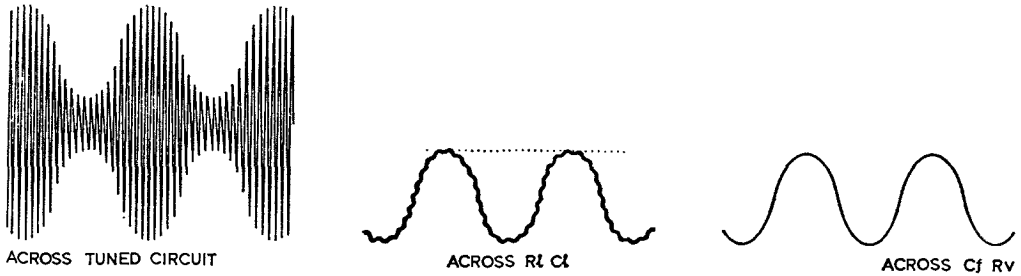BFO INJECTED INTO LAST I.F.A.   BFO INJECTED INTO DETECTOR
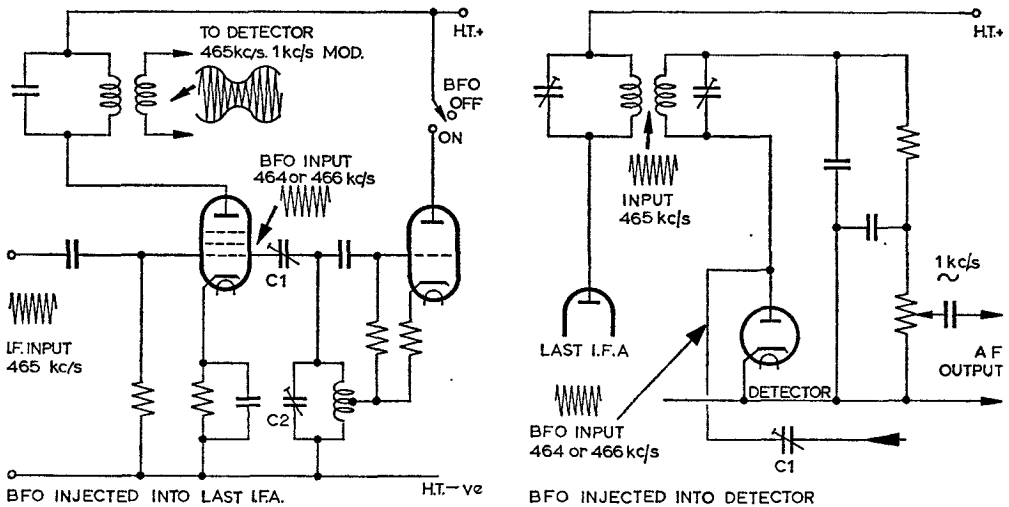
FIG. 5.   RECEPTION OF CW

variation in the pitch of the resulting note.   The amplitude of the b.f.o. signal is constant but can be adjusted for optimum value by varying C_1.   To enable r.t. or m.c.w. signals to be received without the heterodyne action, an ON-OFF switch is provided for the b.f.o.

Automatic Gain Control (AGC)

The a.g.c. system of a receiver is a circuit which reduces gain in proportion to input signal strength.   It thus provides an output which tends to remain constant despite fading and tuning changes.

The simple a.g.c. system outlined in Part 1 has two disadvantages which make it quite inadequate for use in communication receivers.   Firstly, it reduces gain for *all* input signals

109

including the very weak ones. Secondly, in reducing receiver gain it also reduces its own controlling effect. In fact this simple system is most sensitive to weak signals rather than strong ones. In order to overcome these limitations and to provide a better control, two variations have been developed.

The first improvement is that a small bias voltage is applied to the a.g.c. valve to delay the production of an a.g.c. voltage until the input signal exceeds a certain minimum amplitude. Thus small input signals below this minimum amplitude receive the full amplification of the receiver. This system is known as *delayed a.g.c.*

The second improvement, *amplified a.g.c.*, increases the controlling effect of the system by amplifying the a.g.c. voltage before it is applied to the r.f. and i.f. stages. In most communication receivers a combination of these two systems is used providing amplified delayed a.g.c. Fig. 6 shows the input/output characteristics of the various systems.

Without a.g.c. the receiver output rises in proportion to the input until receiver saturation is reached at input V_s.

With simple a.g.c. the output is approximately proportional to the input but amplification is reduced for all signals.

With delayed a.g.c. full gain is effective for small inputs below the threshold voltage V_t. Above V_t the output is practically constant over a large range of inputs.

Notice that, as with any self-controlled system, even an amplified a.g.c. voltage cannot entirely prevent a change in output. There must be some change to operate the a.g.c. system, but by amplifying the a.g.c. voltage the necessary change can be made very small. Typical circuits used to produce delayed and amplified a.g.c. are considered in the following paragraphs.
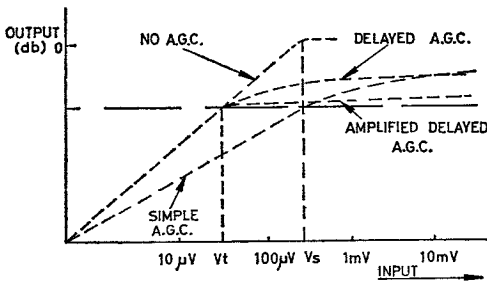
FIG. 6. AGC CHARACTERISTICS
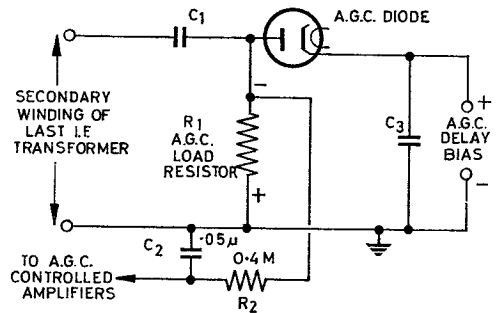
FIG. 7. BASIC DELAYED AGC CIRCUIT

Delayed AGC

A simple circuit for producing delayed a.g.c. is shown in Fig. 7. The cathode of the a.g.c. diode is supplied with a positive bias. As the anode is taken to earth through the a.g.c. load resistor R_1, the cathode is positive with respect to the anode by the bias voltage. The output from the secondary winding of the last i.f. transformer is applied, via C_1, between anode and earth.

If the bias voltage is, say, 2 volts, then the a.g.c. diode will not conduct until the peak i.f. voltage at the anode exceeds 2 volts. When this occurs the 2 volt delay bias is overcome, the a.g.c. diode conducts and a voltage is developed across R_1 with the polarity shown in Fig. 7. The action is illustrated in Fig. 8. This voltage forms the negative a.g.c. bias to the vari-mu valves and is taken to the control grids of these valves via the filter formed by C_2R_2. In this way

110

the receiver gain is made inversely proportional to the signal strength for i.f. voltages that exceed the delay bias.

An alternative method of feeding the a.g.c. diode is shown in Fig. 9. The detector diode is fed, as before, from the secondary winding of the last i.f. transformer, but the input to the a.g.c. diode is taken from the anode of the last i.f. stage.
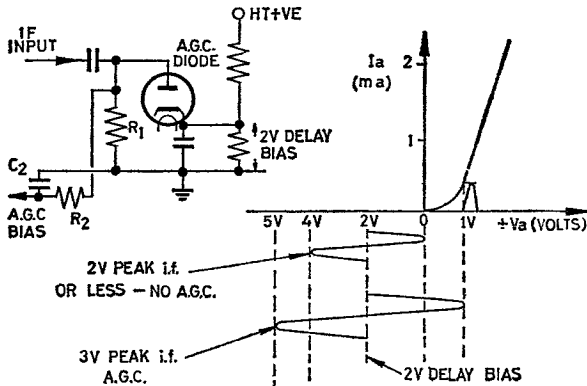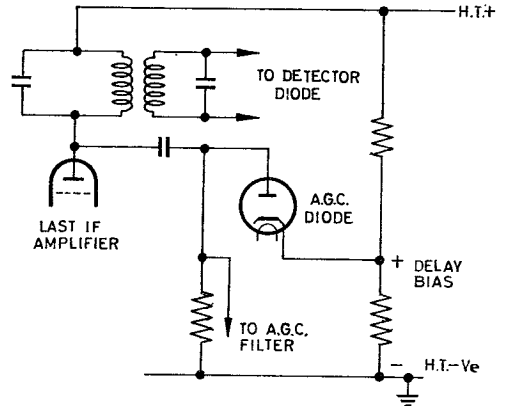

FIG. 8. EFFECT OF DELAY BIAS

FIG. 9. ALTERNATIVE METHOD OF FEEDING AGC DIODE

Since the primary circuit has a flatter response than the secondary the a.g.c. voltage is derived from a wider band of frequencies. In this case the response of the a.g.c. circuit is smoother.

AGC Time Constant

Although the a.g.c. bias must be proportional to signal amplitude it must not follow the modulation. If it did the effect of a.g.c. would be to reduce the effective depth of modulation and so reduce the signal-to-noise ratio. To ensure that the a.g.c. voltage does not fluctuate at a.f. the time constant of the a.g.c. filter circuit (C_2R_2) is made long with respect to the period of the lowest a.f. For example, assuming 300 c/s as the lowest frequency, C_2R_2 is made at least $\frac{5}{300}$ seconds.

The time constant of the a.g.c. filter must not be made too long or the a.g.c. voltage will not follow the variations in signal strength due to fading. A CR value of between 0·1 and 0·2 seconds is usual. Thus typical values of C_2 and R_2 could be 0·1μF and 1 megohm.

Amplified AGC

The second improvement on simple a.g.c. is to increase the controlling effect of the system by some form of amplification. If this is done the a.g.c. bias is altered by a given amount for a smaller change in signal voltage.

The two practical methods of doing this are amplification before or after rectification by the a.g.c. diode. Fig. 10 shows block diagrams of the two arrangements.

In the first arrangement illustrated, the a.g.c. rectifier is fed from an additional i.f. amplifier stage which is not itself controlled by a.g.c. In the second arrangement the rectified a.g.c. bias is amplified before it is fed to the controlled stages. This latter method is more common since there are no screening or alignment problems in a d.c. amplifier.

Fig. 11 shows the circuit of a typical delayed amplified a.g.c. system using a double-diode-triode valve as a d.c. amplifier.

In this circuit a 1 kilohm resistor is connected between h.t. negative and earth. The cathode current of all the other valves (here taken as 50 mA) flows through this resistor and produces a voltage across it of 50 volts, the polarity being as shown.
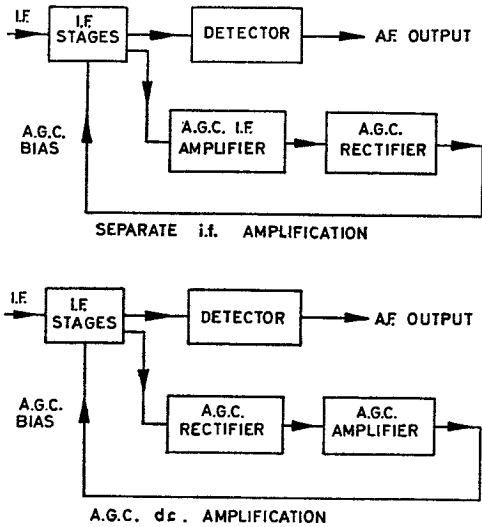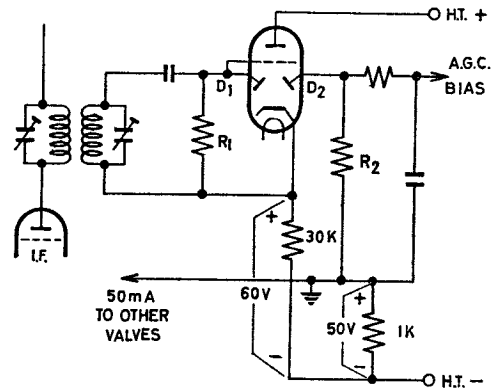

FIG. 10. AMPLIFIED AGC

FIG. 11. DELAYED AMPLIFIED AGC

With no signal voltage across the secondary of the last i.f. transformer the triode portion of the valve conducts 2 mA. This current flows through the 30 kilohm resistor and produces a voltage of 60 volts. Thus, under no-signal conditions, the anode of the a.g.c. diode D_2 is 10 volts negative with respect to its cathode and will not conduct. This -10 volts is the a.g.c. delay voltage.

On receipt of a signal the diode D_1 conducts and produces a negative bias across R_1. This is applied to the grid of the triode and reduces the current passing through the 30 kilohm resistor to below 2 mA. Thus the voltage across the 30 kilohm resistor falls and when it falls below 50 volts the a.g.c. diode conducts producing an a.g.c. voltage across R_2. This voltage is then applied to the controlled valves.

Assume that the triode has a mutual conductance of 1 mA/volt and that the signal voltage rises by 1·5 volts. Then the triode current falls by 1·5 mA, from 2 mA to 0·5 mA. This 0·5 mA produces 15 volts across the 30 kilohm resistor and the anode of D_2 becomes 35 volts positive with respect to its cathode. This is the amplified a.g.c. voltage, a change in signal level of 1·5 volts having produced a change in a.g.c. bias of 35 volts.

Post-detector AGC

It has been explained that to produce a change in a.g.c. bias the input signal must vary. This means that even when amplified a.g.c. is used some variation in the output must be expected. In order to compensate for the small variations in i.f. output necessary to produce a.g.c., the a.f. stage is sometimes provided with a small a.g.c. voltage. This is called *post-detector* or *compensated* a.g.c. and involves the use of a vari-mu a.f. valve.

112

Comparison of AGC Systems

The main features of the different a.g.c. systems are summarized in Table 1.

System	Advantages	Disadvantages
Simple	Minimum circuitry required. Can use detector output.	Affects small signals.
Delayed	Does not affect small inputs.	Limited control.
Amplified	Provides large control.	Requires an additional i.f. stage or a d.c. amplifier.

TABLE I.   COMPARISON OF AGC SYSTEMS

Output Stage

The purpose of the output stage of a receiver is to supply the amplified signal with the necessary power to operate the headset or the loudspeaker.

The theory of a.f. amplifiers has been covered in Part 1: the main points concerning a.f. amplifiers used in communication receivers are summarized as follows:—

a.   The frequency range that the amplifier covers is from 300 to 3,000 c/s.

b.   The load is usually coupled into the anode circuit by means of a transformer.   This ensures that no d.c. power is developed in the load and also provides a convenient method of matching.

c.   AF amplifiers are not normally subject to feedback through valve interelectrode capacitances so triodes, beam power tetrodes and power pentodes can be used.

d.   The output power required to work the headphones is normally less than half a watt.

e.   The usual circuit arrangement is one a.f. amplifier followed by a power output valve, both working in class A.

A typical output section of a communication receiver is shown in Fig. 12.   The input is fed to V_1 the first a.f. amplifier: this stage has a resistive anode load and the 1 kilohm cathode
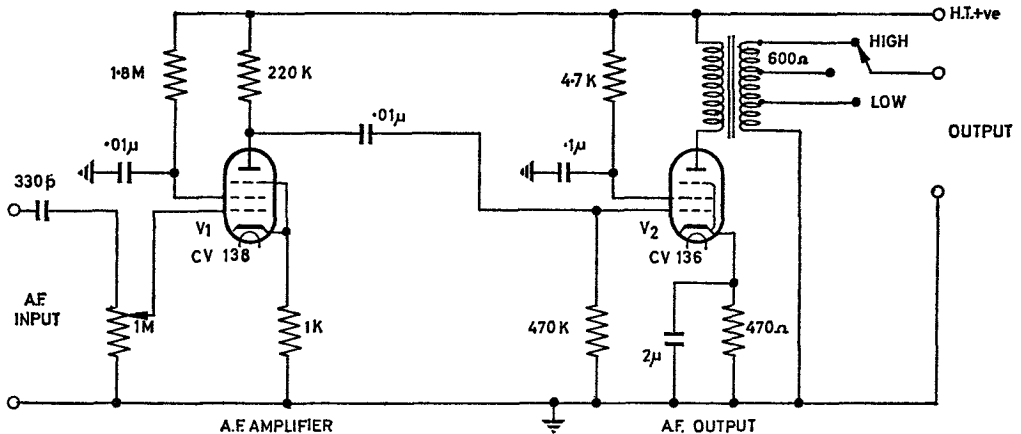

FIG. 12.   TYPICAL AF STAGES

113

bias resistor is undecoupled thus providing negative feedback. For c.w. and m.c.w. reception an a.f. choke could be used in place of the anode load resistor and in some receivers this is made to resonate at about 1,000 c/s so giving a more selective output.

The anode load of V_2 is a transformer with three output taps on the secondary. These provide matching to a high or low impedance headset and to remote lines, a typical value of which would be 600 ohms.

AF Mixing

In some airborne communication receivers the output section is required to perform several functions. Apart from amplifying the a.f. from the detector stage, the a.f. section also acts as a modulator for the transmitter, and as an intercommunication amplifier so that crew members can talk to each other.

To enable this to be done, an a.f. mixing stage is used. A typical mixing and a.f. amplifying stage is shown in Fig. 13. The input from the detector stage of the receiver is developed across
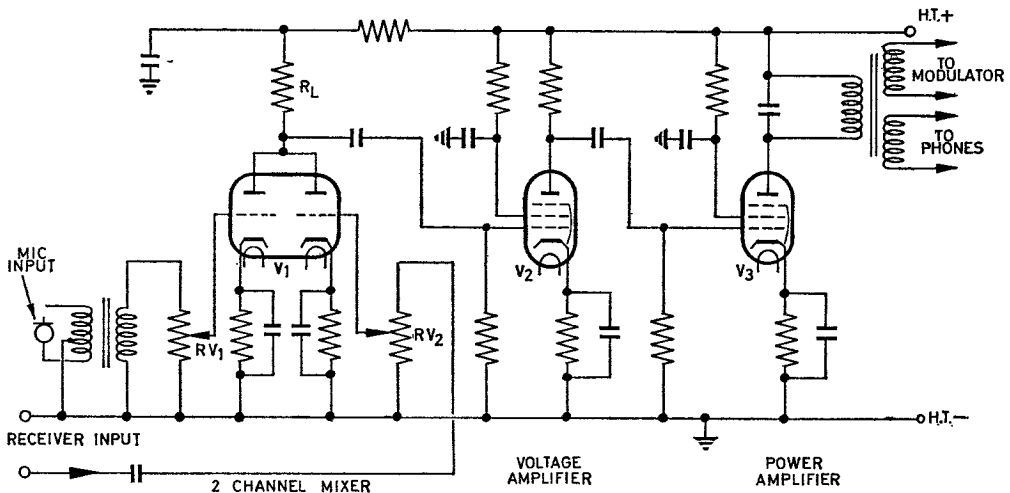
FIG. 13. THREE-PURPOSE AF AMPLIFIER

RV_2 and applied to the grid of one section of the double-triode audio mixer valve V_1. The microphone input, necessary for intercom or transmitter r.t. modulation is fed to the grid of the other triode section via RV_1. Thus the input from either of these two sources is developed across the common load resistor R_L, and fed to the a.f. voltage amplifier V_2 and hence to the modulator and headphones. Suitable relay switching provides the facility required. By adjusting RV_1 and RV_2 similar amplitude outputs are obtained from the microphone and receiver inputs.

Transistor Output Stages

A typical circuit of the a.f. output stages of a receiver using transistors instead of valves is shown in Fig. 14.

The a.f. input is developed across R_1 and fed to the base of the a.f. voltage amplifying stage TR_1. The coupling capacitor C_1 is of large value because of the low input impedance of

FIG. 14. TYPICAL TRANSISTOR OUTPUT STAGES

the transistor. Stabilizing bias is provided by $R_2R_3R_4$, and negative feedback to reduce distortion is obtained from the secondary of the output transformer T_2.
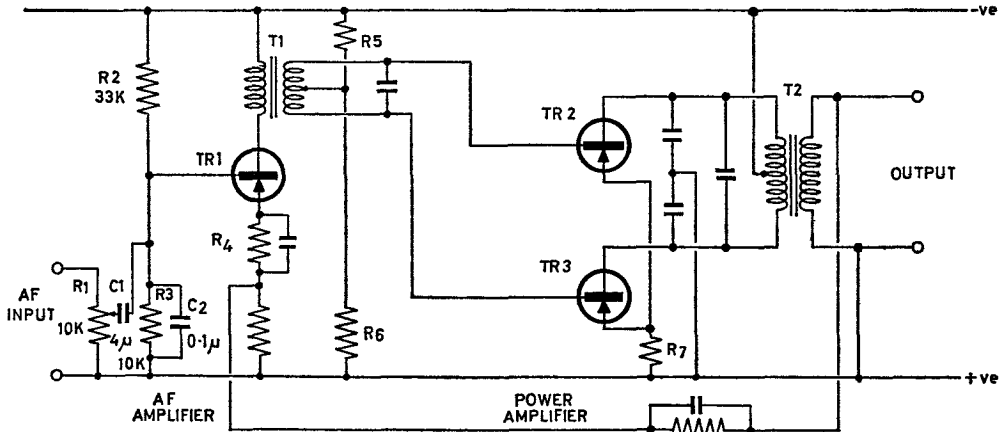
The amplified a.f. current develops a voltage across T_1 secondary to drive the push-pull connected power transistors TR_2 and TR_3. These are usually matched transistors and the bias resistors R_5, R_6 and R_7 are undecoupled to provide negative feedback and so improve the performance of the stage. Normal push-pull action results in considerable power being developed to operate the headset or loudspeaker.

Tuning Indication

The most obvious method of tuning is to turn the tuning control for maximum output in the headphones. However, when a.g.c. is applied it tends to counteract any change in signal strength, including that due to tuning. Thus a.g.c. makes tuning much less decisive and when tuning for maximum sound output the maximum is difficult to judge.

Since the a.g.c. bias is proportional to signal strength it can be used to indicate the correct tuning point.

Tuning Meter

The basic form of visual tuning indication uses a meter to read current changes in a valve controlled by a.g.c. Fig. 15 shows a tuning meter connected in the cathode circuit of an i.f. amplifier valve.
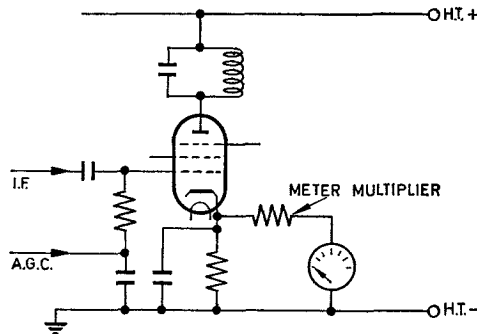

FIG. 15. THE BASIC TUNING METER

115

Since the cathode bias is directly proportional to valve current and this in turn depends on a.g.c. bias, the voltage across the cathode bias resistor will provide a simple tuning indication. Maximum signal will cause maximum a.g.c. bias, minimum valve current and so minimum cathode bias. Therefore a maximum dip in the tuning meter indicates correct tuning.

Limitations of the Tuning Meter

The basic tuning meter just described works quite effectively if the signal is strong enough to cause an appreciable change in valve current. On the other hand, if the signal is very strong the a.g.c. voltage produced will bias the valve so far back on its vari-mu characteristic that very little change of current occurs in spite of considerable change of bias voltage.

The obvious alternative is to measure the a.g.c. bias itself. This is very difficult because of the high output impedance of the a.g.c. diode. Even a meter with a very high ohms-per-volt rating would tend to shunt the source and give false indications.

Magic-eye Tuning Indicator

A device which draws no current from the a.g.c. diode and one commonly used as a visual tuning indicator in receivers, is the magic-eye indicator. Its construction is shown in Fig. 16. It comprises a triode amplifier section and a visual indicator section.

The cathode of the triode section extends into the indicator section which consists of a conical target anode with a coating of fluorescent material. A thin deflector rod or 'shadow-former' projects through a hole in the target anode and is connected to the anode of the triode section.

Fig. 17 shows a typical circuit for a magic-eye indicator operated from a receiver a.g.c. line.

The target anode is connected directly to the h.t. positive line. The shadow-former and the triode anode are connected to h.t. positive via a 1 megohm resistor. The a.g.c. line is applied to the grid of the triode.

Under 'no signal' conditions, no a.g.c. bias is developed and the triode grid is at the same potential as the cathode. Maximum current flows and a large voltage drop occurs across the 1 megohm resistor. The potential of the shadow-former is therefore less than that of the target anode. Thus electrons from the cathode to the target anode are deflected by the shadow-former and a wide shadow (Fig. 18a) is formed on the fluorescent surface of the target anode.
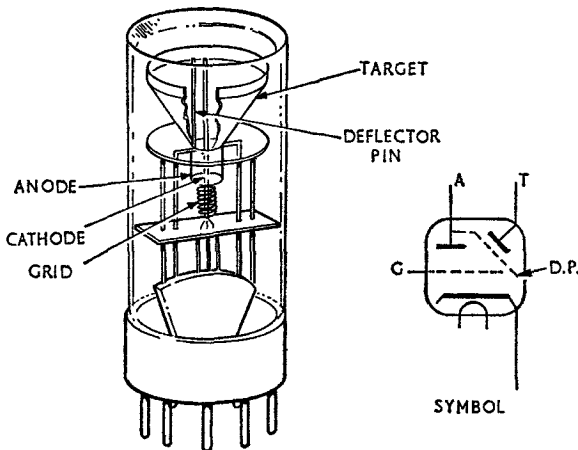

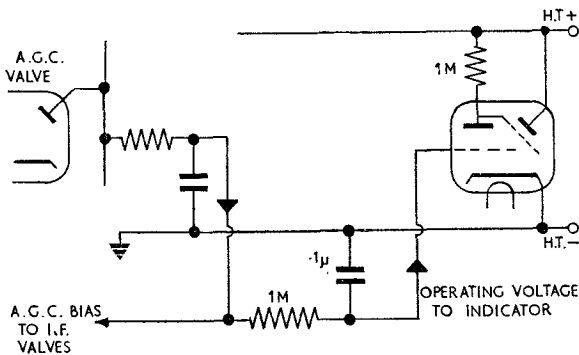FIG. 16. MAGIC EYE TUNING INDICATOR

FIG. 17. TYPICAL MAGIC EYE CIRCUIT

As a signal is tuned in the a.g.c. negative bias is increased and the triode current falls. The decrease in voltage drop across R_L causes the potential of the shadow-former to rise and the shadow narrows (Fig. 18b). When the shadow is a minimum (Fig. 18c), the receiver is correctly tuned.
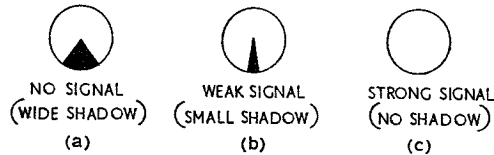
NO SIGNAL
(WIDE SHADOW)
(a)

WEAK SIGNAL
(SMALL SHADOW)
(b)

STRONG SIGNAL
(NO SHADOW)
(c)

FIG. 18. MAGIC EYE INDICATIONS

Noise Limiting

Although with a strong signal a.g.c. can greatly reduce background noise, this is only because the noise level is almost constant. The problem of interference caused by sudden noises still remains, and it is to eliminate this form of interference that noise limiting circuits have been developed.

Fig. 19 shows the waveform of an amplitude modulated signal containing short-duration noise.

The method of eliminating this type of interference is simply to cut off the noise peaks during detection.

Basic Limiting Circuits

There are two basic limiting circuits: one employs a diode connected in series with the signal path: the other uses a diode connected in parallel. These two circuits with their input and output waveforms are shown in Fig. 20. The main difference between these and series or shunt detector circuits is that the limiter diode must be biased to discriminate between the signal and noise.
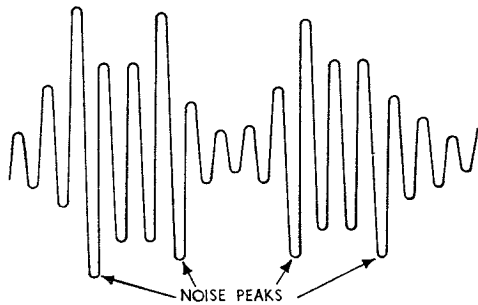
NOISE PEAKS

FIG. 19. NOISE ON AN AM SIGNAL
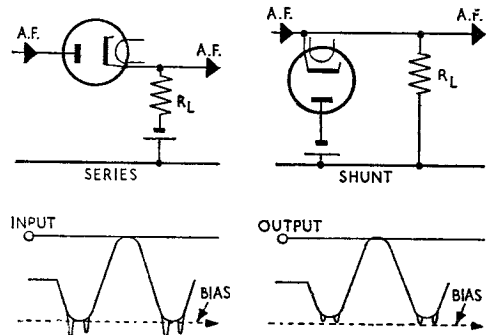
FIG. 20. BASIC LIMITING CIRCUITS AND WAVEFORMS

In the series limiting circuit the diode is normally conducting because the cathode is biased negatively. When the amplitude of the negative-going noise peaks exceeds the bias value, the diode cuts off and noise voltages of greater amplitude than the peak a.f. signal are not developed across R_L.

In the shunt circuit the diode is normally cut off with its anode biased negatively. When a noise voltage on the cathode exceeds the bias voltage the diode conducts, shorting the load resistor R_L and preventing the noise peaks appearing in the output. Thus in either case, if the input signal exceeds the bias, there is no output.

117

Practical Noise Limiter

The simple circuits of Fig. 20 are only effective for one particular value of signal amplitude, i.e. signals equal in amplitude to the bias level. If the receiver were tuned to a weaker signal, noise limiting would be reduced. With a stronger input signal both noise and signal would be limited and distortion introduced. In fact these simple circuits are effective for one amplitude signal only.

In order to provide a limiter which will limit noise on any amplitude of carrier without distorting the a.f., the bias voltage must be proportional to the mean carrier level. Such a bias voltage is available in the d.c. output of a normal detector.

A typical automatic noise limiting circuit is shown in Fig. 21. The anode of the limiting diode is positive with respect to its cathode by an amount proportional to the amplitude of the signal. Therefore V_2 conducts and the a.f. output is taken across R_3.
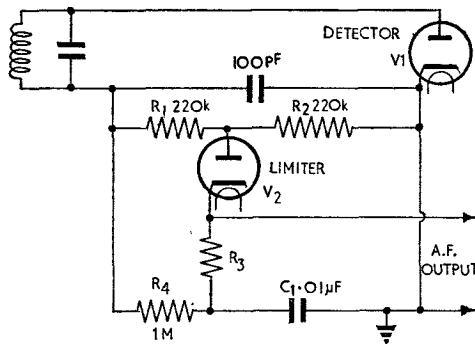

FIG. 21. AUTOMATIC NOISE LIMITING CIRCUIT

A large, short duration noise pulse will make the anode of the limiter diode less positive, but because C_1R_4 is a long time constant (10 milliseconds) the cathode potential of the limiter does not change. If the pulse exceeds 100 per cent modulation V_2 will cut off. Different values of R_1 and R_2 will alter the bias level on V_2 and so change the modulation depth at which limiting occurs.

Receiver Muting

With no input signal there is no a.g.c. bias produced and so the receiver gain is a maximum. This means that the receiver noise gets full amplification and a fairly loud continuous background of noise is heard in the headset. It is necessary to quieten this no-signal receiver noise. This facility is called *muting* and results in the output to the headset being completely cut off unless the signal amplitude exceeds a certain value. This can be done by a relay shorting the receiver output under no-signal conditions. Alternatively, it can be done electronically by biasing the detector valve such that it is cut off for all signals below a certain 'threshold' level.

Muting circuits are sometimes called squelch or interchannel noise suppression circuits.

Conclusion

This chapter completes the circuits which form a typical communication receiver. The function of each circuit and the important points concerning each circuit have been dealt with. In the next chapter details of measurements necessary to gauge the performance of a receiver will be considered.

CHAPTER 3

NOISE AND RECEIVER MEASUREMENTS

Introduction

When a communication receiver is initially brought into use it should meet all the specifications of the designer. After it has been in use for some time however, its performance may have deteriorated. To detect any falling-off in performance the receiver is given certain periodic tests: if it is found to be failing in any of these tests the necessary adjustments or component replacements are made to bring it back to specification.

The purpose of this chapter is to discuss in a general way the need for and methods of making these various performance tests and adjustments. More precise details of the tests which are required to be carried out on a particular receiver will be found in the servicing A.P. relating to that receiver.

The factors which govern the effectiveness of a receiver are sensitivity, selectivity, fidelity and signal-to-noise ratio. The receiver to be tested must be serviceable in that there should be no faulty components or wiring. The tests must be carried out under certain standard conditions which are:—

a. Specified screened input,

b. Specified power supplies,

c. Specified adjustment sequence,

d. Specified operating temperature.

These requirements are illustrated in Fig. 1.
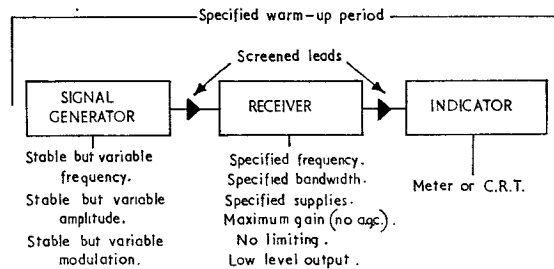
FIG. I. RECEIVER TEST REQUIREMENTS

Receiver Noise

In order to appreciate the reasons for some of the receiver tests it is necessary to know what limits the receiver performance. Noise is an important limiting factor and so the various types and sources of noise will now be considered.

The noise in the input to a receiver is the background from which the receiver selects the wanted signal. There are two types of noise present in the output of a receiver. Firstly, the noise produced *outside* the receiver and introduced into it via the aerial. Secondly, there is noise produced *within* the receiver itself. The sources from which these two types of noise originate are shown in Fig. 2.

FIG. 2.   RECEIVER NOISE SOURCES

Of the external noise, electrical interference can be reduced by suppressing or screening the source of interference.   The internal noise depends on the design of the receiver, and has a considerable effect on the output signal-to-noise ratio.   It is the internal noise level which fundamentally limits the range of a communication receiver.
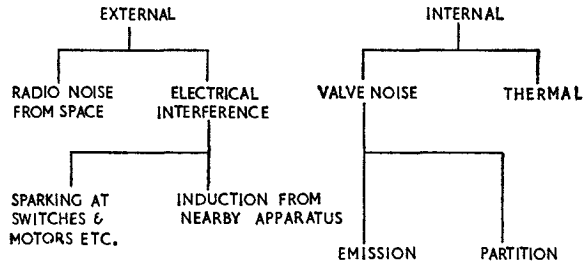
When carrying out performance tests on a receiver the lead connecting the source of the input signal (usually a signal generator) to the receiver input circuit is normally screened.   Thus external noise is eliminated and the term receiver noise means only that noise due to internal sources.   The types of receiver noise are considered in detail in the following paragraphs.

Thermal Noise

In any conductor at room temperature there is a random movement of electrons.   This movement of electrons constitutes a current and therefore a random noise voltage is produced across the conductor.   The magnitude of this voltage is dependent on three factors: the temperature of the conductor, the resistance of the conductor and the bandwidth of the random variations.
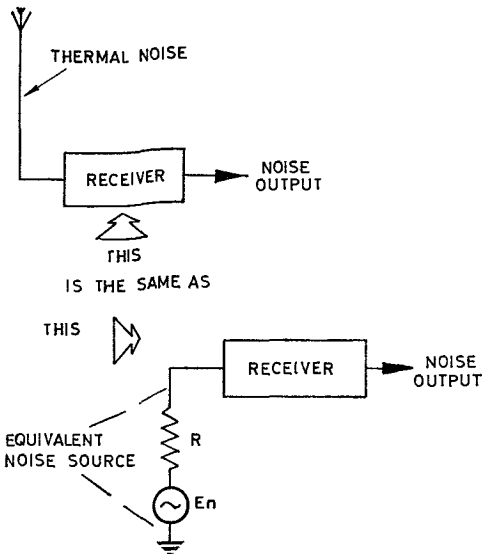
FIG. 3.   EQUIVALENT NOISE SOURCE

The equivalent thermal noise voltage (e_n) is given by the formula:

$$e_n = \sqrt{4KTBR} \quad \text{volts,}$$

where K is a constant, called Boltzmann's constant,

T is temperature in degrees Kelvin,

B is bandwidth in c/s,

R is resistance in ohms.

Thus the equivalent noise source of a receiver aerial, due solely to thermal noise in the aerial wire, can be considered as consisting of a noise generator giving e_n volts, in series with a resistance R equal to the aerial resistance.   This is shown in Fig. 3.

Maximum power is transferred when the source matches the load.   So both signal and noise develop maximum power

120

in the receiver when aerial resistance and receiver input resistance are equal. The noise voltage input is then:

$$\frac{e_n}{2} = \sqrt{KTBR} \text{ volts.}$$

Emission or Shot Noise

The number of electrons leaving the cathode of a valve in which all the electrodes have a steady potential is not constant. The number varies in a random manner and this variation of current flowing through a load resistor sets up a varying voltage across the resistor. This constitutes *emission noise*.

A valve operated under space charge conditions tends to introduce less emission noise since the space charge smooths out the random variations. The higher the valve current the
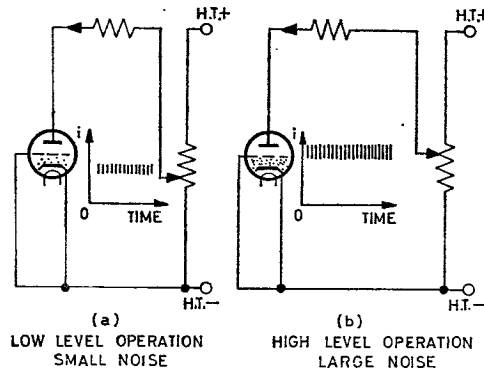

FIG. 4. EFFECT OF CURRENT AMPLITUDE ON EMISSION NOISE

greater is the noise produced. This is illustrated in Fig. 4 which also shows that noise occurs with or without an input to the valve.

In effect the valve current is a direct current whose amplitude fluctuates in a random manner, i.e. it is modulated by noise.

Partition Noise

In a valve with several positive electrodes all maintained at constant potentials, the number of electrons collected by each electrode is not constant. Thus if the screen current varies, the anode current will vary and random noise voltages will be produced across the load resistor.

The more electrodes there are at a positive potential the more variation there is in the division of the original electron stream and thus the greater will be the random variations in the anode current. A pentode is about five times as noisy and a heptode about twenty-five times as noisy as a triode for these reasons.

Transistor Noise

Emission noise and thermal noise are both present in a transistor. The noise introduced by a transistor increases as the current through the transistor increases. This is because the junction temperature rises. Also, a transistor is more noisy at low frequencies than at high frequencies and so noise is greatest in high-gain low frequency amplifiers.

121

Equivalent Noise Resistance

In order to make a quick comparison of the noise generated by different valves it is usual to refer to the *equivalent noise resistance* of a valve. This is the resistance which would produce the same thermal noise as the total noise produced by a valve. Fig. 5 shows three valves and their equivalent noise resistances.
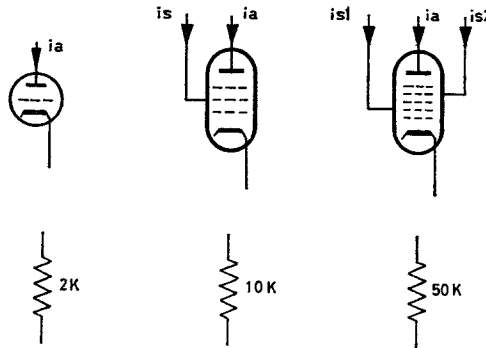

FIG. 5. EQUIVALENT NOISE RESISTANCE

Thus although the pentode is capable of about five times the amplification of a triode, it is also five times as noisy. Because of this, grounded grid triodes are frequently used in the r.f. amplifier stage of a receiver.

Signal-to-noise Ratio

The term 'signal-to-noise ratio' of a receiver is generally taken to mean the ratio of the output signal voltage to the output noise voltage. With an ideal receiver, i.e. one that generates no internal noise, the signal-to-noise ratio at the output of the receiver would be the same as that at the input. In a practical receiver, each stage adds to the thermal noise present in the aerial and this is amplified and added to by succeeding stages. Therefore the *output* signal-to-noise ratio will be much *lower* than the *input* signal-to-noise ratio.

For example, if the input signal is 10 μV and the input noise 1 μV the input signal-to-noise ratio is 10:1, or 20 db. The output signal-to-noise ratio will be less than this. Fig. 6 shows typical signal-to-noise ratios in db at the various stages of a communication receiver.

In this case the output signal-to-noise ratio is 12 db or 4:1. Thus if the amplification of the signal and noise is the same, the receiver increases the input noise by $\frac{10}{4}$ or 2·5 times.

The noise generated by the receiver can therefore be represented by an effective noise source of 1·5 μV in series with the signal and the input noise applied to the input circuit of an *ideal* receiver. This equivalent circuit is shown in Fig. 7.

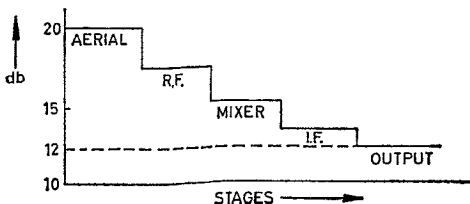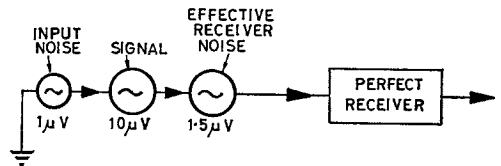FIG. 6. RECEIVER SIGNAL-TO-NOISE
RATIOS

FIG. 7 EQUIVALENT NOISE CIRCUIT OF
A PRACTICAL RECEIVER

122

Since the ideal receiver produces no internal noise the input and output signal-to-noise ratios are the same, i.e. 4:1.

Noise Factor (NF)

The noise factor of a receiver gives a measurement of the noise introduced by the receiver itself. It is given by,

$$\text{Noise factor} = \frac{\text{S/N ratio at output of ideal receiver}}{\text{S/N ratio at output of actual receiver}}.$$

This is the same as: $\dfrac{\text{Input signal-to-noise ratio}}{\text{Output signal-to-noise ratio}}$ of a practical receiver.

For an ideal receiver the noise factor is of course unity. For all practical receivers it must be greater than unity. For the receiver of Fig. 6,

$$NF = \frac{10:1}{4:1} = 2 \cdot 5.$$

Expressed in terms of db the noise factor is simply the difference between the input and output signal-to-noise ratios measured in db. For the receiver of Fig. 6,

$$NF = 20 \text{ db} - 12 \text{ db} = 8 \text{ db}.$$

Measurement of Receiver Sensitivity

The sensitivity of a receiver is defined as the minimum voltage required at the aerial to produce a stated a.f. power output with a given signal-to-noise ratio. For example the sensitivity of the receiver considered in Fig. 6 might be stated as a 10 μV input required to produce 150 mW output with a signal-to-noise ratio of 12 db.

To measure the sensitivity of a receiver the equipment, connected as shown in Fig. 8, is required. The signal generator with an output modulated at 400 c/s to a depth of 30% is
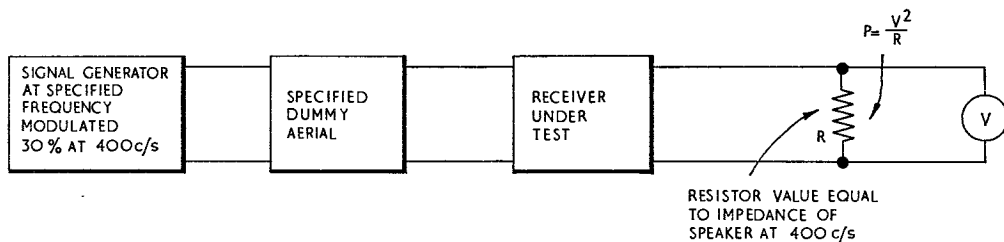
FIG. 8. RECEIVER SENSITIVITY TEST

connected to the receiver through a standard dummy aerial. A resistor of value equal to the loudspeaker or headphone impedance at 400 c/s is connected across the receiver output. A voltmeter is connected in parallel with this resistor.

All the equipment should be allowed to warm up to normal operating temperature and with the signal generator disconnected, the r.f. and i.f. gain controls are turned fully up. The a.g.c. and b.f.o. are switched off. The a.f. gain control is then increased to give a noise voltage across the output resistor of the value given in the receiver servicing A.P.

The signal generator is then tuned to the specified frequency and connected via the dummy aerial to the receiver. The input voltage from the generator is then increased until the specified

123

signal-plus-noise voltage is read in the output meter. The receiver sensitivity in terms of input signal voltage is then read from the voltage (or db) calibration on the signal generator.

Receiver Selectivity Test

Receiver selectivity is normally stated as the i.f. bandwidth in cycles per second for a given number of dbs below maximum output. For example the selectivity of a receiver could be stated as 6,000 c/s at 6 db down.

To measure the selectivity of a receiver the signal generator and receiver are connected as for the sensitivity test. With the receiver a.g.c. off, the signal generator is set to the specified receiver frequency (f_o) and the output of the signal generator is adjusted to give the specified receiver output. The signal generator frequency is then varied either side of the receiver frequency and the attenuator output is adjusted to give the standard output. The db attenuation on the signal generator output at various frequencies from f_o is thus obtained to give the selectivity. A graph as shown in Fig. 9 can then be drawn.
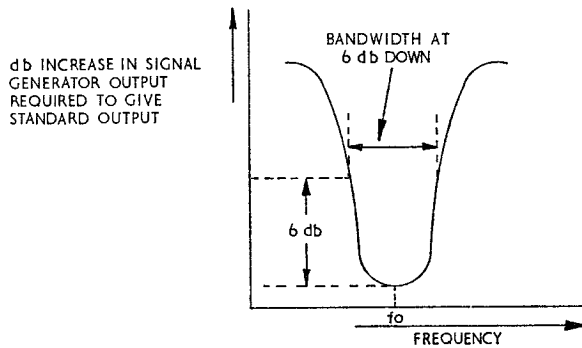

FIG. 9. RECEIVER SELECTIVITY CURVE

Image Channel Ratio Test

The ratio $\dfrac{\text{Response at second channel frequency}}{\text{Response at wanted signal frequency}}$ is called the *image channel ratio* and is usually expressed in dbs.

This test is made by connecting the signal generator to the receiver in the same manner as for the previous tests. By tuning the signal generator first to the wanted signal and then to the unwanted second channel signal and adjusting the db attenuation of the signal generator to give the standard output in each case, the necessary increase in input at the second channel frequency will give the db ratio.

Non-linear Distortion Test

The non-linear distortion ratio is given by:

$$\frac{\text{RMS value of harmonics in output}}{\text{RMS value of fundamental in output}}.$$

To carry out the test, the signal generator and receiver are connected as previously and a filter network is inserted in the receiver output to eliminate the fundamental. The r.m.s. value of the harmonics can then be measured. Then a filter to reject harmonics is inserted and the fundamental filter removed. The r.m.s. value of the fundamental is then measured and the non-linear distortion ratio found.

Amplitude-Frequency Response (Fidelity) Test

A fidelity test can be carried out by observing the variation in a.f. output when a r.f. carrier modulated to a constant depth by a variable a.f. is applied to the receiver input.

A signal generator modulated 30% at an a.f. which can be varied between the limits laid down in the receiver servicing A.P., is connected to the receiver input. The signal carrier level is kept constant while the modulating audio frequency is varied. The a.f. output of the receiver is then measured and a response curve plotted to show the variation in output over the a.f. response band.

Receiver Alignment

In a superhet receiver the tuned circuits are normally provided with trimming and padding pre-set components. The purpose of these in ganging and tracking have already been discussed, and they are set to the correct values by the manufacturers. However, as the components age, or due to major modification to the receiver, the pre-set components may have to be adjusted. The process of adjustment is called *receiver alignment*.

Alignment Procedure

There are two basic methods of aligning a receiver, depending on the test equipment available. In both cases the sequence of adjustment is the same: the circuits are aligned in turn, from the final tuned circuit to the first tuned circuit of the receiver, i.e. from the last i.f. to the first r.f. tuned circuit.

Alignment Using an Output Meter

Fig. 10 shows the basic arrangement and adjustment sequence for aligning a receiver when using an output meter and signal generator.
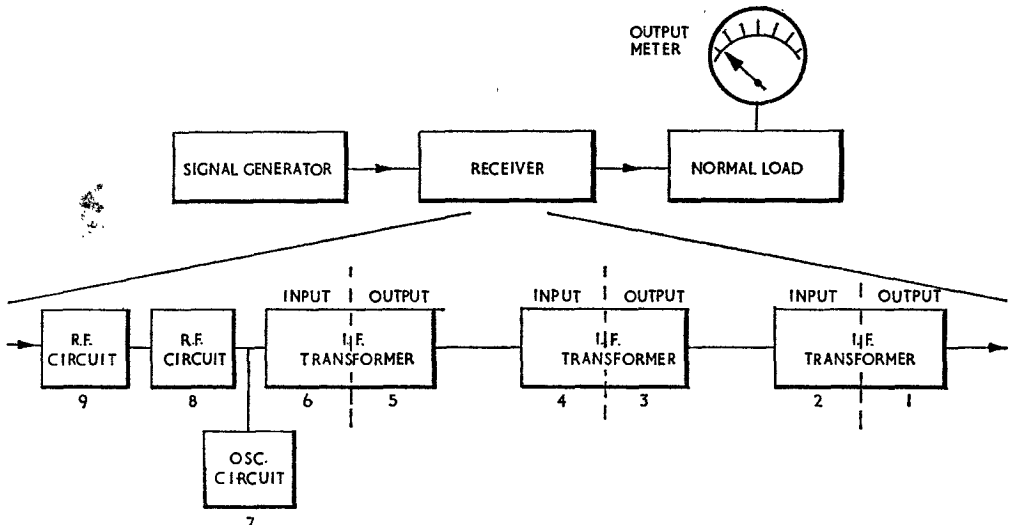
FIG. 10. RECEIVER ALIGNMENT USING AN OUTPUT METER

Notice that the diagram shows only the tuning circuits and not the valve stages. Thus the mixer stage is omitted since the mixer output circuit is the input circuit of the first i.f. transformer and the mixer input tuned circuit is the last r.f. circuit.

125

In a normal communication receiver there will be a number of frequency bands and each band will have its own oscillator tuned circuits and the tuned circuits associated with the r.f. amplifying stages. For example in a six-band receiver consisting of one r.f. stage, a mixer and three i.f. amplifiers there could be 18 r.f. tuned circuits and 6 i.f. tuned circuits which would have to be adjusted.

Alignment Using a Wobbulator and CRO

A very convenient method of aligning a receiver is to display the receiver response curve on an oscilloscope and adjust the receiver pre-set components for optimum response. This method is particularly useful when a special form of response is required.

The device which replaces the normal signal generator in this test is called a *frequency sweep generator* or *wobbulator*. This equipment provides an input test signal whose frequency varies (or 'wobbles') about a centre frequency. The rate of wobble is controlled by the time-base voltage of the c.r.o.

The wobbulator consists of a r.f. oscillator whose frequency is controlled by a reactance valve. The time-base voltage is applied to the reactance valve which then varies the output frequency above and below the centre frequency, in synchronism with the c.r.o. timebase.

Fig. 11 shows the arrangement of test equipment required to align a receiver. The detected output from the receiver is fed to the Y deflection system of the c.r.t. Normally the a.f. circuits cannot be adjusted, so it is not necessary to check their response during alignment.
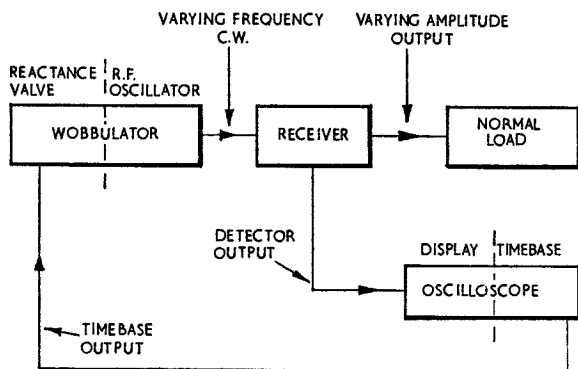

FIG. 11 RECEIVER ALIGNMENT USING A WOBBULATOR AND CRO

Since the time-base and wobbulator are synchronised, the c.r.t. spot will complete one sweep while the signal generator completes one sweep. At the 'on tune' frequency the receiver output should be a maximum: at frequencies above and below this the output will fall according to the response of the receiver.

Fig. 12 shows the equipment suitably connected for an alignment test on a broad-band receiver. The response curve on the c.r.t. would be typical for greater-than-critical coupling of the i.f. tuned circuits.

At the start of the time-base (point X in Fig. 12) the input frequency from the wobbulator is outside the response of the receiver and the gain is a minimum. As the time-base voltage increases so does the input frequency and therefore the gain of the receiver increases. At the centre frequency (point Y), receiver response is in the 'saddle'. At maximum time-base voltage the input frequency has increased to a maximum and is again outside the receiver response.

SET TO CHOSEN
SWEEP FREQUENCY

SET TO SAME R.F.
FREQUENCY

WOBBULATOR

RECEIVER

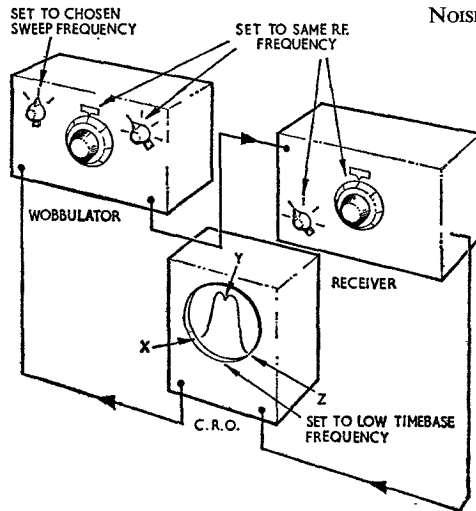Y

X

Z

C.R.O.

SET TO LOW TIMEBASE
FREQUENCY

FIG. 12. TEST EQUIPMENT CONNECTIONS

The whole cycle then repeats and provided it occurs at least 15 times per second the trace displays a steady picture of the receiver response curve. Fig. 13 shows the effect on the display of a change in input frequency variations and also a correctly and incorrectly aligned display in each case.

To enable measurements to be taken, a squared graticule is sometimes fitted to the c.r.t. screen. The bandwidth at the half power points can then be found by taking measurements

TIMEBASE

INPUT f f_0

DISPLAY

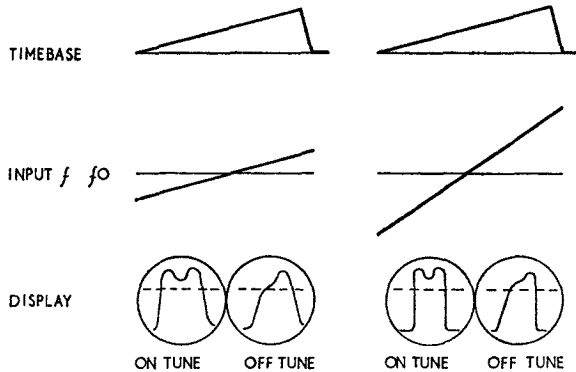ON TUNE   OFF TUNE   ON TUNE   OFF TUNE

FIG. 13. EFFECT OF SWEEP AND ALIGNMENT CHANGES

on the graticule. For example if a 10 kc/s frequency sweep covers 15 cm on the graticule, then 5 cm at half power points would indicate a bandwidth of 3·3 kc/s.

An alternative method of calibration is to mix frequency marker pips with the wobbulator output. The bandwidth can then be read by reference to the markers.

Conclusion

The receiver tests and adjustments mentioned in this chapter are typical of those carried out in practice. Whenever such a servicing has to be carried out on a communication receiver the A.P. referring to the servicing of that type of receiver should be read and details necessary to carry out the servicing obtained.

CHAPTER 4

THE SINGLE SIDEBAND RECEIVER

Introduction

An outline of the s.s.b. system of communication was given in Section 2, Chapter 8. It was pointed out that a normal amplitude modulated signal consisting of carrier plus upper and lower sidebands was really quite an inefficient means of communication. The intelligence is contained in one sideband, yet both sidebands and the carrier are generated, amplified and transmitted. In addition to the waste of power at the transmitter and the unnecessarily wide bandwidth occupied by the double sideband signal, the system also introduces several disadvantages at the receiver.

The wide bandwidth, apart from causing frequency crowding, means that the receiver has to accept and amplify much more noise than is necessary. A s.s.b. receiver with its narrow bandwidth has a much better signal/noise ratio. The effect of selective fading, inherent in any multi-frequency h.f. signal, is greatly reduced in the case of a s.s.b. signal since there is no balance between upper and lower sideband to upset.

Because of these advantages the s.s.b. system is often used for multi-channel radio teleprinter signalling on the Commonwealth Air Forces network, and in air-to-ground and air-to-air communication.

Single Sideband Systems

It has already been pointed out that the main problem associated with reception of s.s.b. signals is the re-insertion of the carrier at the correct frequency. The carrier must be present
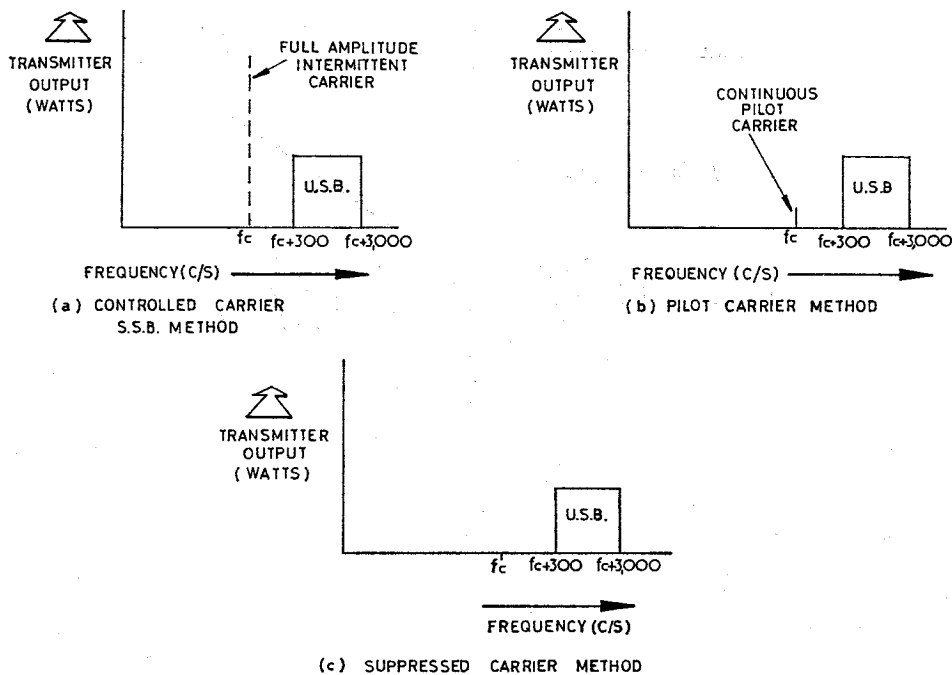

FIG. I. FORMS OF A SSB SIGNAL

at the receiver along with the sideband before detection can take place, but if its frequency is incorrect the detected signal will be distorted. The s.s.b. signal can be transmitted in one of three forms, as illustrated in Fig. 1. The first form illustrated consists of a sideband plus a 'controlled' carrier where the carrier is sent at full strength between breaks in the modulation. Fig. 1b shows the sideband plus a continuous low power or 'pilot' carrier. Fig. 1c illustrates the sideband by itself without any form of carrier and this is known as the suppressed carrier method.

These are the three forms of a true s.s.b. system. However the number of communication channels can be increased if an upper and a lower sideband, without a carrier is transmitted. Each sideband can carry a different message, thus doubling the number of communication channels. This system is known as *independent sideband* (i.s.b.) transmission.

The Triple Superhet

The equipment necessary to receive teleprinter signals must be able to operate with very weak signal inputs and under the most adverse fading conditions. Atmospheric interference and noise generated within the receiving equipment itself must be rejected or minimised to avoid mutilation of the code characters.

These requirements are largely met by using triple superhets, usually in triple-space diversity. In principle the triple superhet is merely an extension of basic frequency changing. It has three frequency changers and three local oscillators. The stages before the demodulator stage are shown in block form in Fig. 2, with the frequency at each stage, assuming a 10 Mc/s input.
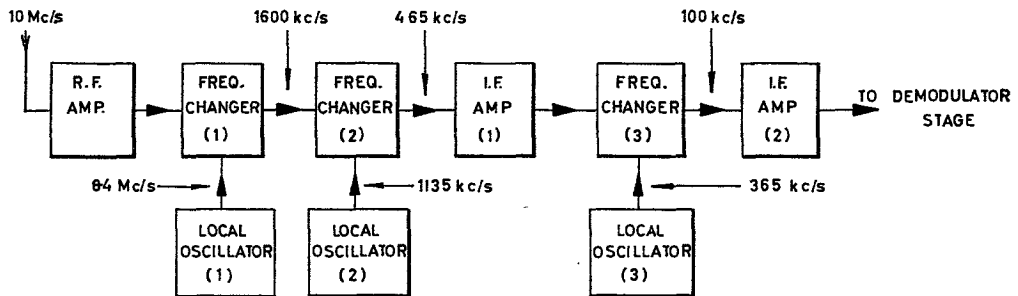

FIG. 2. THE TRIPLE SUPERHET

The incoming signal is fed to the r.f. amplifier and combined in the first frequency changer with the output from the first local oscillator. This produces the first i.f. of 1,600 kc/s. The signal is then passed to the second frequency changer, where, by combining with the 1,135 kc/s output from the second oscillator, the second i.f. of 465 kc/s is produced. This is amplified in the first i.f. amplifier and then passed to the third frequency changer. After combining with the output from the third oscillator, operating at 365 kc/s, the final i.f. of 100 kc/s is formed. This signal is amplified in the second i.f. amplifier and passed to the demodulator stage.

A big advantage in using three frequency changing stages is that a much more accurate final i.f. can be obtained than if one frequency changer were used. This is because if each oscillator is accurate to within, say, $0 \cdot 1\%$ then the first oscillator in Fig. 2 would be accurate to within $8 \cdot 4$ kc/s. The third oscillator on 365 kc/s would be within 365 c/s. A single frequency changer with the oscillator on $10 \cdot 1$ Mc/s, and with an accuracy of $0 \cdot 1\%$ ($10 \cdot 1$ kc/s) would produce an i.f. accurate to less than 10%.

The triple superhet has a high gain and high selectivity which make it most suitable for reception of the narrow-band s.s.b. signal.

Reception of Controlled or Pilot SSB Signals

When the s.s.b. signal is transmitted with a controlled carrier or with a weak pilot carrier, the composite signal is amplified in the receiver and the carrier is selected by a filter and used to produce an a.f.c. voltage with which the frequency of the second oscillator is controlled. An a.g.c. voltage is also obtained from the carrier and used to control the gain of various stages of the receiver. A block diagram of such a receiver is shown in Fig. 3.
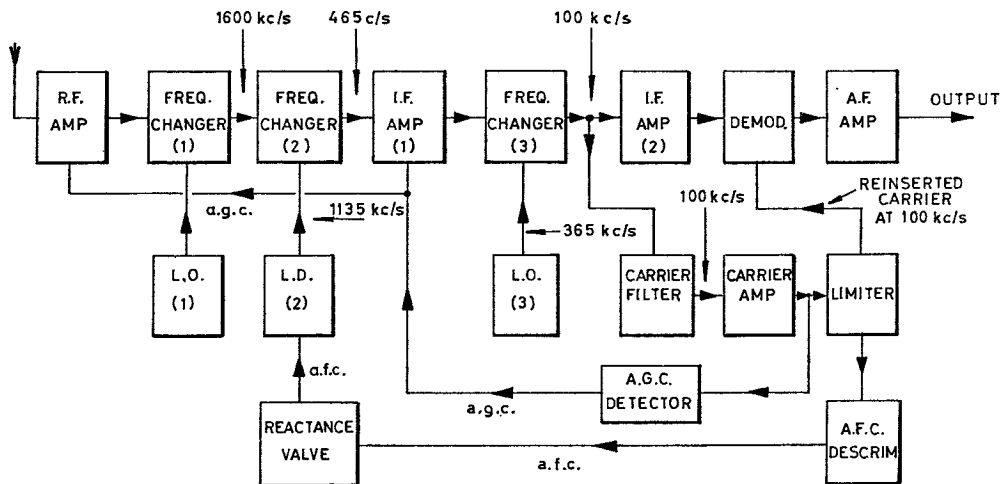
FIG. 3. SSB PILOT CARRIER RECEIVER

The s.s.b. signal, consisting of sideband and carrier, is amplified in the r.f. amplifier, changed to 1,600 kc/s and then to 465 kc/s in the first and second frequency changers. After amplification in the first i.f. amplifier, the third frequency changer reduces the signal to 100 kc/s plus the audio modulating frequencies. Then, after further amplification in the second i.f. amplifier, it is applied to the demodulator stage.

Before the s.s.b. signal can be demodulated it is necessary to provide a carrier of higher level than the sideband. This carrier must be of fairly constant amplitude and free from sidebands. Thus a portion of the signal output from the third frequency changer is passed through a filter which rejects the sideband frequencies and passes the carrier frequency only. After amplification and limiting, the carrier is fed into the demodulator stage where it combines with the sideband frequencies and enables demodulation to take place in the normal way.

After amplification in the a.f. stage, the audio output is fed to the line or teleprinter equipment.

Automatic Frequency Control

The required frequency stability is obtained by feeding part of the carrier voltage from the limiter stage to an a.f.c. discriminator circuit. The output from the discriminator will be a positive or negative voltage proportional to the *frequency* of the carrier. This voltage is applied to a reactance valve which controls the frequency of the second oscillator.

With this system the frequency of the second i.f. is held such as to maintain the carrier frequency within a few cycles of 100 kc/s thus compensating for any drift in the frequency of the first oscillator.

If the transmitter drifts in frequency, the receiver will follow this drift because the pilot carrier indicates the direction and extent of the drift. With suppressed carrier signals there is no such indication since there is no carrier, and therefore frequency drift in both the transmitter and receiver must be negligible. Hence the transmitter and receiver oscillators in a suppressed carrier system are thermostatically controlled to provide this high standard of stability.

Automatic Gain Control

To provide a.g.c., a portion of the filtered and amplified carrier is taken from the output of the carrier amplifier. This voltage is then rectified in the a.g.c. circuit and applied to the r.f. amplifier and first i.f. amplifier to provide automatic control of the gain of these stages.

Reception of Suppressed Carrier SSB Signals

The suppressed carrier s.s.b. system is the most desirable since *all* the transmitted power is contained in the intelligence-carrying sideband, thus giving an effective increase in transmitted power. There is, however, a major problem which must be overcome before this method can be effective.

In order to interpret the intelligence contained in the sideband the original carrier frequency must be re-inserted at the demodulator stage of the receiver. To avoid distortion of the audio signal, the frequency of the re-inserted carrier must be within a few cycles per second of the original carrier. This means that the receiver must contain a very high stability oscillator. Oscillators with the required degree of stability are now available and suppressed s.s.b. equipments are fitted in some types of Service aircraft.

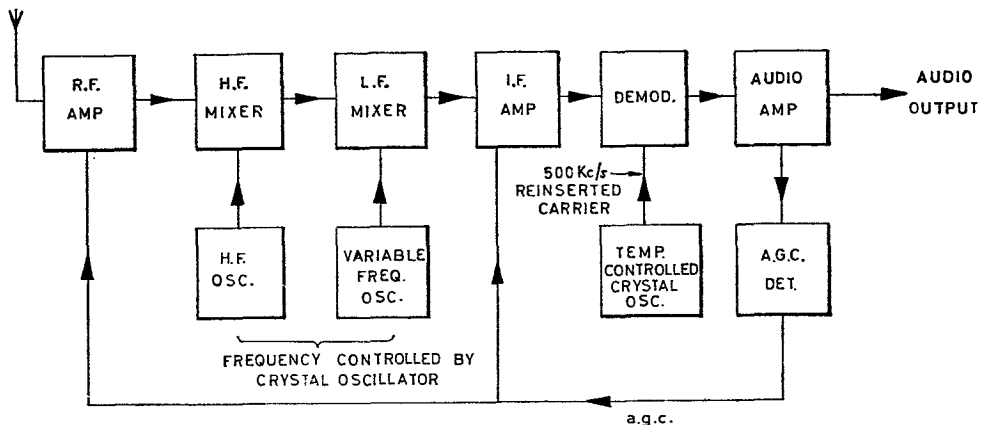A block diagram of a suppressed carrier s.s.b. receiver is given in Fig. 4.

FIG. 4. SUPPRESSED CARRIER SSB RECEIVER

The receiver is a double superhet, both local oscillators being controlled by the frequency of a very stable temperature-controlled crystal oscillator.

Incoming s.s.b. signals are amplified by an r.f. amplifier and then heterodyned down to a variable i.f. in the h.f. mixer. This signal is then mixed with the output from a variable frequency oscillator in the l.f. mixer, the output of which is 500 kc/s plus the audio frequencies. This is then amplified by the i.f. amplifier and passed to the demodulator stage. Here it combines with the re-inserted 500 kc/s carrier obtained from the crystal oscillator. The envelope of the

131

combined wave varies in accordance with the original modulation, and after demodulation, the audio signal is amplified and fed to the headset.

Automatic gain control voltage cannot be obtained from the carrier, since there is not one, so the voltage is taken from the a.f. amplifier. The audio signal is rectified by the a.g.c. detector and the d.c. voltage thus obtained is applied to the r.f. and i.f. stages.

Conclusion

This chapter has reviewed the advantages of a s.s.b. system over normal amplitude modulation and the different types of s.s.b. signal have been discussed. The important features of receivers suitable for the reception of pilot carrier s.s.b. signals and suppressed carrier s.s.b. signals have been considered.

One of the main uses of s.s.b. systems is long-distance high-quality telegraphy and to improve reliability of reception under varying propagation conditions diversity reception is normally used. This is the subject of Chapter 6.

CHAPTER 5
FM AND FSK RECEIVERS

Introduction

The principles of frequency modulation have been covered in Part 1, and f.m. transmitters were dealt with in Section 2. The main differences between a.m. and f.m. receivers lie in the bandwidth required for the i.f. amplifiers, and in the demodulation circuits. A block diagram showing the essential circuits of a f.m. receiver is given in Fig. 1.
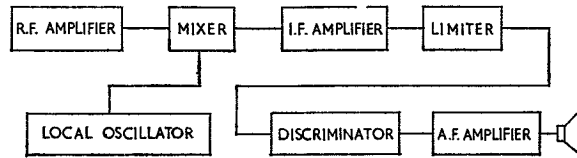

FIG. I. BLOCK DIAGRAM OF A TYPICAL FM RECEIVER

The methods used to increase the bandwidth of the i.f. amplifiers, and the purpose and action of the limiter were explained in Part 1. The two types of discriminator that were discussed, the Travis and the Foster-Seeley discriminators, both have to be preceded by a limiter in order to prevent noise reaching the audio amplifier. A form of discriminator which does not require a separate limiter valve is the *ratio detector*.

The Ratio Detector

The circuit of a basic ratio detector is shown in Fig. 2. It is similar to the Foster-Seeley circuit, but has two major differences in that the connections to one of the diodes are reversed and the capacitor C_3 is a large electrolytic capacitor which maintains a comparatively constant voltage between points A and B. If the strength of the input carrier increases, the voltage across C_3 will increase, but the voltage across C_3 cannot increase at an audio rate.
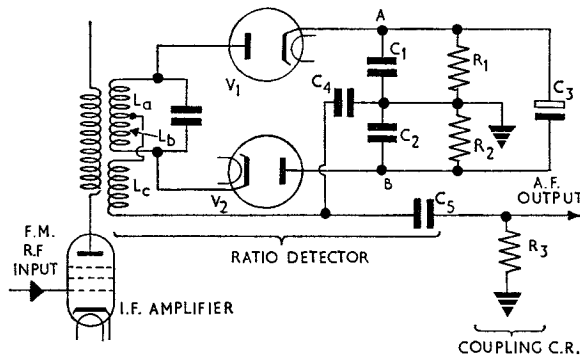

FIG. 2. THE RATIO DETECTOR

Because of this, amplitude variations due to noise do not affect the a.f. output. Such amplitude variations are effectively 'smoothed' by C_3 which is usually of the order of 20 μF.

133

Thus the ratio detector does not have to be preceded by a limiter valve and for this reason it is widely used in f.m. receivers.

Tuning a FM Receiver

Because of the wide bandwidth of an f.m. receiver, it is difficult to tune for maximum output. De-tuning can however cause distortion, especially when large deviations occur. Thus magic-eye tuning indicators are often used to enable accurate tuning to be carried out.

Automatic Gain Control

The f.m. receiver operates at v.h.f. and thus it is used for short range communication with a fairly strong input signal. Hence a.g.c. is only necessary in fringe areas where pronounced fading may cause the incoming signal to become too small to drive the limiter stage correctly.

Where an a.g.c. circuit is necessary, the a.g.c. diode would normally take its input from the last i.f. stage just as in an a.m. receiver.

Pre-emphasis and De-emphasis

In a f.m. receiver noise is considerably reduced by the limiter circuit. However, in practice this circuit does not completely eliminate noise, and a further reduction in noise is obtained by using a pre-emphasis circuit in the transmitter and a de-emphasis circuit in the receiver.

In a f.m. system noise is far more troublesome at the higher audio frequencies. This is because many sidebands in f.m. transmissions are above the audio range and contribute little to the audio signal whereas they contribute a full quota of noise.

To reduce this noise all higher audio frequencies are cut down at the receiver by a de-emphasis circuit. The circuit is usually included in the discriminator output and as shown in Fig. 3, is very similar to a simple r.f. filter.
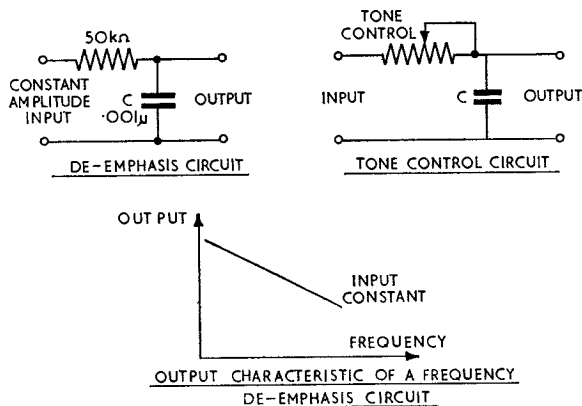

FIG. 3. COMPARISON OF DE-EMPHASIS AND TONE CONTROL CIRCUITS

This circuit cuts down the noise, but unfortunately it also cuts down the higher audio frequencies. To compensate for this a pre-emphasis circuit is included in the transmitter. This circuit and its characteristic are shown in Fig. 4. Its effect is to accentuate the higher audio frequencies in the modulation. Since these are attenuated by the same amount at the receiver as they are accentuated at the transmitter, there is no distortion, but the noise contained in the higher audio frequencies is reduced.
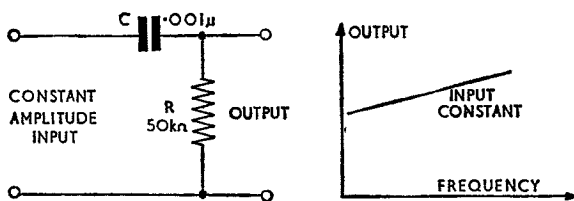
FIG. 4. PRE-EMPHASIS CIRCUIT AND CHARACTERISTIC

The pre-emphasis is usually expressed in microseconds, a standard figure being 50 μs. This means that the time constant of the de-emphasis circuit in the receiver should be 50 μs to restore the receiver response to its correct level. Thus if the pre-emphasis of a transmitter is 50 μs suitable values of C and R in the receiver de-emphasis circuit would be 0·001 μF and 50 kilohm.

FSK Reception

The principle and advantages of a frequency shift keying system were discussed in Section 2. This system is more reliable than simple ON/OFF keying and is widely used on teleprinter tributary networks of the RAF.

The input signal to a f.s.k. receiver shifts about its nominal frequency by $\pm$ 425 c/s. As shown in Fig. 5, the mark input is at the nominal frequency minus 425 c/s and the space input
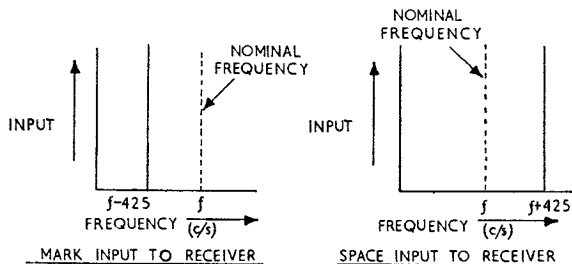

FIG. 5. INPUT TO FSK RECEIVER

is at the nominal frequency plus 425 c/s. Thus the input is really frequency modulated with a very small deviation and so the basic requirements for the reception of a f.s.k. signal are the same as those for the reception of a frequency-modulated signal. The main differences between a f.m. and f.s.k. receiver are summarized as follows:—

a. Narrower bandwidth for f.s.k.

b. More stringent frequency stability precautions required for f.s.k.

c. Selective filters are employed in the demodulator stage to separate the mark and space frequencies of f.s.k.

FSK Receiver

A block diagram of the latter stages of a f.s.k. receiver is given in Fig. 6. The r.f. amplifier, mixer and i.f. amplifier stages would be similar to those of a normal communication superhet, although the bandwidth of these stages could be reduced.
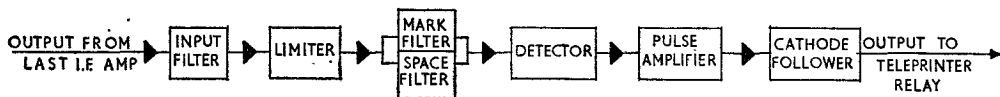
135

FIG. 6.   POST-IF STAGES OF A FSK RECEIVER

The output from the last i.f. amplifier is fed to the input filter which accepts the mark and space frequencies and rejects unwanted interfering signals. The limiter passes constant amplitude mark and space signals to selective filters which separate the mark and space frequencies and pass them to the detector. The positive (mark) and negative (space) pulse outputs from the detector are then amplified in the wideband pulse amplifier whose output is well-shaped pulses of $\pm$ 80 volts suitable for operating the teleprinter relay. The cathode-follower provides a correct match to a 600 ohm line without distorting the pulses.

FSK Diversity Reception

To reduce the effects of fading and thus to increase further the reliability of f.s.k. reception, space diversity reception is normally used. For dual space diversity, two separate aerials spaced several wavelengths apart and feeding two separate receivers are required. The high standard of frequency stability required in f.s.k. reception is catered for by using a common crystal-controlled r.f. oscillator for both receivers. If double or triple superhets are used the stability precautions outlined in Chapter 6 would be adopted.

CHAPTER 6

DIVERSITY RECEPTION

Introduction

The s.s.b. system discussed in Chapter 4 is widely used for long distance telegraphy. To improve the reliability of reception under varying propagation conditions, 'diversity reception' is normally employed. This chapter will be concerned with diversity reception of long distance telegraphy signals using the s.s.b. system.

Sky Wave Propagation Problems

Long distance communication in the h.f. band uses the sky wave; this wave is reflected by the ionosphere and by the earth's surface so that it is 'bounced' round the earth. During its passage of several thousand miles the signal will be subject to extreme variations of propagation conditions. It may start from a temperate zone and pass through either arctic or tropic zones. It may start in daylight and be received at a station on the other side of the world at night. In fact the telegraph signals of the Commonwealth Air Forces network undergo all extremes of path conditions, from day to night, from summer to winter, from tropic to arctic.
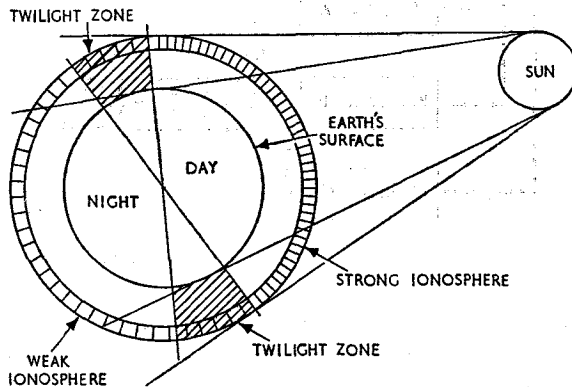

FIG. I. IONOSPHERIC VARIATION

Because the signal path conditions vary, acute fading can occur, particularly when the signal path is via a twilight zone. Fig. 1 shows the variation of ionospheric conditions between day and night.

In the region of sunlight the ionosphere density is maximum and in the shadow of the earth it is minimum. Consequently the bending effect on the radio wave varies and if the variation is of short duration it is termed 'fading'.

Selective Fading

Fading caused by interaction between the ground wave and sky wave can be counteracted by using a.g.c. However, a.g.c. is only effective when the fading occurs over several cycles of the modulating signal, and affects the whole of the signal equally. Propagation conditions are also dependent on frequency. Thus if a signal is modulated, the carrier and the sidebands may be subjected to different degrees of fading. This effect is known as *selective fading* and may cause one part of the signal to fade and other parts to be unaffected.

137

For example, with a s.s.b. signal consisting of carrier and one sideband, selective fading may momentarily make the carrier much stronger than the sideband. If normal a.g.c., which is proportional to the amplitude of the carrier, were used, it could completely suppress the sideband and render the signal meaningless.

Diversity Reception

A more effective way of counteracting all types of fading, including selective fading, is to use *diversity reception*. This means simultaneously receiving the same signal by two or three different means; for example using two or three aerials spaced several wavelengths apart (*space diversity* reception) or by receiving the same signal on two or three different frequencies (*frequency diversity* reception).

Space Diversity

In this form of diversity reception, two or three separate receivers fed by separate aerials, provide a common output. If the aerials are spaced several wavelengths apart it is very unlikely that the signal will fade *simultaneously* in all the aerials. Thus at least one receiver will supply
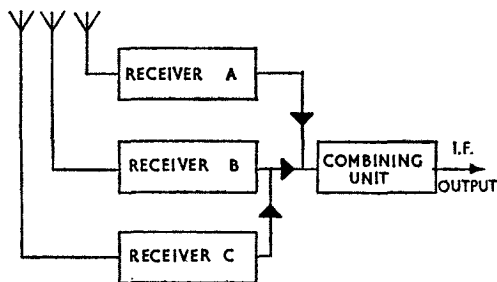

FIG. 2. SPACE DIVERSITY RECEPTION

a useful amplitude output. Fig. 2 shows the outline of an arrangement for *triple diversity* reception.

The combining unit consists basically of three detectors feeding a common load resistor. Thus the three detectors isolate the three inputs in an arrangement similar to that shown in Fig. 3.

Notice that the CR load is fed in parallel from the diodes. Thus only the diode receiving the largest input signal will provide an output. For example, if input B exceeds inputs A and C, the capacitor C will charge to the maximum voltage determined by input B and the voltage across the CR load is of such a polarity as to cut off V_1 and V_3.

In practice a similar arrangement is used for the a.g.c. system, so that an a.g.c. voltage proportional to the *strongest* signal is produced and fed to *all* receivers. In this way receivers with momentarily weaker signals contribute very little in the way of signal *or noise* to the actual output. Thus the signal amplitude and signal-to-noise ratio are always the best of the three inputs.

Use of Common Oscillators

As noted earlier, it is common practice to use double or triple superhets for diversity reception. In principle the double or triple superhet is merely an extension of basic frequency changing; however, in the case of diversity receivers the stability of the local oscillators causes
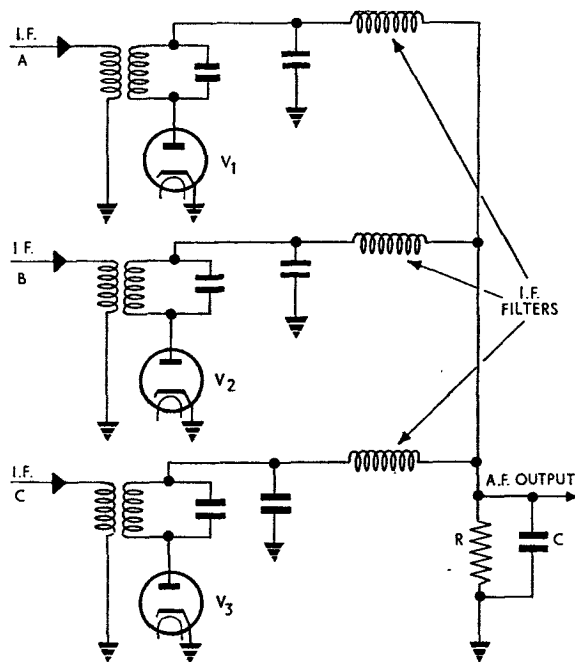
FIG. 3.   COMBINING DETECTORS

a problem.   Since the three receivers provide a common output the effect of random frequency drift in the nine local oscillators of a triple diversity system employing triple superhets, would cause considerable error in the final output.

In order to prevent this, it is usual to make one or two of the local oscillators common to all receivers.

Remember that only the first oscillator need be variable; the second and third oscillators can be pre-set in frequency, since they have only to change one fixed i.f. to another.   Often, the second oscillator is made slightly variable ($\pm$ 5 kc/s) and its frequency is controlled by an a.f.c. voltage in order to stabilise the final i.f.   Thus the accuracy of the i.f. is dependent on the stability of the third oscillator which is usually crystal controlled.   Fig. 4 shows the frequency changing sequence of such a receiver.   Also shown are typical frequencies for oscillators and i.f.'s.
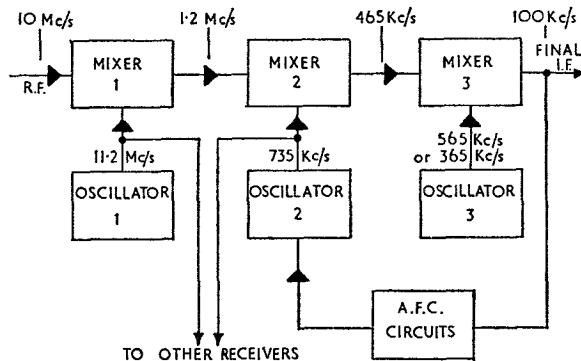

FIG. 4.   FREQUENCY CHANGING SEQUENCE

Triple Diversity SSB Receiver

The outline of a typical triple-superhet, triple-diversity s.s.b. receiver arrangement is shown in Fig. 5. Only one of the three receivers is shown in any detail together with the common oscillator and combining units.
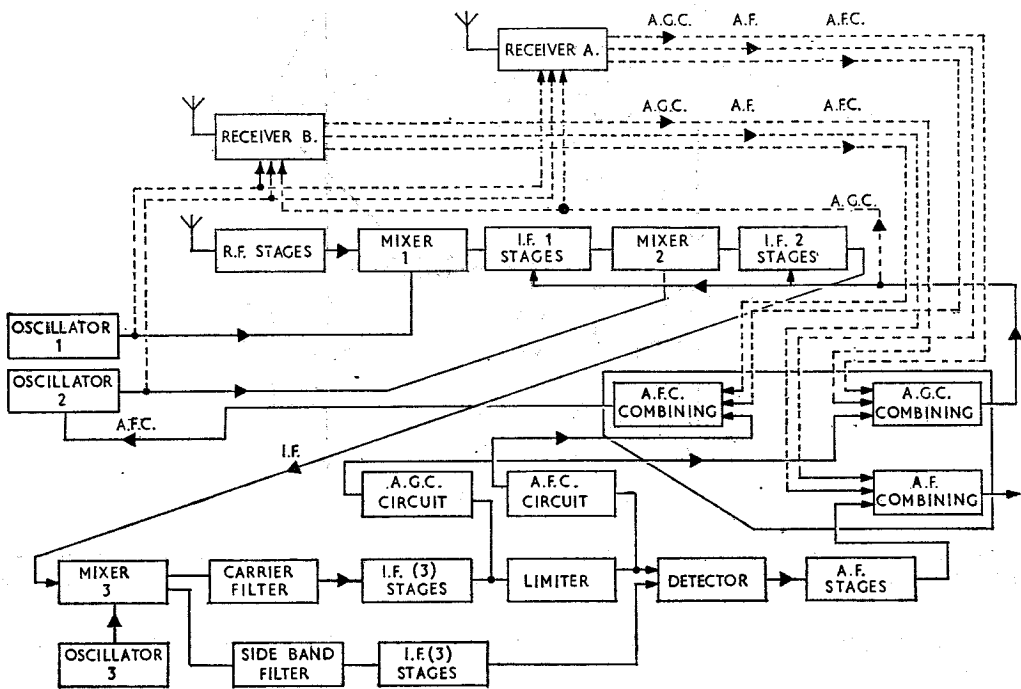

FIG. 5. BASIC SSB DIVERSITY RECEIVER

The received s.s.b. signal is changed down in three successive frequency changer stages, the output from the final mixer being 100 kc/s plus modulation. At this point the carrier is filtered out to produce an a.g.c. voltage proportional to the amplitude of this particular carrier. Also, after limiting, the carrier produces an a.f.c. voltage proportional to the frequency of the carrier.

The output from the sideband filter is amplified and then applied to the detector stage. As the amplified and limited carrier is also applied to the detector stage, the sideband frequencies and carrier frequency combine and demodulation takes place in the normal way.

At the combining unit, all three receiver a.g.c., a.f.c. and audio outputs are compared so that only the highest a.g.c. voltage, the strongest audio output and the mean a.f.c. voltage are effective as final outputs. The a.g.c. and audio outputs are d.c. voltages, but since the a.f.c. comes from a discriminator it can be either a positive or a negative voltage.

In this way the best possible reception under varying propagation conditions is obtained. However, three aerials, three receivers, three frequency changing processes with common oscillators, combining units and power units are necessary. Consequently this system is only used where maximum reliability in long-distance communication is essential.

Frequency Diversity

In this form of diversity reception the signal is transmitted with a slight modulation of frequency, in addition to the normal amplitude modulation which produced the s.s.b. signal. The slight frequency modulation is not intended to convey any intelligence but solely to wobble the carrier frequency slightly and so reduce the chances of total fading.
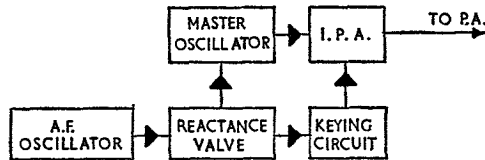
FIG. 6. FREQUENCY DIVERSITY TRANSMISSION

Fig. 6 shows the outline of the transmitter arrangement. No special features are involved in the receiver since the frequency deviation is only a matter of a few hundred cycles per second.

There are also frequency diversity systems in which the same signal is transmitted on two different frequencies simultaneously, sometimes from two different aerials. The signals are received in two separate receivers and then fed to a combining unit which selects the better output. This system, which is employed on microwave links, is discussed more fully in the next section.

Other Diversity Systems

One cause of fading is due to the change in polarisation of the wave as it passes into an ionised layer during reflection. Thus a wave which leaves the transmitter aerial with vertical polarisation may arrive at the receiver aerial with horizontal polarisation. This wave will induce only a small signal voltage into the vertical receiver aerial. This form of fading can be combated by a *polarisation diversity* system. The signal is radiated from two separate aerials, one vertical and the other horizontal. With a similar arrangement of vertical and horizontal aerials at the receiver, the stronger of the two signals can be selected.

Another form of diversity reception, sometimes employed in scatter propagation, is known as *wave angle* diversity. This is discussed in the next section.

SECTION 4

TELEGRAPHY

CHAPTER 1

LINE TELEGRAPHY

Introduction

Telegraphy is a system used to communicate the written word over long distances. In its simplest form, a telegraphic system could consist of some form of "sender" joined to a distant "receiving" device by wires. This is *line telegraphy*. If the sending and receiving equipments are linked by electromagnetic waves instead of wires, the system is known as *radio telegraphy*.

Basic Telegraphic Systems

In all telegraphic systems the intelligence to be sent is converted into some form of code. Perhaps the best known telegraphic code is the Morse code which consists of combinations of dots and dashes representing letters and numbers. In the simple telegraph system illustrated in Fig. 1, the sender is a morse key and the receiver is a buzzer. This buzzer is rather like an electric bell without the gong. When it receives the long and short pulses of current sent down the line, it makes a buzzing sound corresponding to the keyed dots and dashes of the Morse code. A trained operator can convert these symbols back into letters and numbers.
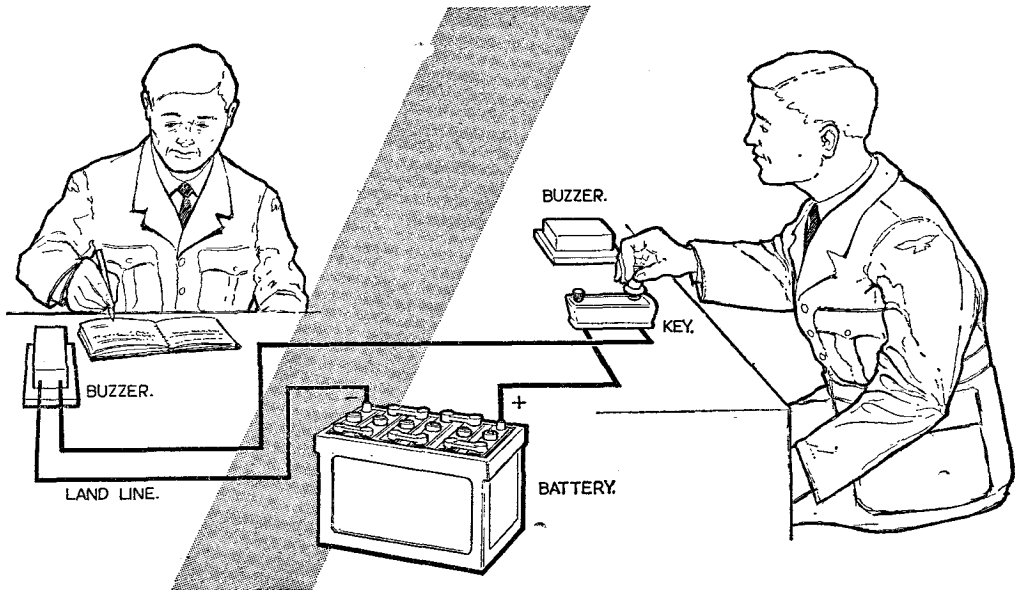

FIG. 1. SIMPLE LINE TELEGRAPHY

This simple system shows the principle on which all line telegraphy is based. It has many disadvantages, the main one being that it is hand operated. The volume of signals traffic in modern communication systems could not be handled by such a simple device. Most modern systems are automatic and operate at high speed. The Morse code is not suitable for use on automatic systems so a code known as the *Murray code* and others similar to it are used.

Various types of machines are used on automatic systems, the best known one being the *teleprinter*. This machine looks like a large typewriter but when the keys are pressed they operate a mechanism which makes and breaks the connection between a source of voltage and the

lines. In this way, pulses of current corresponding to the Murray code are transmitted. A teleprinter at the other end of the line then "translates" the received current pulses into corresponding movements of its own keys and the received message is recorded.

It is not always convenient or necessary to transmit a signal as soon as it is written. In such cases the message is typed out on a machine with a keyboard similar to that of a teleprinter but instead of feeding current pulses to a line, the machine punches holes in a paper tape (Fig. 2). The punched tape is then stored and some time later it can be fed into another machine which "reads" the holes in the tape and transmits the appropriate current pulses.
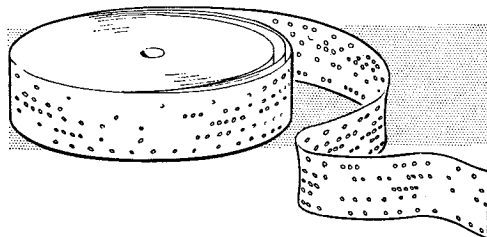

FIG. 2. TELEGRAPH TAPE

At the receiving end, the current pulses can operate a teleprinter or another perforating machine which will again record the message on tape. The tape can then be stored, used for retransmitting the message, or it can be printed later.

Telegraphic machines such as teleprinters which are connected to the landline are termed "on line" devices; a perforator, which need not be connected to line is often called an "off line" device.

Another method of telegraphic communication is known as *facsimile*. With this system a photo-electric device scans the message and sends along the line current pulses of different amplitudes corresponding to the different shades of the message. At the receiving end, the pulses make marks of different shades on electro-sensitive paper. In this manner a complete picture of the original message is reproduced. Facsimile is not used for transmitting written messages because it is too slow and it is not suitable for fine detailed photographs. However, it is ideal for sending documents such as meteorological charts.

Telegraph Codes

There are several codes used in the various telegraphic communication systems just outlined. The code used for a particular system depends upon the following factors:—

a. The form of the intelligence to be transmitted, i.e. whether it consists of letters, numbers, line drawings or half-tone photographs.

b. Whether all the available code combinations can be fully used.

c. Whether the message is to be sent by hand signalling or by automatic means.

d. The degree of accuracy required.

Three of the best known codes will now be considered.

The International Morse Code

With this code the letters of the alphabet, numbers, punctuation and other signs are represented by combinations of dots and dashes. These can be transmitted by two methods. With one method, the two conditions for signalling are a current for the dot or dash state and no current for the interval between dots or dashes. This method is illustrated in Fig. 3.

With the alternative method a positive current represents one condition and a negative current, i.e. one flowing in the opposite direction, represents the other condition (see Fig. 4.).

The two methods are known as *single current* working and *double current* working respectively.

Morse code is satisfactory for manually operated telegraph systems but the unequal length of the dots and dashes (dash = 3 dots) and the unequal lengths of letters in the code make it unsuitable for teleprinter operation.
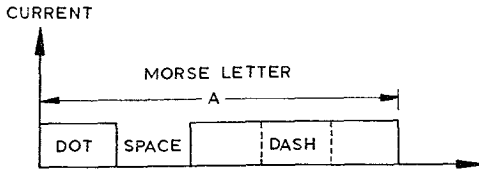

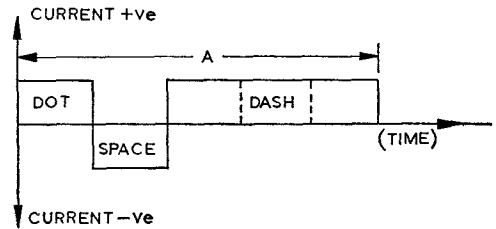FIG. 3. MORSE CODE WAVEFORM (SINGLE CURRENT WORKING)

FIG. 4. MORSE CODE WAVEFORM (DOUBLE CURRENT WORKING)

The Five-unit Code

This code is used on automatic teleprinter networks. The particular type of five-unit code used in the Service is called the *Murray code*. In this code the number of elements is the same for every character (letter, figure, etc.), and the duration of each element is constant. This code therefore overcomes the disadvantage of the Morse code in which the characters are composed of long and short elements.

Each character of the Murray code consists of five elements and each element is in one of two signalling conditions: either the marking condition or the spacing condition. With this arrangement a total of $2^5 = 32$ combinations are possible. The letter F in the five-unit code is shown in Fig. 5.
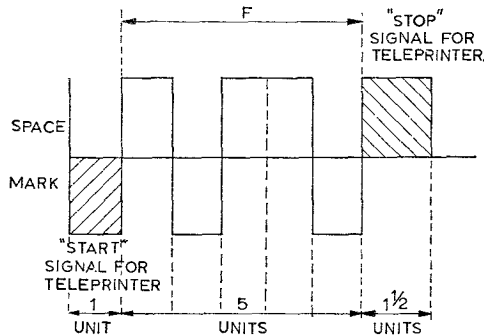

FIG. 5. THE MURRAY CODE WAVEFORM

The 32 combinations available with this code are insufficient for the transmission of 26 alphabet letters together with figures, so each combination is given two meanings. This is done by sending a "letter shift" signal prior to letter characters and a "figure shift" signal prior to figure characters. Thus to send "20 AIRCRAFT" the figures 2 and 0 on the keyboard would be pressed, followed by "letter shift" and then the word "AIRCRAFT".

In order to start and stop the receiving teleprinter motor as each character is sent, a "start" signal (1 unit) is added to the beginning of each character and a "stop" signal (1½ units) follows the character. Thus the basic 5 unit code really contains 7½ units.

The time taken to send each element or unit of code is 20 milliseconds. Thus, neglecting the teleprinter start and stop signals, each five unit character occupies 100 milliseconds, i.e. 1/10th of a second.

The Seven-unit Code

All five-unit codes suffer from the disadvantage that attenuation of the signal, or interference, can change the meaning of a sequence of five pulses by filling in a space or dropping out a mark. Thus in Fig. 5, if interference had caused the first element of the letter F to be a mark instead of a space, the teleprinter would have registered the letter N. By using a seven-unit code such errors can be made self-detecting. The principle is that all characters are made up of 3 mark pulses and 4 space pulses. The loss of a mark produces a character having 2 marks and 5 spaces which would be registered by a gap in the signal or by the printing of a special sign, e.g. TEL?PRI?TER.

Bandwidth

For faithful reproduction by radio equipment, the bandwidth must be related to the frequencies which make up the signal. Bandwidth requirements are equally important in telegraphic line communication.

If the steady current in a circuit is keyed with a code waveform, the circuit no longer carries only direct current. The keyed pulses have a square waveform and this can be considered as consisting of a large number of sine waves. In theory a square wave consists of a fundamental, the frequency of which depends on the keying speed, plus an infinite number of odd harmonics of the fundamental. Thus, theoretically, the telegraph system would have to pass an infinite band of frequencies for perfect reception.

Since there is a limit to the working bandwidth of lines, keying waveforms must be limited in bandwidth as much as possible without introducing excessive distortion. Another reason for limiting the bandwidth of the keyed signals is that with certain systems several messages can be sent simultaneously along the same line. Each message occupies a different band of frequencies and the overall bandwidth must not be excessive if distortion is to be kept to a minimum.

A reasonable compromise between bandwidth and distortion is obtained by filtering out all harmonics above the third. If this is done the resulting keyed waveform is as shown in Fig. 6.
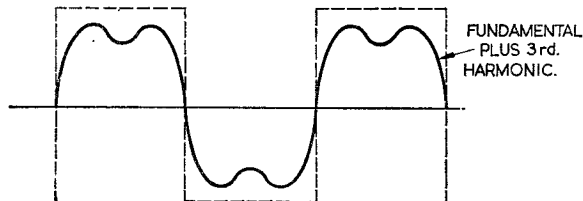

FIG. 6. PRACTICAL KEYING WAVEFORM

Morse Code Bandwidth

To obtain an approximate figure for the bandwidth needed to send a signal in Morse code the following assumptions are made:—

a. The average word consists of 5 letters.

146

b. The average letter occupies 9 dot-units. For example, the morse symbol for the letter **V** consists of 3 dots and a dash, the dash being equivalent in length to 3 dots; each dot is separated from every other dot (or dash) by a space equivalent to a dot; thus to transmit the letter V, 9 dot-units are required. Some characters require less than this, others require more.

c. The space between adjacent letters is equal to 3 dot-units.

d. The space between adjacent words is equal to 7 dot-units.

As the average word consists of 5 letters (45 dot-units) with 4 inter-letter spaces (12 dot-units) and one word space (7 dot-units), the number of dot units required to transmit the average word is:—

$45 + 12 + 7 = 64$ dot-units.

If the speed of transmission is 30 words per minute, then 30×64 dot-units are transmitted per minute, i.e. 32 dot-units *per second*.

Two dot-units constitute one cycle of the fundamental keying frequency; therefore the fundamental frequency is 32/2 c/s or 16 c/s and the third harmonic is 48 c/s. Thus a bandwidth of 0 — 50 c/s would suffice for a keying speed of 30 words per minute.

Murray Code Bandwidth

In the Murray code the time taken to transmit each element is 20 milliseconds. Thus the time taken to transmit two adjacent elements is 40 milliseconds and this corresponds to a fundamental keying frequency of 25 c/s. The third harmonic of this fundamental is 75 c/s and so a bandwidth of 75 c/s will suffice.

From the two examples given it can be seen that the greater the signalling speed the greater is the bandwidth required.

Note that the bandwidth required for normal speech reception is about 4,000 c/s compared with 75 c/s for the Murray code. This means that more telegraphic channels can be accommodated in a given band of frequencies and a better signal-to-noise ratio obtained than is the case with telephony.

Telegraph Signalling Speed

The unit of telegraphic signalling speed is the *baud* which is defined as the reciprocal of the time taken (in seconds) to transmit the shortest element. Thus the signalling speed, measured in bauds, gives a direct indication of the bandwidth required.

a. **Morse code.** In the example used above, if the sending speed is 30 words per minute each dot-unit takes 1/32 seconds to send. Thus the telegraphic speed in bauds is 32.

b. **Murray code.** Each element takes 20 milliseconds to send and thus the telegraphic speed is 50 bauds.

Telegraph Distortion

If the bandwidth of the telegraph circuit is insufficient for the signalling space required, the signal will be distorted, i.e. it will not be faithfully reproduced at the receiving end. Thus signalling speed is limited by bandwidth.

Other factors can also introduce distortion of the telegraphic signal. For example, the code elements can be distorted by various parts of the telegraph system. Amplitude distortion can be caused by attenuation of the current pulses as they travel along the line. This type of distortion can easily be corrected by placing amplifiers along the circuit between transmitter and receiver.

The most serious form of telegraph distortion is the alteration in length of the code elements. Because of this, telegraph distortion is defined in terms of the element length.

$$\text{Distortion} = \frac{\text{Change in length of the shortest element.}}{\text{Normal length of the shortest element.}}$$

Distortion is divided into the following classes:—

a. **Characteristic distortion.** This is peculiar to special signal combinations and to the electrical characteristics of the line or relay in use.

b. **Fortuitous distortion.** This is caused by erratic variations due to irregularities in the circuit. These irregularities can be caused by sparking at relay contacts or by interference from external sources such as electrical storms or power cables.

c. **Bias distortion.** This form of distortion causes lengthening of the marking or spacing elements due to asymmetry in the transmitting or receiving apparatus, e.g. differing battery voltages.

Telegraph Relays

Magnetic relays are widely used in line telegraphy and they play an important part in the efficient working of a telegraphic system.

A relay possesses inductance and therefore, when a voltage is applied to it, the current through the relay coil does not rise immediately to its final value but increases relatively gradually as shown in Fig. 7.

Thus the relay contacts will not close immediately a voltage is applied to the coil, but will take some time to respond. When used for telegraphy, the response time of a relay is often more important than its sensitivity since the latter can be aided by amplifying the applied signal. The response time is largely determined by the time constant L/R of the relay. As can be seen from Fig. 7, unless the input voltage exists for at least L/R seconds, the relay current will be negligible and the relay contacts will not close. To reduce the time constant and so shorten the response time, a resistance (R) can be placed in series with the relay coil. Alternatively, the effective inductance (L) of the coil can be reduced by placing a capacitor (C) in series with the coil. If a capacitor is used, it must be shunted by a resistor in order to allow direct current to flow. Both these arrangements are shown in Fig. 8.
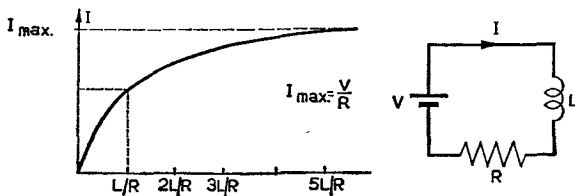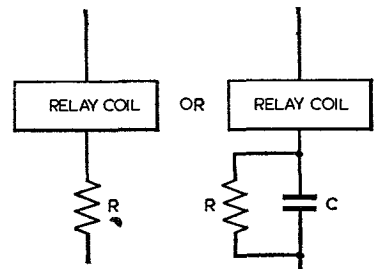

FIG. 7. RISE OF CURRENT IN A RELAY COIL

FIG. 8. METHODS OF SHORTENING RELAY RESPONSE TIME

There are two main classes of telegraphic relay, polarised and non-polarised relays.

a. **Non-polarised relays.** A non-polarised relay is one in which the direction the armature moves is independent of the direction of the controlling current through the windings. It is therefore suitable for single current working only and requires some means of opening

148

the contacts once the energising current has ceased. A spring, or magnetic bias is necessary to do this.

In modern telegraph circuits the non-polarised relay has been replaced by the polarised relay.

b **Polarised relay.** A polarised relay is one in which the direction of movement of the armature depends on the direction of the *resultant* controlling current through the relay coil. For double current working, i.e. when current flows in either direction in the relay coil, the armature is adjusted neutrally, so that neither contact is made. This is illustrated in Fig. 9, which shows the magnetic circuit of one form of polarised relay.

When no signal is being received from the telegraph line, the only magnetic flux present is that due to the permanent magnet. The armature, placed centrally between the two U-shaped pole pieces, will therefore remain stationary since the pull from each of the pole pieces is the same. When a current from the line flows through the armature coil, a magnetic field is set up and the armature acts as a magnet. Under the conditions shown in Fig. 9, the armature would have a north pole at the top and a south pole at the bottom. This is shown in Fig. 10.
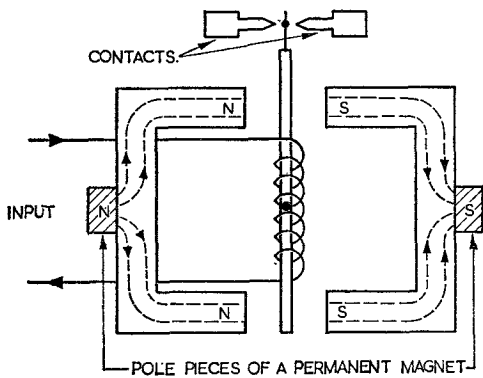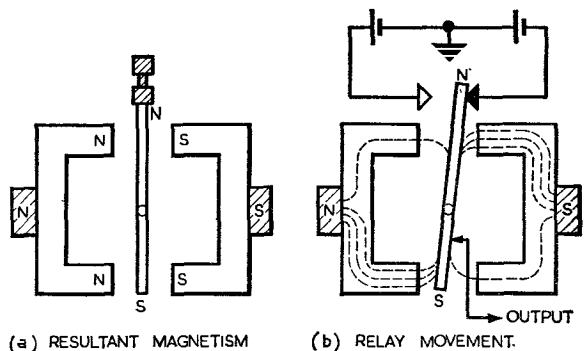

FIG. 9. POLARISED RELAY

(a) RESULTANT MAGNETISM   (b) RELAY MOVEMENT.

FIG. 10. OPERATION OF A POLARISED RELAY

The attraction between unlike poles will cause the armature to swing to the position shown in Fig. 10*b*. The movement is assisted by the repulsion force of the like poles. When the direction of the current through the coil reverses, the polarity of the armature reverses and the armature moves to the opposite pole piece.

The polarised relay has the following advantages over the non-polarised relay:—

a. There is no residual magnetism, since the current reverses on each signal.

b. It is more sensitive for small currents.

c. It is essential for double-current working.

d. It reduces the mark-to-space distortion due to variations in signal amplitude. This last point is illustrated in Fig. 11.
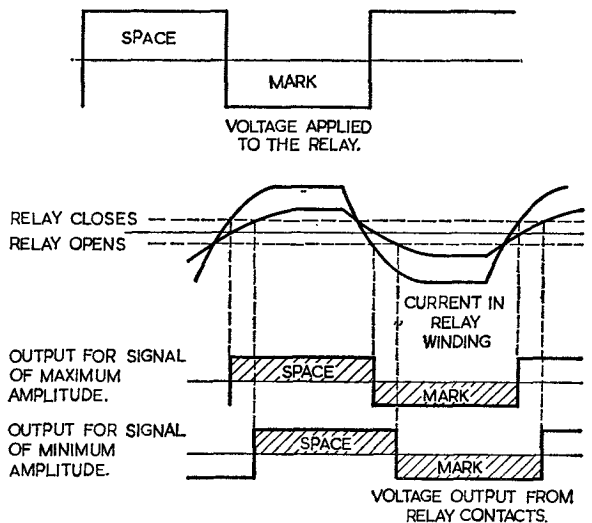

FIG. 11. ADVANTAGE OF POLARISED RELAY

149

It can be seen that the lengths of the mark and space signals remain almost the same despite a change in signal amplitude. The corresponding results with single current working, using a non-polarised relay, are shown in Fig. 12. Note the change in length of the mark and space signals between maximum and minimum amplitudes.
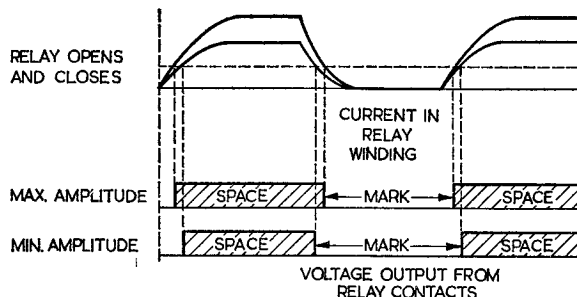
FIG. 12. DISADVANTAGE OF NON-POLARISED RELAY

Basic Line Systems

The usefulness of the simple form of line system consisting of key and buzzer described earlier is limited by voltage drop along the line.

This simple system can be improved by using a relay, operated by the line current pulses, to apply a *local* source of voltage to operate the buzzer. Fig. 13 illustrates a simple relay-operated system.
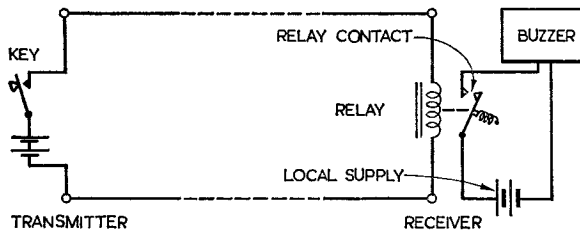
FIG. 13. USE OF A RELAY IN A SIMPLE TELEGRAPH SYSTEM

As long as the line current is of sufficient amplitude to operate the relay, the relay contact closes and applies the full local battery voltage to the buzzer. All practical line communication systems are based on this type.

The system shown in Fig. 13 is a single current system; this means that current flows only one way through the relay coil. With double current working, current is sent in one direction through the relay to indicate a mark and in the opposite direction to indicate a space. This means that a polarised relay must be used.

The outline of a double current system is shown in Fig. 14. When the transmit key is on its
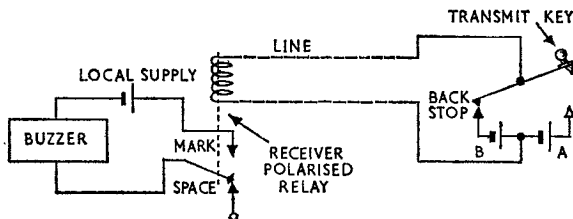
FIG. 14. DOUBLE CURRENT WORKING

150

back-stop, battery B is connected to the line and a current is sent through the polarised relay coil so that the relay contact moves to the space position. When the key is pressed down, battery A sends a current of the opposite sense down the line and the polarised relay contact moves over to the mark position.

The advantages of double current working are as follows:—

a. Higher working speeds are possible.

b. Less frequent adjustment of the relays is required.

c. Mark-to-space distortion is reduced, as explained earlier.

Direct Current Line Telegraphy

The line circuits shown in Figs. 13 and 14 are known as *simplex* circuits. Only one message can be sent along the line at one time and in one direction only. Thus for inter-communication two separate circuits would be needed.

With the *duplex* system of telegraphy, messages can be sent in opposite directions along one line simultaneously. The duplex system thus greatly increases the communication efficiency of the system.
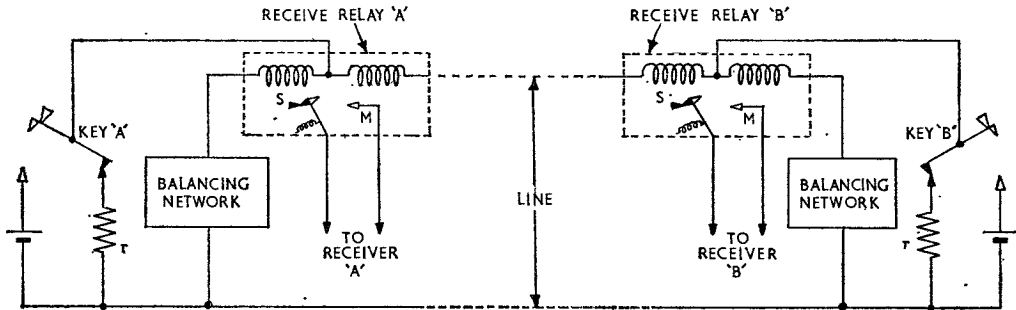

FIG. 15. BASIC DUPLEX SINGLE CURRENT SYSTEM

The outline of a basic duplex single current circuit is shown in Fig. 15. Each relay winding is centre-tapped; one half is connected to the line and the other half to a balancing network which simulates the line. The network consists of inductance, capacitance and resistance so that when the key is pressed equal currents flow to the line and to the balancing network. The resistance
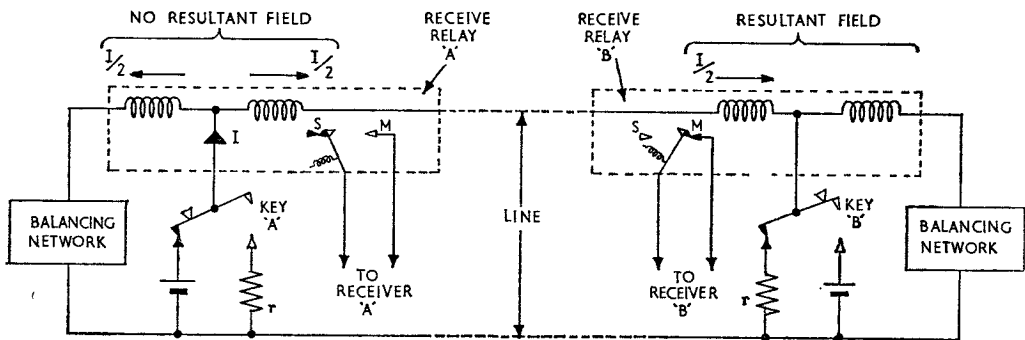

FIG. 16. CURRENTS IN THE DUPLEX SYSTEM

"r" is equal to the internal resistance of the battery so that the impedance of the circuit is the same with the key either up or down.

Fig. 16 illustrates what happens when key A is pressed and key B is raised. Equal currents flow in opposite directions through relay A and so there is no resultant field. This relay is not energised and stays in the "space" position. Received current from the line flows in one direction through both parts of relay B winding and produces a field which moves the contact over to the "mark" position.
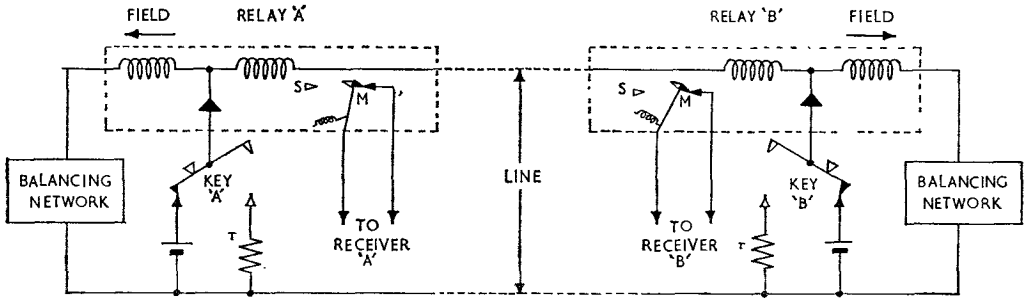

FIG. 17. SIMULTANEOUS TWO-WAY WORKING

Fig. 17 shows the conditions when both keys A and B are in the send position. No current flows through the line, since both batteries have similar poles connected to the line. However, both relays are energised by current through the coils connected to the batteries through the balancing networks. Thus both relays move over to the mark contact and both receivers register a mark.

In this way, for either position of key A and key B a corresponding mark or space is recorded at the distant receiver.

The system just described is a single current duplex system but the principle applies also to double current working. Polarised relays and double-current keying would, of course, have to be used.

Telegraph Repeaters

When a signal is transmitted over long lines or over radio links, the amplitude of the pulse may be considerably reduced. Telegraph *repeaters* consist of sensitive relays controlling local batteries which "repeat" the signal with sufficient power for the next stage of its journey. (Repeater is also the name given to electronic amplifiers placed at intervals along a line such as transoceanic cables). Fig. 18 shows the basic idea of a double current duplex telegraph repeater.
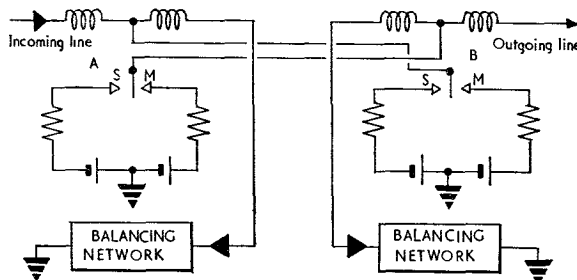

FIG. 18. DOUBLE CURRENT DUPLEX REPEATER

The arrangement consists of two centre-tapped polarised relays A and B and two balancing networks.   The relay contacts can switch a local battery of either mark or space polarity to the outgoing line.

Thus in Fig. 18 a space signal from the left operates relay A and the negative pole of a local battery is switched to the centre point of relay B.   Equal currents will flow in opposite directions through the two halves of relay B winding, and so relay B will remain unenergised.   Thus the amplified space signal will be transmitted down the outgoing line.

Similarly, if a mark signal is received from the right, relay B switches a positive potential to the centre of relay A windings; relay A is not energised, but the amplified version of the incoming mark signal is sent on down the line.

Regenerative Repeaters

When the signalling speed is very high or the transmission link is poor, the signal may not only be weakened but may also be distorted.   In this case amplification merely adds to the distortion and in extreme cases the mark and space elements become blurred and unrecognisable. To overcome this in teleprinter circuits, *regenerative* repeaters are used.   These generate a new signal produced by "sampling" the incoming signal at the centre of each element.

Fig. 19 shows the basic components of a regenerative repeater.   The incoming signal is fed through an input relay to an electronic input switch which controls the voltage on lines A and B. When a mark is received, line A is at a higher voltage than line B, and when a space is received, line B has a higher voltage than line A.
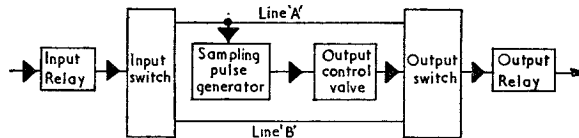

FIG. 19. BLOCK DIAGRAM OF A REGENERATIVE REPEATER

At the beginning of a $7\frac{1}{2}$ unit code signal the START element triggers the sampling pulse generator via line A.   The generator then issues a train of seven narrow pulses, each pulse occurring at the instant at which the centre of each succeeding element occurs in a $7\frac{1}{2}$ unit code signal.   This is shown in Fig. 20.
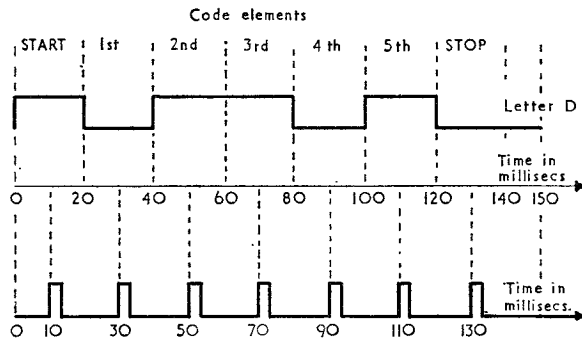

FIG. 20. TIMING OF SAMPLING PULSES

The sampling pulses are applied, via a valve, to the output switch which is another electronic switch similar to the input switch.   The output switch measures the voltage of the lines A and B during the instants of the sampling pulses and causes the output relay to transmit a mark or space element depending on the line voltage.

153   *(A.L. 3, September, 1964)*

Thus the output of the repeater is a series of marks and spaces in the same order as received but at the correct intervals as determined by the sampling pulses. Because only the centre of the "start" signal is sampled, no notice is taken of any changes in the duration or timing of the succeeding elements. This is illustrated in Fig. 21.
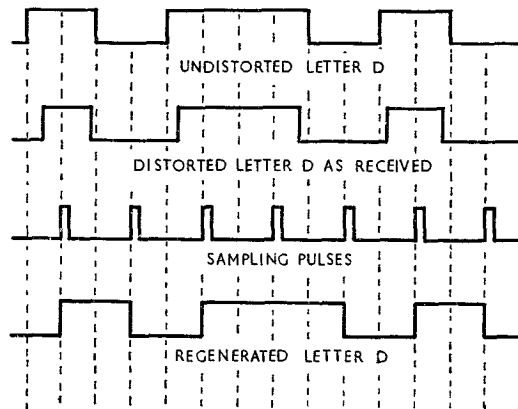

FIG. 21. REGENERATIVE ACTION

The instant at which each element of the regenerated signal begins is determined by the sampling pulse generator and therefore occurs at exactly 20 millisecond intervals, irrespective of the timing of the distorted input. Whether the element is a mark or a space is determined from the incoming signal by the input switch. In this way, the appropriate elements of the correct spacing are re-transmitted.

The circuits used in the various stages of the regenerative repeater are dealt with in Chapter 3.

Summary

Some of the important points of line telegraphy dealt with in this chapter are summarised as follows:—

Codes. The Morse code is used in telegraphy for hand signalling and for automatic high speed morse signalling. It is not suitable for use with teleprinters, where the 5 unit Murray code is used. 7 unit codes overcome the possibility of a character being wrongly received due to distortion or interference.

Bandwidth. For satisfactory reception of code pulses, the fundamental and third harmonic must be passed by the line. The bandwidth of a signal depends on the type of code used and on the signalling speed. The unit of telegraphic signalling speed is the baud.

Distortion. Changes in the length of code elements result in telegraphic distortion. Distortion can be caused by line characteristics, faulty relays and incorrect battery voltages.

Relays. Polarised relays make double-current working possible. The response time of a relay can be improved by reducing the time constant of the winding.

Duplex working. With this method, messages can be sent simultaneously in both directions along the line.

Repeaters. These are placed at intervals along a long line to compensate for line losses. They reproduce the incoming signal by switching a local battery supply to the outgoing line. Regenerative repeaters correct distortion in the incoming signal.

CHAPTER 2

FILTERS AND ATTENUATORS

Introduction

The general principles and properties of filters have been considered in previous chapters dealing with power supplies and other electronic circuits where it was desired to separate a frequency or band of frequencies from some other frequency. In the case of the power supply circuit the "ripple" had to be separated from the d.c. An i.f. filter "filters out" the i.f. component of the detected wave, leaving the a.f. component to be applied to the a.f. amplifier stage.

If a signal is reduced in strength it is said to be *attenuated*. Attenuation is thus the opposite of amplification. Sometimes it is required to reduce deliberately the amplitude of the whole of an incoming signal; this means that all the component frequencies of the signal must be attenuated by the same amount. This is done with a device known as an *attenuator*.

Filters and attenuators are widely used in line and radio telegraphy and this chapter will deal with the properties and types of these devices. The difference between a filter and an attenuator is that the filter will attenuate signals at different frequencies by *different* amounts, i.e. it is frequency sensitive, while an attenuator will reduce the amplitude of signals of different frequencies by the *same* amount.

Filters

A filter consists of a network of inductors and capacitors and ideally has the following properties:—

a. It will pass, without attenuation, at least one range of frequencies.

b. It will completely block at least one other range of frequencies.

c. It will form a correct match between the source of energy it is filtering and the load.

The range of frequencies over which attenuation is ideally zero is called the *pass-band*, whilst the range which is subject to a large amount of attenuation is called the *attenuation-band* or *stop-band*.

The frequency which separates the pass-band and the attenuation-band is called the *cut-off frequency*.

Classification and Types of Filter

There are four main tasks which filters must perform and the filter is classified accordingly:—

a. **Low-pass.** Attenuates the high frequency range and passes the low frequency range.

b. **High-pass.** Attenuates the low frequency range and passes the high frequency range.

c. **Band-pass.** Attenuates all frequencies except a chosen band.

d. **Band-stop.** Attenuates all frequencies within a chosen band and passes all other frequencies.

The *ideal* attenuation characteristics for each of these four types are shown in Fig. 1.

In practical filters, the cut-off frequency is not as clearly defined as in the ideal case although filters have been designed with characteristics which closely approach the ideal ones shown in Fig. 1.

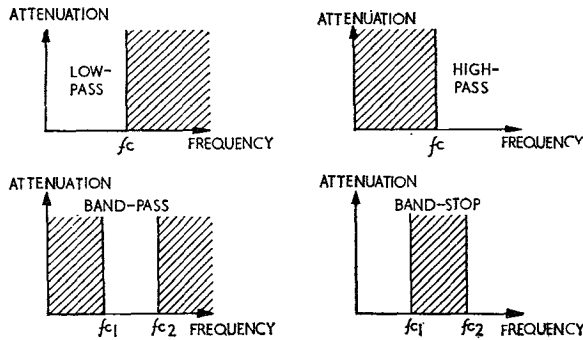Each of the four main types of filter will now be considered.

FIG. 1. IDEAL ATTENUATION CHARACTERISTICS

Low-pass Filter

The circuit of a low-pass filter is given in Fig. 2. Since the reactance of the inductor increases as the input frequency increases and the reactance of the capacitor decreases as the input frequency increases, the output, taken across the capacitor, will fall as the frequency rises.

High-pass Filter

With the high-pass filter shown in Fig. 3, the positions of L and C are reversed. Thus as the input frequency increases, greater voltage is developed across the inductor, and less across the capacitor. Hence higher frequency inputs will be less attenuated than lower frequency inputs.
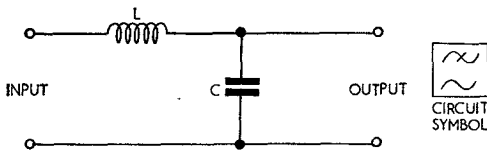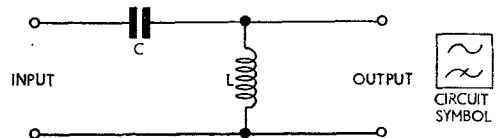

FIG. 2. LOW-PASS FILTER   FIG. 3. HIGH-PASS FILTER

Band-pass Filter

Resonant circuits are used in the construction of filters which will pass without attenuation all frequencies within a chosen band and attenuate frequencies outside that band. The response curves of series and parallel tuned circuits are shown in Fig. 4.
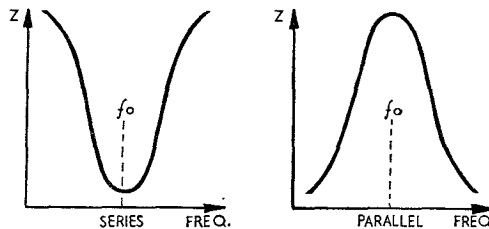

FIG. 4. TUNED CIRCUIT RESPONSE CURVES

Fig. 5 shows a band-pass filter circuit which uses a series tuned circuit and a parallel tuned circuit. Both circuits are tuned to the same resonant frequency f_o. For frequencies near f_o, the impedance of the series tuned circuit is very small whilst that of the parallel tuned circuit is

156

very high. Hence voltages at these frequencies will pass through the filter with only slight attenuation, whilst at other frequencies voltages will be attenuated and will not appear in the output.

Band-stop Filter

A filter which passes all frequencies outside a given band and "stops" frequencies within the band is shown in Fig. 6. In this case the output is taken across the series tuned circuit which, for frequencies near f_o, will have a very low impedance. Thus the output at frequencies near f_o
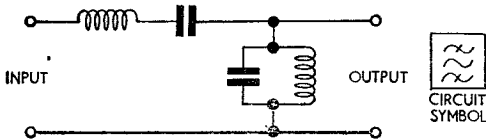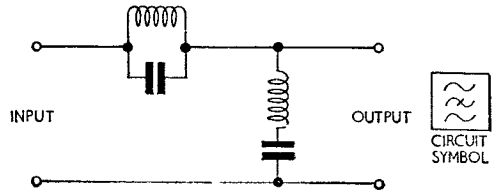

FIG. 5. BAND-PASS FILTER   FIG. 6. BAND-STOP FILTER

will be very low. For frequencies outside this band the series circuit will present a high impedance and the output will be similar to the input.

Practical filter circuits are more complicated than the simple examples discussed but they work on the same principles. They consist of a number of sections, similar to the simple circuits shown, connected in cascade. This arrangement produces a sharper cut-off frequency and better attenuation in the stop-bands. They also provide better matching at all frequencies in the pass-band.

Narrow-band Magneto-striction Filters

A disadvantage of a filter constructed of inductors and capacitors is that its characteristic is not stable. Further, a filter with a sharply defined cut-off frequency must consist of several sections and, for low frequency filters, this means using many heavy chokes and capacitors.

A very stable small filter has been developed based on the magneto-striction property of ferrite. The filter consists of a rod of ferrite, free to vibrate about its centre and excited by a single-turn primary coil. This is wound on a paxolin tube surrounding the ferrite rod and an adjacent winding serves as an output coil. Alternating current in the primary causes the rod to vibrate. When the input frequency is near the natural resonant frequency of the rod, there is a sharp rise in the electro-mechanical coupling between the input and output coils resulting in a very peaked response. A slug of iron dust provides fine tuning.

This type of filter has a very narrow bandwidth and is used as a sideband filter in s.s.b. transmitters.

Crystal Filters

Another alternative to the multi-section LC filter is the quartz crystal filter. The selectivity factor of a quartz crystal cut to a certain frequency is many times that of its equivalent LC circuit; thus its discrimination between two nearby frequencies is better. A crystal filter has a very clearly defined cut-off frequency and is very stable. Two crystals can be used to provide a band-pass filter, one crystal resonating at the higher cut-off frequency and the other at the lower cut-off frequency.

Matching

One of the requirements of a filter is that it must be capable of providing a correct match between a generator and its load. A generator is correctly matched to its load if the internal

impedance of the generator equals the load impedance. Thus if a filter is placed between the generator and its load, the input and output impedance of the filter must equal the load impedance, Z. (See Fig. 7).
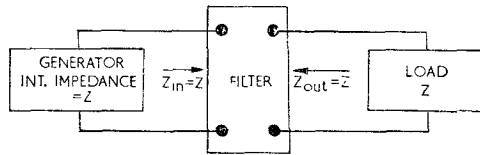

FIG. 7. MATCHING

To maintain a correct match, the impedance which is "seen" on looking back into the filter when the load is removed must equal Z. Also the impedance seen on looking into the filter when the generator is removed must be Z. Thus the filter network must have the same value of input and output impedance.

By suitable choice of component values the filter can be made to have the required impedance at the frequencies at which it is intended to operate. This impedance is known as the *characteristic impedance* of the filter. The filter in Fig. 7 would have a characteristic impedance of Z ohms.

The Filter as a Load

Because filters can be designed to present a particular impedance at a given frequency, they are often used as dummy loads to take the place of, say, a transmission line. Filters used for this purpose are called *artificial lines* and to the generator they appear equal to the line. The balancing networks used in the duplex system are examples of the use of artificial lines.

Attenuators

Often it is required to reduce the strength of a signal at a certain point in a circuit. For example, several separate signals may be fed into a common system which has a maximum permissible power or voltage rating. Each signal must be attenuated to a common level by separate attenuators so that the total voltage or power does not exceed the permissible rating. Since the signal may consist of many different frequency components all of which must be equally attenuated, a resistive network, known as an *attenuator*, is used. The principle of an attenuator is shown in the simple circuit of Fig. 8.

V_{out} is at a lower value than V_{in}, the difference between the two values depending upon the values of R_1 and R_2. The attenuation can be continued by feeding the output across R_2 into another potential divider network, and so on. This is shown in Fig. 9.
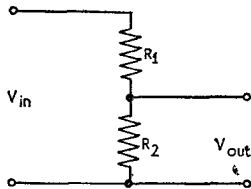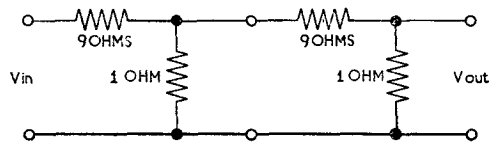

FIG. 8. RESISTIVE POTENTIAL DIVIDER

FIG. 9. ATTENUATOR NETWORK

Such an attenuator network can be designed to reduce the input by a certain amount proportional to the original input. It can also be designed to have a particular characteristic impedance. Thus an attenuator can be inserted into a circuit to provide the required attenuation without upsetting the matching between source and load.

The amount of attenuation introduced by an attenuator is usually measured in decibels (it can also be measured in nepers). Thus, if the signal leaving the attenuator is 5 db below the input to the attenuator, then it is called a 5 db attenuator.

CHAPTER 3

PULSE CIRCUITS

Introduction

Although often considered peculiar to radar, pulses are widely used in communication systems. In fact the original form of pulse transmitter was the simple hand-operated morse transmitter. With high speed automatic telegraph equipment the need has arisen to produce pulses of short duration; "short" means of the order of milli- or micro-seconds. Such pulses are used in pulse communication systems.

The circuits which handle these pulses are called *pulse circuits* and include all circuits dealing with pulses of voltage or current—pulse generators, pulse amplifiers and so forth. This chapter is mainly concerned with the production of voltage pulses.

Pulse Shape

The pulses used in telegraphy are called *square waves* and are illustrated in Fig. 1. A square wave is one whose waveform has right-angled corners. To be more accurate, they are rectangular waves, but the term square wave is usually applied to all right-angled waveforms.

As shown in Fig. 1, the pulse edge which occurs first in time is called the *leading* edge, and the edge which terminates the pulse is termed the *trailing* edge.

When the interval between pulses is very long compared to the duration of the pulse itself, the waves are often called *narrow pulses*. (Fig. 2).
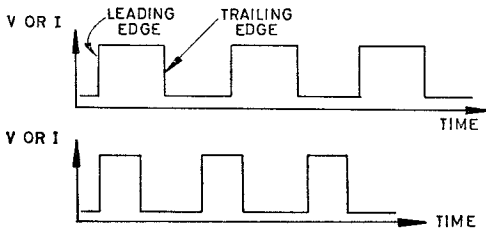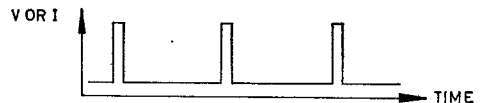

FIG. I. SQUARE WAVES

FIG. 2. NARROW PULSES

Squaring the Sine Wave

Perhaps the simplest method of producing an approximate square wave is by merely cutting off the top and bottom of a sine wave. All that is required to obtain a fairly square wave shape, is that the amplitude of the sine wave be large. Fig. 3 shows the effect of the input sine wave amplitude on the shape of the square wave output.

Square wave 1, produced by cutting off the top and bottom of a small amplitude sine wave (i.e., *limiting* the sine wave) has pronounced slopes to the leading and trailing edges. In other words the rise time of the leading edge and the fall time of the trailing edge are large.
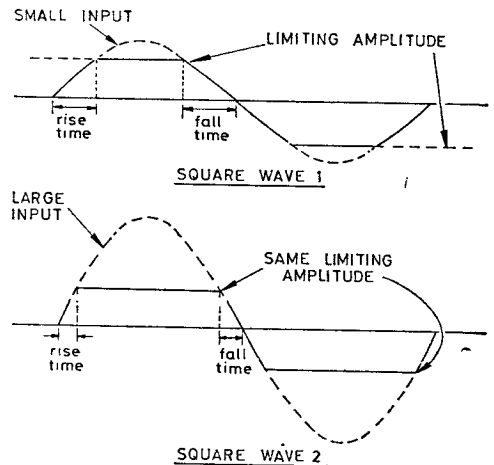

FIG. 3. EFFECT OF SINE WAVE AMPLITUDE

The amplitude of the sine wave used to produce square wave 2 is fairly large and the resulting square wave has much steeper leading and trailing edges.

Waveform 2 is much nearer a square wave than is waveform 1. Thus by making the output only a small fraction of the input, quite a good square wave is obtained.
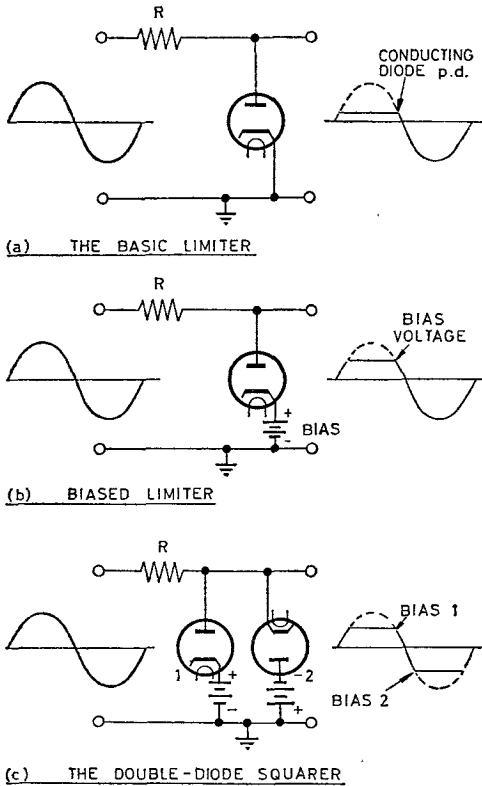
Squaring with Diodes

One method of producing a square wave from a sine wave input is shown in Fig. 4.

The circuit of Fig. 4a is basically a potential divider network consisting of the resistor R and the diode. R must be of a fairly high value compared with the resistance of the diode when it is conducting. A typical value of R would be about 100 kilohms. The input sine wave is applied across the network and the output is taken across the diode; this will be very small when the diode conducts during the positive half-cycle of input.

Thus the output from the circuit of Fig. 4a consists of the negative half-cycle of input and a very small voltage necessary to cause conduction on the positive half-cycle.

In the circuit of Fig. 4b the diode is biased so that it will not conduct until the positive half-cycle of input exceeds the bias value. Thus the output from this circuit is the whole of the negative half-cycle, plus that part of the positive half-cycle below the bias level.

In Fig. 4c a second diode is connected across the output with its *anode* taken to earth via a *negative* bias. Diode 1 conducts on the positive peaks and diode 2 conducts on the negative peaks and hence the peaks of the input sine wave are eliminated. The output is thus an approximate square wave.

FIG. 4. DIODE SQUARERS OR LIMITERS

These circuits are sometimes called *limiters* or *clipping circuits* because they limit or clip off a part of the input waveform above or below a certain level decided by the value and polarity of the bias. They are used to eliminate unwanted parts of a waveform but are not often used to produce a square wave from a sine wave; the amplifying circuits described in the next paragraph are generally used.

Squaring with Triodes and Pentodes

If the input voltage to an amplifier is large enough to drive the valve into saturation or cut-off, then limiting occurs. (See Fig. 5).

The anode voltage has the same waveform as the anode current but it is in antiphase to it, i.e., when I_a rises to a maximum at the saturation point, V_a falls to a minimum. An amplifier can be used in this way to produce a square wave from a sine wave input. It has a big advantage over the diode limiter in that the triode or pentode acts as an amplifier and a small amplitude sine wave input will result in a steep-sided square wave output. An amplifier used in this way is sometimes called an *overdriven* amplifier.

Another method of limiting the positive half-cycle of grid input is by placing a resistor in the grid lead. The grid-to-cathode circuit now becomes similar to the circuit of Fig. 4a, the grid of the triode (or pentode) replacing the anode of the diode. This is illustrated in Fig. 6. The series grid resistor is called a *grid stopper*.

When the grid is driven positive by the input, grid current flows and most of the input voltage is dropped across the grid stopper, only a very small part appearing between grid and cathode. On the negative half-cycle, no grid current flows and all the input is applied between grid and cathode.

The waveforms associated with grid limiting are shown in Fig. 7.

When a high value anode load resistor is used in a pentode amplifier, an effect known as *bottoming* occurs. If the grid voltage is increased above a certain value, there is no corresponding increase in anode current and so the anode voltage remains constant at a low value.

This effect can be used to limit the tips of the positive half-cycles of an input sine wave. If the pentode is biased so that part of the negative half-cycle is eliminated by the cut-off point of the valve, a square wave output results. This is illustrated in Fig. 8.
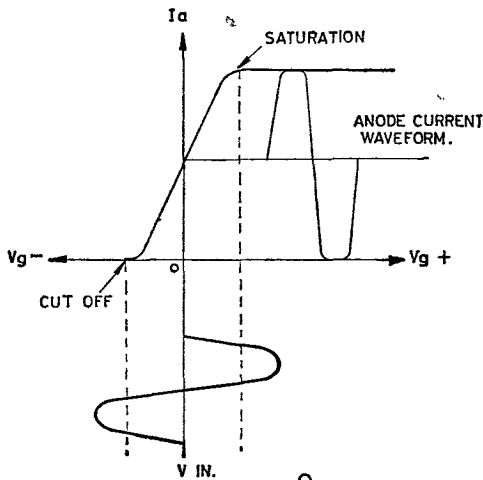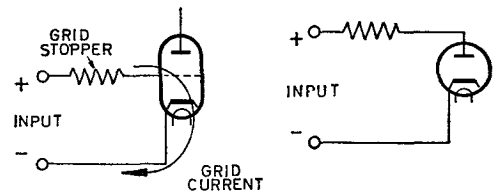
THIS IS THE SAME AS THIS

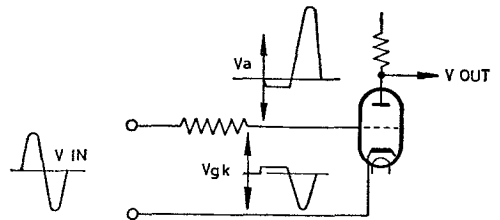FIG. 6. GRID CURRENT LIMITING

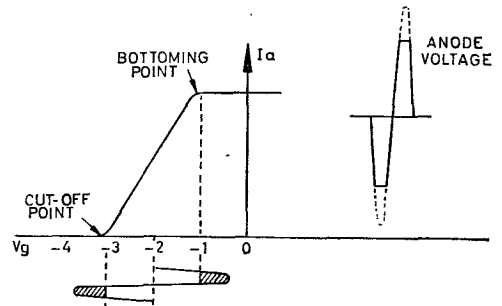FIG. 7. GRID LIMITING WAVEFORMS

FIG. 5. OVERDRIVEN AMPLIFER

FIG. 8. PENTODE SQUARER USING BOTTOMING EFFECT

Transistor Squaring Circuit

Transistors and semiconductor diodes, instead of thermionic valves, may be used to square a sine wave. Semiconductor diodes would be just as effective in the circuits of Fig. 4 as the valve diodes.

161   (*A.L.* 3, *September*, 1964)

A transistor squaring circuit is shown in Fig. 9*a* and the associated waveforms in Fig. 9*b*.

Only the negative going half-cycles of input sine wave cause current to flow. If the collector load resistor R is large, the transistor will bottom, as in the case of a valve, and the positive peaks
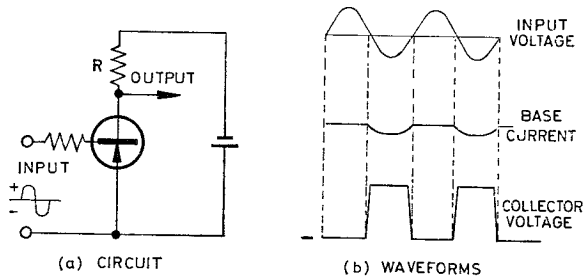

FIG. 9. TRANSISTOR SQUARING CIRCUIT

are not reproduced. The output, taken from the collector, is thus an approximate square wave. An improved square wave is obtained by feeding the output from this circuit into a similar circuit.

The Differentiating Circuit

A method of producing narrow pulses and "pips" is to apply a square wave to a *differentiating* circuit. Such a circuit consists of a capacitor and resistor in series, such that the time constant CR of the circuit is short compared to the time of half a cycle of the square wave applied to the circuit.

Consider a voltage step (the start of a square wave) applied to the CR circuit shown in Fig. 10. It is already known that:—

a. the charge on a capacitor cannot change instantly.

b. the capacitor charges exponentially at a rate which depends on the time constant of the circuit.

Therefore, since the charge in a capacitor is proportional to the voltage across the capacitor, at the instant the voltage step is applied no voltage appears across C and all of the step appears across R. As C charges, the voltage across it rises and less and less appears across R.

When C is fully charged, there is no charging current and no voltage across R; all the applied voltage is across C. This is shown in Fig. 11.
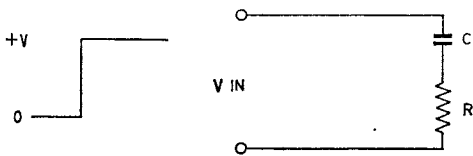

FIG. 10. VOLTAGE STEP APPLIED TO A CR

FIG. 11. VOLTAGE ACROSS R

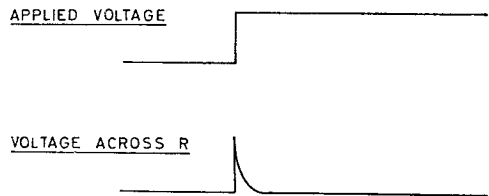If the time constant of the circuit is short compared to the length of time the voltage is applied to the circuit, the voltage across R is a narrow pulse or pip with a steep leading edge and a curved trailing edge.

If a complete square wave is applied to the circuit the same thing happens at the trailing edge as happened at the leading edge, except that the capacitor now discharges through the

162

resistor. Since the current is reversed, the voltage across R is a negative-going pip. Thus a train of square waves produces a series of positive and negative pips as shown in Fig. 12.
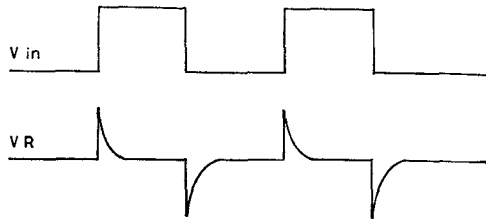
FIG. 12. A DIFFERENTIATED SQUARE WAVE

Either positive or negative pips can be removed by a limiter, leaving a train of narrow pulses of the required polarity. The shape of these pulses can be improved by applying them to an amplifier followed by a limiter.

Relaxation Oscillators

It is often necessary to know accurately the instant at which the leading edge of a pulse occurs. In telegraphy this is necessary in order to avoid variations in the length of code elements. The rise time of the leading edge of a pulse produced by squaring a sine wave is often too long and a pulse with a steeper leading edge is required. Also, a differentiating circuit requires a square wave input in order to produce a narrow output pulse.

Such square waves are produced by a family of oscillators called *relaxation oscillators*. One of these, the *multivibrator* will now be considered.

The Multivibrator

In order to maintain oscillations in an amplifier circuit, part of the output of the amplifier must be fed back in phase with the input. If an oscillator using two amplifier stages is used, the output of the second stage is in phase with the input to the first stage, and a part of the output could therefore be fed back to the input to maintain oscillations. A Franklin oscillator uses this principle.

The basic anode-coupled multivibrator shown in Fig. 13 also uses this principle, the whole of the output of the second valve being fed back to the input to the first. The circuit consists of
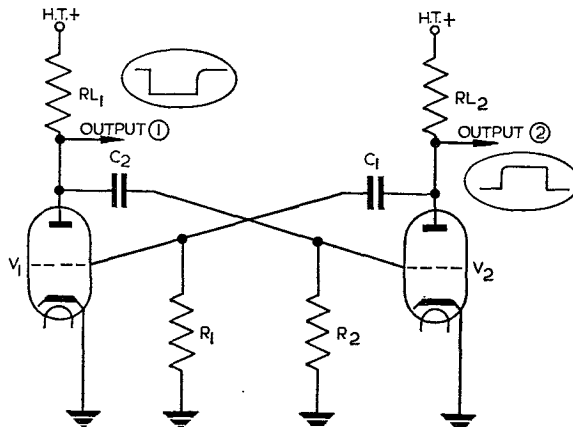
FIG. 13. BASIC ANODE COUPLED MULTIVIBRATOR

163

two RC-coupled amplifiers V_1 and V_2 in cascade, with the output of V_2 coupled via C_1 and R_1 to the grid input circuit of V_1.

The multivibrator is really a two-valve high-speed switching circuit. If V_2 is conducting, V_1 is switched off by a negative potential on its grid: then, due to an *avalanche* or *cumulative* action which takes place very rapidly, V_2 is switched off and V_1 conducts. This switching action continues from the time the supplies are switched on to the circuit until they are switched off. No input other than h.t. and l.t. supplies is required and because of this the circuit is known as a *free-running* square wave generator.

An output taken from V_1 anode consists of a train of square waves, and one taken from V_2 anode consists of similar square waves but in anti-phase to those at V_1 anode. The length of time that either valve is switched off depends upon the value of the grid CR. Thus the values of C_1 and R_1 and of C_2 and R_2 govern the width of the square waves and hence the frequency.

Switching Action

Assuming that V_2 is conducting, there will be a heavy current flowing through R_{L2} and the anode voltage of V_2 will be low (Fig. 14). V_1 is not conducting and its grid voltage is below cut-off. This voltage is not constant but is risin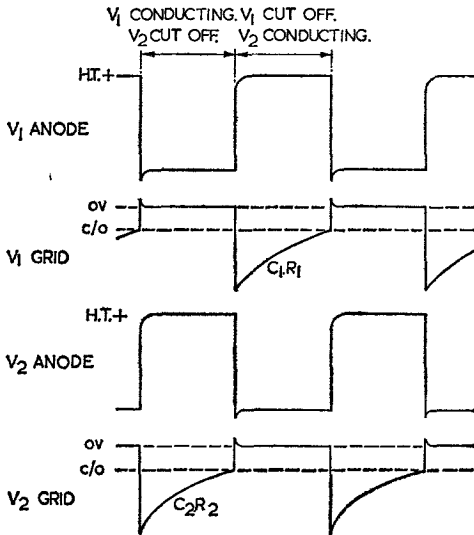g towards cut-on. When this cut-on value is reached, V_1 starts to pass current which flows through R_{L1} and the anode voltage of V_1 therefore falls. This fall is much greater than the grid rise which caused it because V_1 is an amplifier: the fall is applied as a voltage step to C_2 and R_2. Thus the grid voltage of V_2 falls, as in the differentiating circuit of Fig. 10.

Because the grid voltage of V_2 falls, the current through R_{L2} falls and the voltage at V_2 anode rises sharply. This rise is applied as a voltage step to $C_1 R_1$, and V_1 grid voltage rises. Current through R_{L1} increases and V_1 anode voltage falls. This further fall is again applied to V_2 grid via $C_2 R_2$ and results in V_2 passing more current.

This action occurs very rapidly, because both valves are amplifiers, and is called an *avalanche* or *cumulative* action. The result is that V_2 grid voltage is driven down below cut-off almost instantaneously and the potential of V_1 grid is taken up to zero volts very rapidly. Thus V_1 is switched on and V_2 is cut off.

FIG. 14. MULTIVIBRATOR VOLTAGE
WAVEFORMS

After this almost instantaneous action comes the relaxation period in which C_2 discharges through R_2 and the conducting valve V_1 to the new low voltage at V_1 anode (see Fig. 15). Thus the negative grid voltage of V_2 will rise gradually towards earth. The rate of this rise will depend on the time constant $C_2 R_2$.

When V_2 grid voltage reaches cut-on another avalanche action occurs, V_2 grid voltage rises rapidly to zero volts, and V_2 anode voltage falls rapidly from h.t. + to a low positive value. V_1 grid voltage is driven well below cut-off and the voltage at V_1 anode rises to h.t. +.

It is now the turn of C_1 to discharge slowly through R_1 and the conducting valve V_2 and bring the grid voltage of V_1 to cut-on. The cycle is thus continued and square waves with very rapid leading and trailing edges are formed at the anodes of V_1 and V_2. The anode and grid waveforms of these two valves are shown in Fig. 14.

The length of the positive-going pulse at V_1 anode is governed by the time constant $C_1 R_1$, and $C_2 R_2$ controls the length of the positive-going pulse at V_2 anode. Thus by making R_1 and R_2 variable, the width of the pulses and hence the frequency of the square wave outputs can be varied.

The curves on the anode waveforms of both valves is due to the presence of the capacitors C_1 and C_2 and could be removed by a limiting circuit.

Frequency Stability

Any variation in the values of the components of a multivibrator circuit due to temperature changes, or change in value of the supply voltages, can cause an alteration in the frequency of the output square waves. The frequency stability of the basic multivibrator can be considerably improved by taking the grid leaks R_1 and R_2 to a positive voltage instead of to earth. (See Fig. 16).
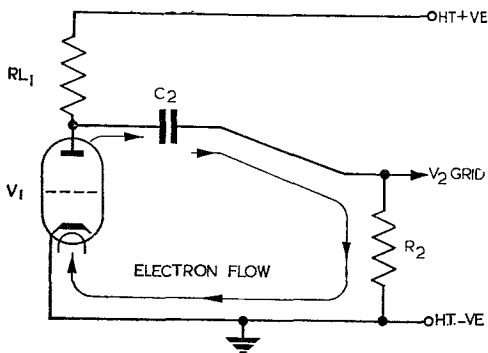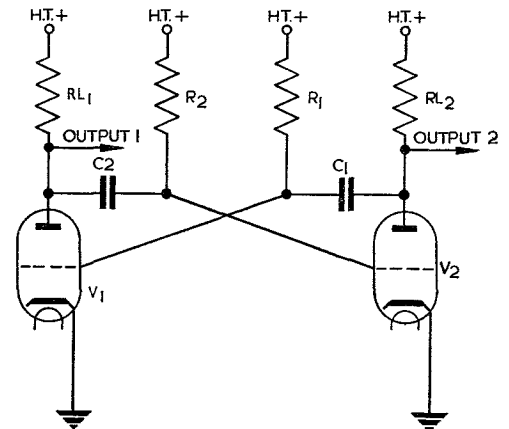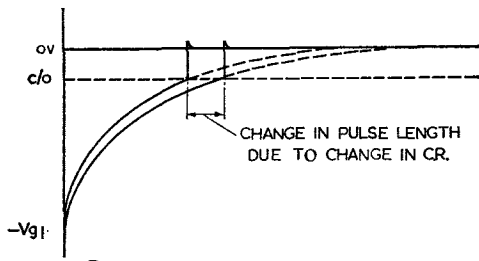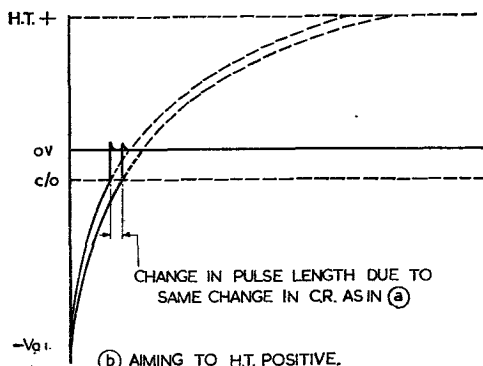

FIG. 15. DISCHARGE CIRCUIT

FIG. 16. MODIFIED MULTIVIBRATOR CIRCUIT

(a) AIMING TO ZERO VOLTS.

(b) AIMING TO H.T. POSITIVE.

FIG. 17. FREQUENCY STABILITY

Fig. 17a shows the grid waveform of V_1 when R_1 is taken to earth. A slight variation in the value of R_1 will result in a slightly different rise time. Since C_1 is almost discharged when the curve reaches the cut-on line, an appreciable change in the pulse width of the anode waveform, and hence in the frequency results.

Fig. 17b shows the grid waveform when R_1 is taken to h.t. positive; C_1 is "aiming" at a final voltage of h.t. and is only slightly discharged when its voltage reaches cut-on. Therefore the curve is fairly steep when it reaches cut-on, and for the same change in the value of R_1 the change in pulse length is much less than when the aiming voltage is zero volts. Thus the frequency stability of the circuit is increased.

It can be seen from Fig. 17 that by taking the grid leaks of V_1 and V_2 to h.t. positive the output frequency is increased. This can be compensated for by increasing the values of the grid

165　　　　　　　　　　　(*A.L.* 3, *September,* 1964)

CR's. If the grid leaks are taken to a variable aiming voltage as shown in Fig. 18, frequency of the output square wave can be conveniently varied.
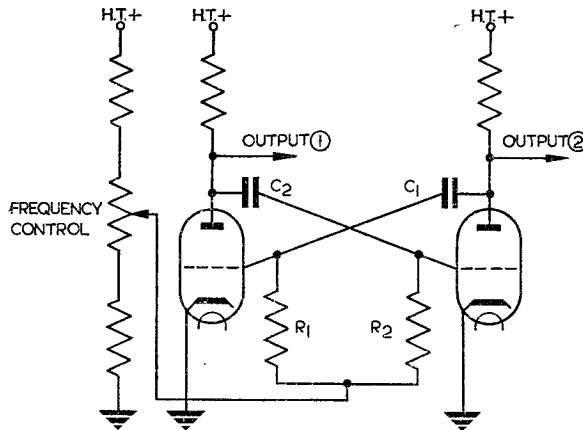

FIG. 18. MULTIVIBRATOR FREQUENCY CONTROL

Synchronisation

The anode coupled multivibrator is a free running or *astable* oscillator. The frequency of its output is decided by the values of components used and by the applied voltages. It is often necessary to have several such square wave generators operating at *exactly* the same frequency. This can be done by *synchronising* all the multivibrator circuits with a master timing circuit. Thus the trailing edge of each cycle of square wave oscillation in each multivibrator exactly coincides with the master timing pulse. Fig. 19 illustrates the principle.
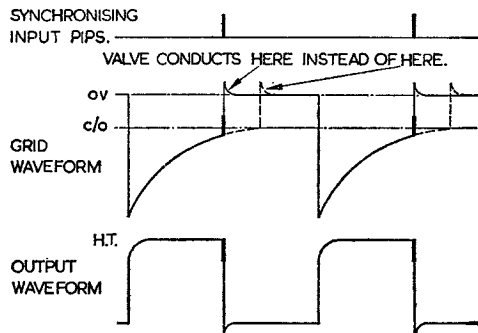

FIG. 19. SYNCHRONISATION OF THE MULTIVIBRATOR

The synchronising pips from the master timer are fed to the grid of either V_1 or V_2 of the multivibrator. The frequency of these pips is slightly *higher* than the free-running frequency of the multivibrator. Thus the valve is brought into conduction slightly before it would normally conduct and at the exact instant that the synchronising pip occurs. The falling edge of the output wave form therefore coincides with the input pip and the multivibrator is synchronised.

If the frequency of the input pips is considerably increased, the effect shown in Fig. 20 results. The first input pip cuts the valve on just before it would normally reach cut-on and the
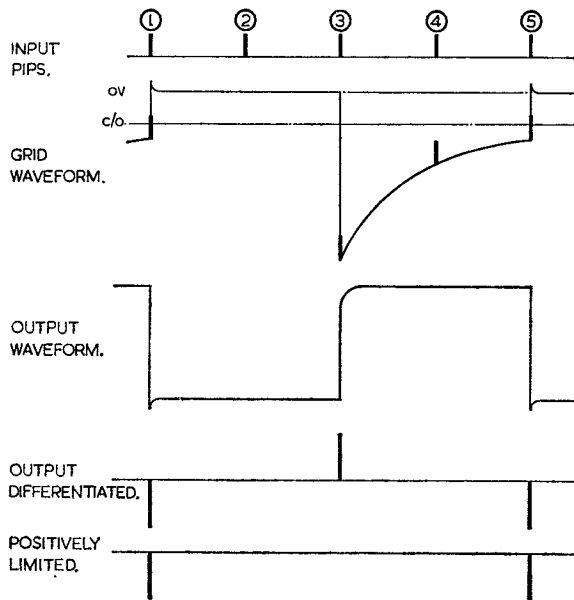
FIG. 20. COUNTING-DOWN BY MULTIVIBRATOR

multivibrator action is started. The second, third and fourth pips have no effect on the action, but the fifth pip starts another multivibrator cycle. Thus one cycle of multivibrator output is produced for every four input pips and the circuit is said to have *counted-down* by a ratio of 4 : 1. Multivibrators can be used in this way to count down by a ratio of up to 10 : 1.

Transistor Multivibrator

A basic multivibrator circuit using transistors in place of valves is shown in Fig. 21. The similarity between this circuit and that of Fig. 16 is immediately apparent, and the action is very similar.

The collector and base waveforms for this circuit are shown in Fig. 22. Initially TR_1 is conducting with its base at a negative voltage and its collector almost zero. TR_2 base is at a
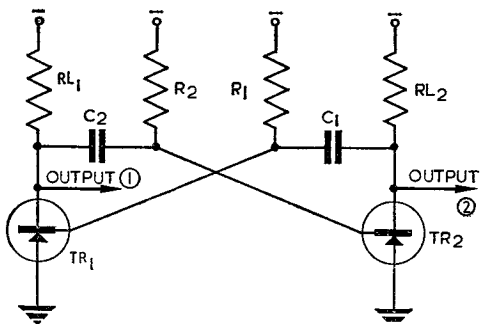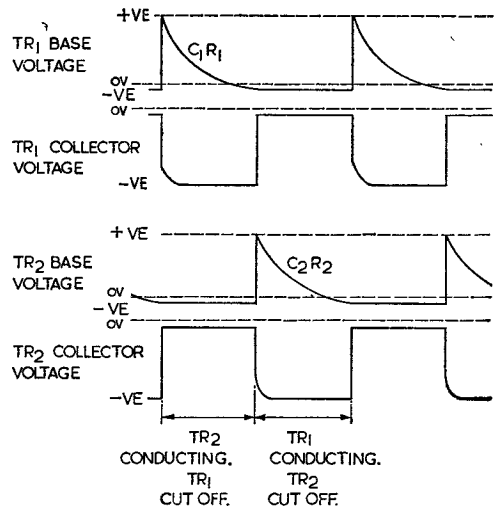

FIG. 21. BASIC TRANSISTOR MULTIVIBRATOR

FIG. 22. TRANSISTOR MULTIVIBRATOR WAVEFORMS

167

positive voltage but is falling towards zero volts as C_2 discharges through R_2. When the base voltage of TR_2 is sufficiently negative, TR_2 conducts and a cumulative action results in TR_1 being abruptly cut-off and its collector voltage falls sharply to the negative voltage of the supply. The collector voltage of TR_2 rises sharply towards zero volts. The collector waveforms remain in this condition until, as C_1 discharges through R_1, the base of TR_1 falls sufficiently negative to cause TR_1 to conduct. A further cumulative action then causes TR_1 to conduct fully and TR_2 to cut off. C_2 then discharges through R_2 until TR_2 conducts, cutting TR_1 off and completing the cycle.

The voltage rise at the collector when the transistor is cut on is sharper than the voltage fall when the transistor is cut off. One reason for this is the presence of capacitors C_1 and C_2 in the collector circuits. Another cause is peculiar to transistors which have been driven hard into conduction so that the collector-emitter voltage is almost zero. When this happens, the emitter injects more holes into the base region than are required to give the collector current; the excess holes are stored in the base ready to be swept into the collector to prolong the collector current for a short period after the emitter current has been cut off. This effect is known as *hole storage*.

Another result of high current conduction in common emitter circuits is that the transistor readily "bottoms". This is because the transistor output characteristic curves are similar to those of a pentode.

The Flip-flop

A flip-flop is a form of multivibrator which can be held in a stable state until an input trigger pulse is applied. It then completes one cycle of oscillation and stays stable until the next trigger pulse. It is sometimes known as a *monostable* multivibrator.

The circuit of an anode-coupled flip-flop is shown in Fig. 23a. There is an important difference between this circuit and that of a free-running (or *astable*) multivibrator: the grid of V_1 is taken to a negative bias voltage. Thus in this circuit V_1 cannot conduct and normal multivibrator oscillations cannot start. V_2 is, however, conducting.

When a positive trigger pulse, of sufficient amplitude to overcome the bias, is applied to V_1 grid (Fig. 23b), V_1 is driven into conduction and a cumulative action occurs which results in V_2
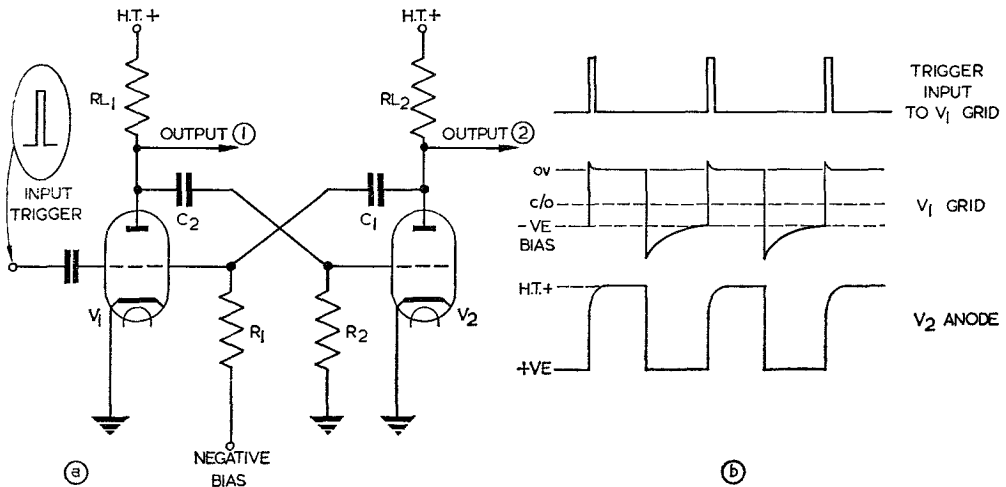

FIG. 23. ANODE-COUPLED FLIP-FLOP CIRCUIT AND WAVEFORMS

grid being driven below cut-off. V_2 grid voltage then rises towards cut-on on a time constant of $C_2 R_2$. At cut-on a second cumulative action occurs and V_1 grid is now driven below cut-off and starts to rise on a time constant of $C_1 R_1$. By this time, however, the trigger pulse has finished and V_1 grid cannot rise above the bias voltage until another trigger pulse occurs.

The output taken from V_2 anode (Fig. 23b) is one complete square wave for each trigger pulse; an antiphase square wave is available at V_1 anode. The frequency of the output waveform is thus decided by the frequency of the input triggering pulses. The circuit could also be triggered by applying a *negative* pulse to V_2 grid.

In the transistor collector-coupled flip-flop shown in Fig. 24 TR_1 collector is directly coupled to TR_2 base via R_1 and TR_2 collector is capacitively coupled to TR_1 base via C_1. The stable state in this circuit is with TR_1 conducting and TR_2 cut off by a positive bias on the base. The positive trigger pulse applied to TR_1 base switches TR_1 off and TR_2 on. The collector of TR_2 rises sharply to zero volts, taking TR_1 base positive, thus cutting off TR_1. C_1 now discharges through R_2 and the circuit switches back when the base-emitter voltage of TR_1 is approximately zero. This state is held until the next trigger pulse.

The switching time can be reduced by shunting R_1 with a capacitor C_2. The duration of the first part of the output cycle is decided by the time constant $C_1 R_2$ and the frequency of the output is that of the triggering pulses.

FIG. 24. COLLECTOR-COUPLED FLIP-FLOP

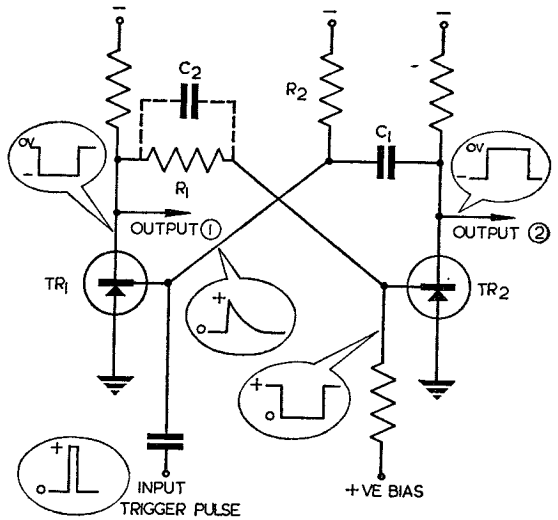## Cathode-coupled Flip-flop

Another circuit arrangement for a flip-flop is shown in Fig. 25: the resistor R_k is common to the cathode circuits of both V_1 and V_2 and is known as a *cathode-coupling* resistor.

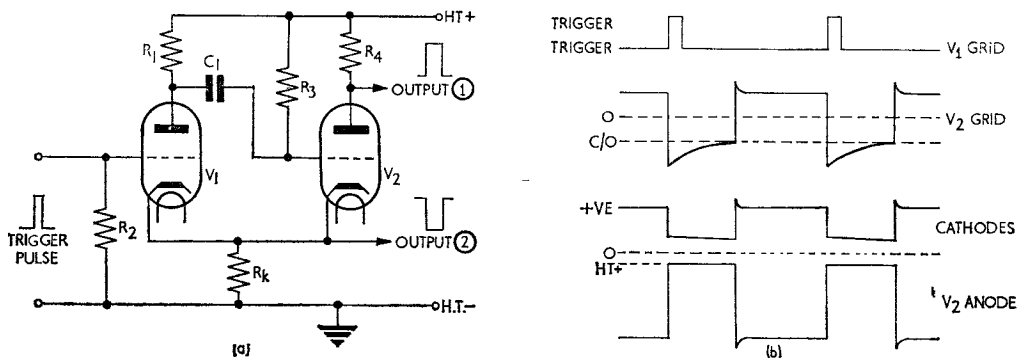The circuit's stable state is with V_2 conducting: the current through R_k produces a voltage

FIG. 25. CATHODE-COUPLED FLIP-FLOP

169   *(A.L. 3, September, 1964)*

which makes both cathodes positive with respect to earth. Since R_2 is returned to earth the grid of V_1 is negative with respect to its cathode, and V_1 is not conducting.

When a positive trigger pulse of sufficient amplitude is applied to V_1 grid, V_1 conducts. Its anode voltage therefore falls and V_2 grid voltage falls with it (Fig. 25b). This causes the current through R_k to fall, thus reducing the cathode voltage. This reduced bias on V_1 causes more anode current to flow and a cumulative action results in V_2 being cut off and V_1 conducting fully.

C_1 now discharges on a time constant of approximately C_1 R_3 and V_2 grid voltage rises towards cut-on. At cut-on V_2 conducts and starts another cumulative action which results in V_1 being cut off and V_2 conducting fully. This stable state is held until the next trigger pulse is applied.

The duration of the positive pulse taken as output from V_2 anode depends upon the discharge time of C_1. This time constant is actually $C_1 (R_3 + r)$ where r is the resistance of V_1 and R_k. In practice as R_3 is about 1 megohm and r is only a few kilohms, r can be ignored.

The frequency of the output is the same as that of the input triggering pulses. An output which is in antiphase to that taken from V_2 anode is available at V_2 cathode.

Bistable Multivibrator

The bistable multivibrator, sometimes called an Eccles-Jordan circuit, has *two* stable states and changes from one state to the other only when a trigger pulse is applied. In the circuit shown in Fig. 26a TR$_1$ collector is directly coupled to TR$_2$ base via R_1, and TR$_2$ collector is in turn directly coupled to TR$_1$ base via R_2. The bases of both transistors are taken to a positive bias.
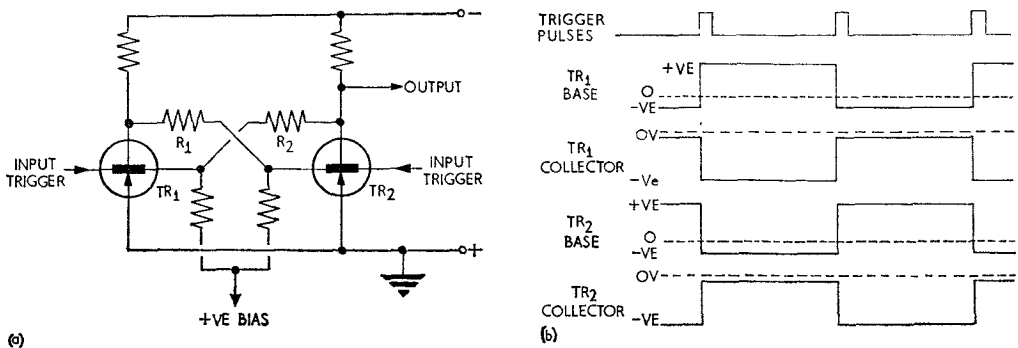

FIG. 26. BISTABLE MULTIVIBRATOR

The two stable states are (a) TR$_1$ conducting and TR$_2$ cut off: (b) TR$_2$ conducting and TR$_1$ cut off. If valves were used in place of transistors, the circuit would be similar but the grids would be taken to a negative bias.

If TR$_1$ is conducting and TR$_2$ cut off, a positive trigger pulse on TR$_1$ base reduces TR$_1$ collector current and TR$_1$ collector voltage goes negative (Fig. 26b). This negative voltage is coupled to TR$_2$ base and TR$_2$ conducts; TR$_2$ collector voltage rises towards zero volts taking TR$_1$ base voltage with it. A cumulative action occurs, resulting in TR$_1$ being cut off and TR$_2$ conducting fully.

This stable state is held until a positive pulse applied to TR$_2$ base switches TR$_2$ off and TR$_1$ on. Thus the bistable circuit produces *one complete cycle* of square wave for every *two* trigger pulses. The frequency of the output is therefore half that of the input. The output is normally

taken from TR_2 collector. Small capacitors across R_1 and R_2 make the edges of the output waveform steeper.

The Schmitt Trigger Circuit

The Schmitt trigger circuit is a cathode-coupled form of bistable multivibrator. It produces a well-shaped square wave from very imperfect input waveforms.

A basic Schmitt trigger circuit is shown in Fig. 27a and two possible input waveforms and an output waveform are shown in Fig. 27b. If V_1 is cut off and V_2 is conducting, a rising voltage
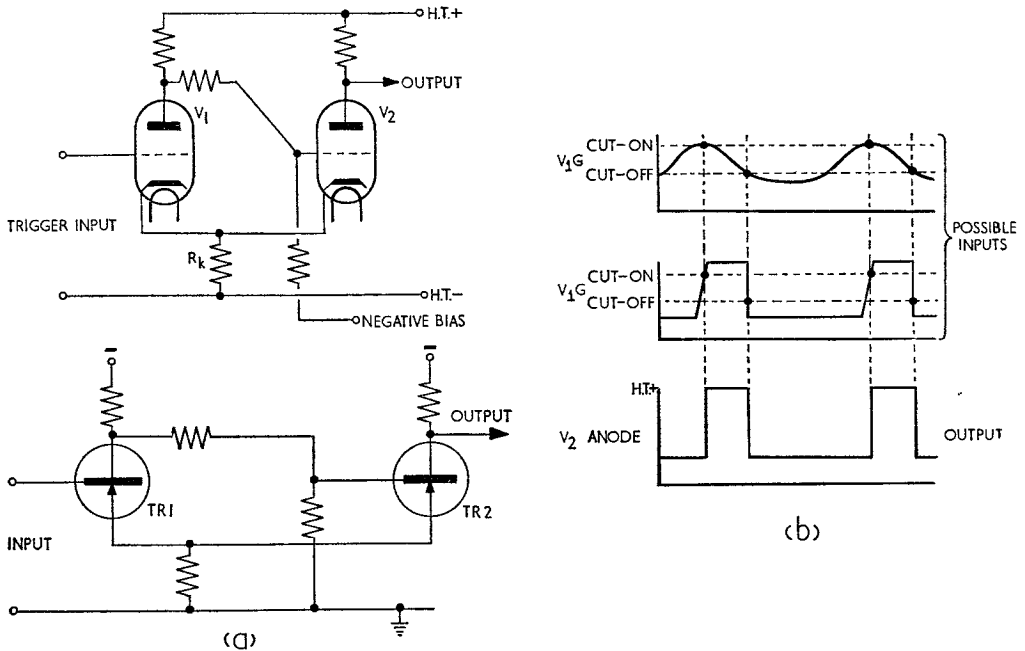
FIG. 27. SCHMITT TRIGGER CIRCUIT AND WAVEFORMS

applied to V_1 grid causes V_1 to conduct. Due to feedback via R_k, a cumulative action occurs; V_2 grid is driven below cut-off and held there by the negative bias. V_1 now conducts and a smaller bias voltage is developed across R_k.

This state is maintained until the input voltage takes V_1 grid below cut-off, when a further cumulative action results in V_1 being cut off and V_2 conducting. Because of the different voltages across R_k, V_1 does not cut off until its grid reaches a lower voltage than the cut-on value.

The Blocking Oscillator

The blocking oscillator is a pulse generating circuit which produces very narrow pulses. The basic circuit (Fig. 28) is similar to that of an inductively-coupled LC oscillator. In such a circuit, if the feedback between anode and grid is great enough and the CR time constant of the bias circuit is long enough, the charge which builds up on C when grid current flows during the
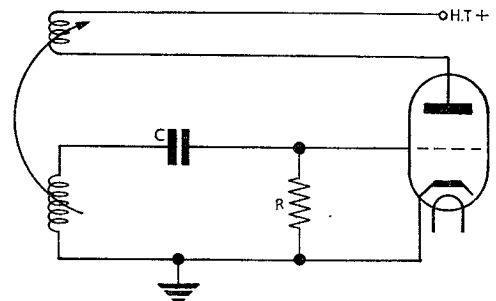
FIG. 28. BASIC BLOCKING OSCILLATOR

171

positive half-cycles of grid voltage is sufficient to bias the valve well beyond cut-off.   Oscillations then cease until C has discharged to cut-on voltage.   As a result, oscillations occur in "bursts"; this is known as *self-quenching* or *squegging*.

The blocking oscillator is made to "squegg" so that it cuts off after the *first half-cycle* of oscillation.   The output is then a single voltage pulse.

The circuit and waveforms of a *free-running* blocking oscillator are shown in Fig. 29.   In this circuit the oscillator is allowed to squegg all the time.   At the end of the first half-cycle of
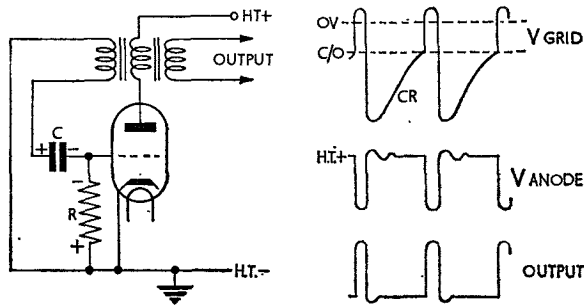

FIG. 29.   FREE-RUNNING BLOCKING OSCILLATOR

oscillation the grid is biased well below cut-off by the voltage to which C has charged and oscillation ceases.   The grid voltage then rises towards cut-on as C discharges through R and after a time—dependent on the time constant CR—reaches cut-on and oscillation recommences.

The anode waveform consists of a series of narrow pulses followed by damped oscillatory swings due to "ringing" in the anode coil.   The *duration* of the pulse depends upon the inductance and self-capacitance of the transformer.   The interval between pulses depends mainly on the grid CR time constant.   The output is often taken from a third winding on the transformer and either a positive or a negative pulse can be obtained, depending on the connections.   The damped oscillations due to ringing can be removed by limiting.

A transistor *triggered* blocking oscillator circuit is shown in Fig. 30.   Transformer coupling between collector and emitter provides feedback and the transistor is biased beyond cut-off by R_1 and R_2.

Positive trigger pulses are applied to the emitter via C_1, thus raising the base-emitter voltage above the bias level.   The transistor conducts and the rising collector current through L_1 induces
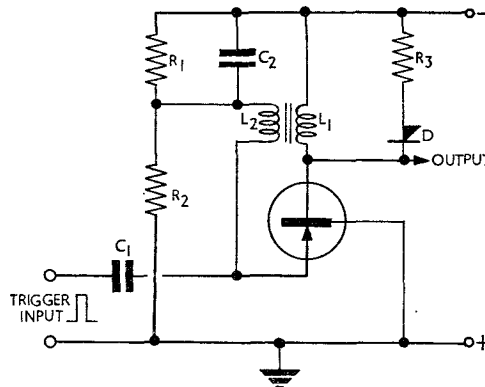

FIG. 30.   TRIGGERED BLOCKING OSCILLATOR

172

a voltage into L_2 which causes the collector current to *increase*. When the emitter current is no longer sufficient to maintain the collector current, the voltage across L_1 falls and induces a voltage into L_2 which tends to *reduce* the collector current. This causes a cumulative action which results in the transistor being rapidly cut off. C_2, which charged negatively during the conducting period, now discharges through R_2 and the emitter voltage rises towards the bias level. The next input trigger pulse starts another cycle.

The output is taken from the collector and, to prevent ringing, a diode D and damping resistor R_3 are placed across L_1.

Pulse Amplification

For the purpose of amplification, narrow pulses can be considered as video signals. Video amplifiers have been considered in Part 1 of these notes and all the points concerning video amplifiers apply to pulse amplifiers. The important features of a pulse amplifier are summarized as follows:—

a. Pentode valves with small interelectrode capacitances, or drift transistors with short transit times, are used as the amplifying devices.

b. The load resistors used are of low value so that a steep leading edge is preserved.

c. High and low frequency compensating networks are employed.

d. Negative feedback is widely used.

The output from a pulse amplifier is often fed into a cathode follower circuit in order to match the impedance of the amplifier to that of the line.

The Rectifier as an Isolating Device

When two or more inputs are applied to the same point in a circuit, some device is required which will isolate the inputs from each other, otherwise current from the largest amplitude input will flow, not only through the load, but also through the other input circuits and cause interaction between the inputs. This is illustrated in Fig. 31a: battery 1 sends a current through the load resistor and also sends a loop current through battery 2.
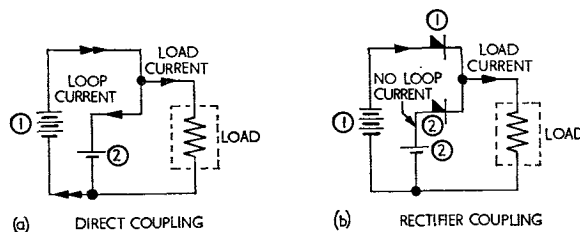
FIG. 31. ISOLATION OF INPUTS

The device used to isolate the inputs from each other is a rectifier (thermionic or semiconductor diode), connected as shown in Fig. 31b. Rectifiers 1 and 2 act as switches operated by the amplitude and polarity of the input voltages. When an input voltage of one polarity is applied to the rectifier, the rectifier acts as a low resistance, and current flows. When the polarity of the input voltage is reversed, the rectifier acts as an open switch. Thus in Fig. 31b current from battery 1 can flow only through the load. No loop current can flow through battery 2 since it is blocked by rectifier 2.

(A.L. 3, September, 1964)

Fig. 32*a* shows two inputs fed to a common load through two back-to-back connected rectifiers.   In Fig. 32*b* input 1 is much larger than input 2; rectifier 1 is therefore a low resistance and current due to input 1 flows through the load.   The voltage developed across the load biases
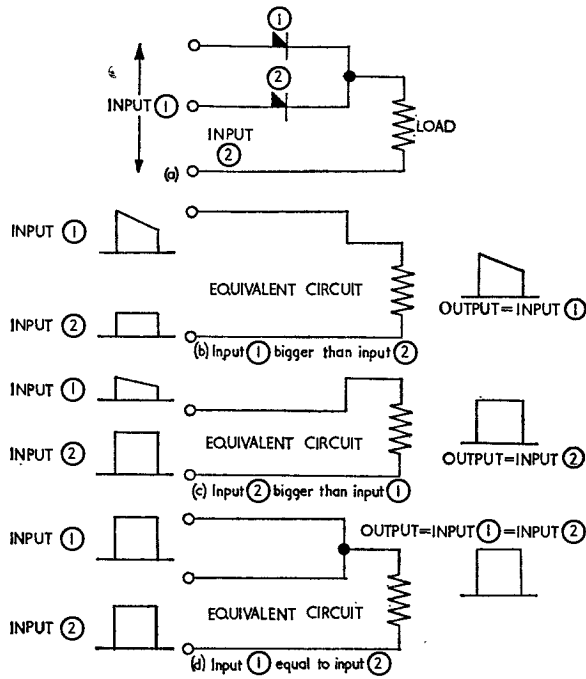

FIG. 32.   AUTOMATIC SELECTION OF INPUTS

rectifier 2 such that it is a high resistance and it does not conduct.   The equivalent circuit is as shown and the output is almost equal in amplitude to input 1.

In Fig. 32*c* input 2 is greater than input 1; rectifier 2 conducts and rectifier 1 is cut off.   The output is now almost equal to input 2.

In Fig. 32*d* both inputs are of equal amplitude and the effect is that the two inputs are connected in parallel in such a way that no loop current flows from source of input to the other.

A typical example of the use of input isolating rectifiers is in the selection of an input trigger pulse from two unstabilised sources (Fig. 33).
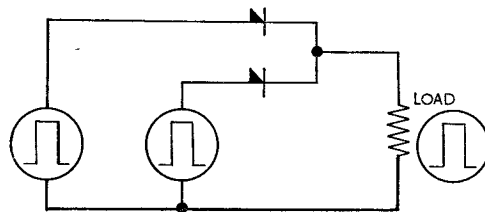

FIG. 33.   USE OF ISOLATING RECTIFIERS

In such an arrangement the output is always of the maximum possible amplitude and interaction between the sources is avoided.

CHAPTER 4

AUTOMATIC RADIO TELEGRAPH SYSTEMS

Introduction

All high-speed automatic telegraph machinery relies on the *synchronous* rotation of motor-driven shafts at send and receive ends. If each signal is made to stop and start the motors, the receive instrument can instantly reproduce the transmitter modulation. This can take the form of code (morse tape) or normal printing by morse printer.

Automatic Morse Telegraphy

Fig. 1 shows in block form the arrangement for automatic morse transmission by radio. The written message is changed into morse symbols by perforating a paper tape with a *morse keyboard perforator*. The perforated tape is then fed into a *Wheatstone transmitter* which
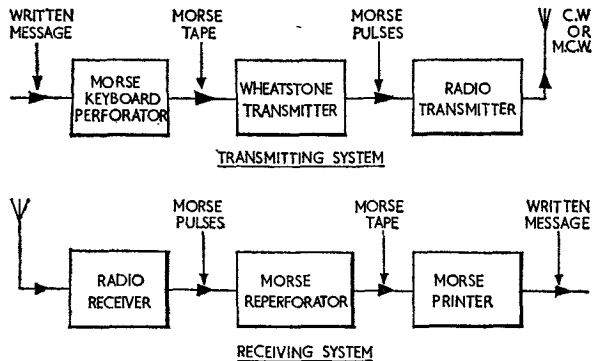
FIG. I. AUTOMATIC MORSE SYSTEM

gives an output of positive and negative voltage pulses. These are used to modulate a radio transmitter. At the receiving end the morse code pulses from a radio receiver are changed into perforations on a tape. The tape is then fed to a morse printer which reproduces the original written message.

Automatic Morse Tape

In order that the dots and dashes of the morse code can operate a Wheatstone transmitter they are represented by holes perforated in a paper strip. A spool of unperforated paper tape is fed into the selector mechanism of a morse keyboard perforating machine. The keyboard of this machine is similar to that of a typewriter; when a key is pressed a selector bar operates the selector mechanism and mark and space holes representing the appropriate character of the morse code are punched into the tape. Central feed holes are also punched into the tape. A piece of perforated tape is shown in Fig. 2.
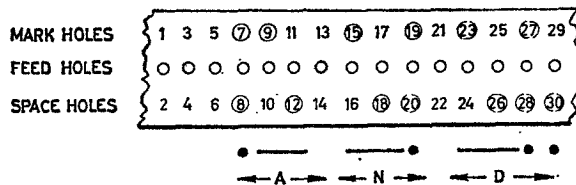
FIG. 2. AUTOMATIC MORSE TAPE

Wheatstone Transmitter

This machine is not a radio transmitter but rather the morse key which operates an automatic high speed radio transmitter. The main part of the machine consists of a contact assembly, the armature of which is actuated by push-rods. Fig. 3 shows that the contact assembly is in effect a light morse key connected for double current working.

The contact spring, which corresponds to the bar of a morse key, is pivoted at the centre and either the mark or the space contact is closed by the movement of one or other of two horizontal push-rods. These push-rods are actuated through a bell crank system by two slender vertical rods known as "peckers". The peckers, one for mark and one for space, move up and down in a reciprocating manner against the morse tape (Fig. 4a).
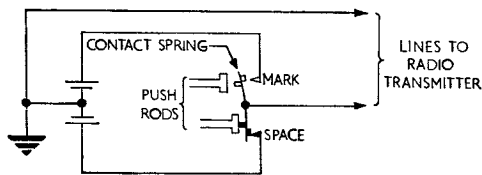
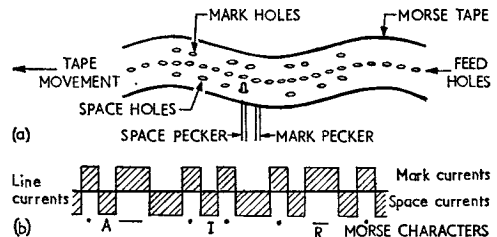FIG. 3. SIMPLIFIED CIRCUIT OF CONTACT ASSEMBLY

FIG. 4. PECKER OPERATION

Assuming the space contact is made, then if both peckers encounter paper there is no movement of the push-rods and the space contact remains closed. The peckers continue to rise and fall and as a mark hole approaches, the mark pecker (which is positioned slightly to the right of the space pecker) will rise through the hole. This extra rise is transmitted through the bellcrank lever and the mark push-rod closes the mark contact and opens the space contact on the contact assembly.

Thus a marking current is sent down the lines and the mark contact remains closed until the space pecker rises and enters a space hole; then the mark contact is broken and the space contact made, thus starting a space current. Mark and space currents for the word AIR are shown in Fig. 4b.

In this way mark and space holes in the morse tape are translated into mark and space currents which operate the radio transmitter keying circuit.

The Morse Reperforator

This machine is used at the receiving end of a circuit and produces a tape perforated in accordance with the received signals. The tape is therefore a replica of the original tape passed through the Wheatstone transmitter.

Unperforated tape is fed into the machine and the output voltage pulses from the radio receiver, representing dots and dashes of the morse code, are used to operate an electro-magnet in the reperforator. This electro-magnet controls the perforating mechanism and holes, representing the transmitted characters of the morse code, are punched into the paper tape.

Teleprinter Systems

A teleprinter is in effect an electrically operated typewriter with the keyboard operating a remote printing mechanism. Fig. 5 outlines the teleprinter system and compares it with a normal type-written letter system.
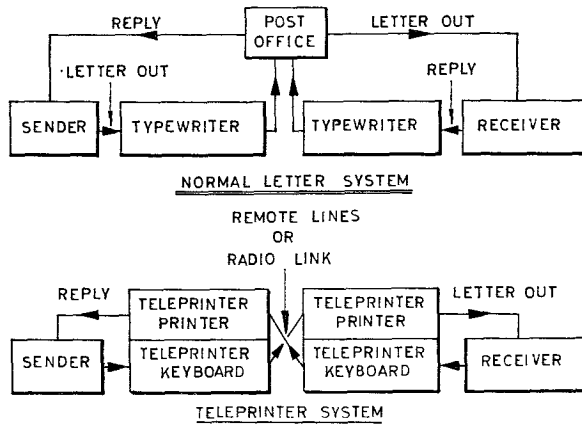
FIG. 5. COMPARISON OF TELEPRINTER AND LETTER SYSTEMS

In its simplest form the teleprinter link would need as many lines between keyboard and remote printer as there are keys on the keyboard. Operation of a particular key would then energise a particular line and the appropriate character would be printed at the remote receiver.

The mechanism of a practical teleprinter enables the machine to respond to the Murray code and to select the appropriate character by the operation of five motor-driven cams. In order to start each cycle of movement in phase with the remote (transmitter) keyboard, the receive motor is automatically started and stopped by each transmission. These 'start' and 'stop' impulses are added to the basic five unit code to produce the $7\frac{1}{2}$ unit code shown in Fig. 6.

Notice that the stop signal is $1\frac{1}{2}$ units long. This ensures that the receive operation is completed before the next character arrives, in spite of random variations in the time of operation. For the same reason the receiver cam shaft rotates at a slightly higher speed than that of the transmitter cam shaft, i.e. cam operation at the receiving end is slightly faster than at the transmitting end.
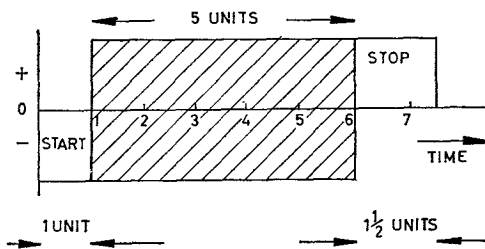

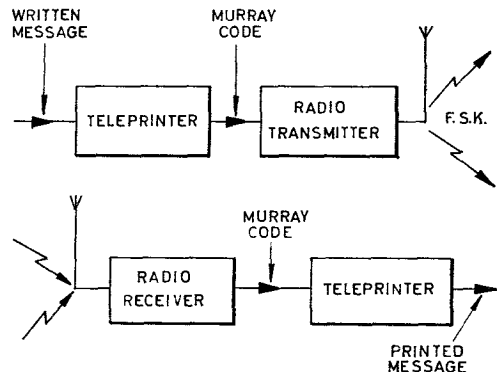FIG. 6. THE $7\frac{1}{2}$ UNIT CODE

FIG. 7. RADIO TELEPRINTER SYSTEM

Radio Teleprinter System

The arrangement for automatic *telegraph printing* by radio is outlined in Fig. 7.

The message is first translated into Murray code by the teleprinter and the output pulses are used to modulate the radio transmitter.

At the receiving end the reverse process occurs. The a.f. output of the radio receiver is made to operate the receive teleprinter and so the original written message is reproduced.

A landline could replace the radio link where this is more convenient. In this case, the teleprinters work directly into each other and merely form remotely operated keyboards.

Voice Frequency (VF) Telegraphy

In the voice frequency system of telegraphy, audio (voice) frequency tones are keyed instead of a d.c. supply. The advantage is that *several tones* differing in frequency by a predetermined amount (see Table 1) can be keyed and transmitted as a composite signal on a *common* line and thus several messages can be sent simultaneously over one system. The received signals can be separated by means of filter circuits. The basic arrangement is outlined in Fig. 8.
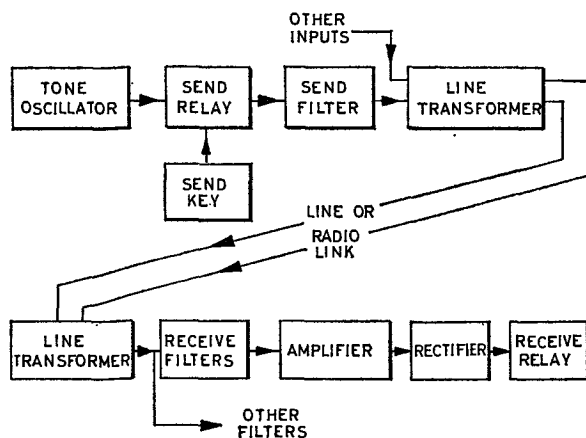

FIG. 8. SINGLE-TONE VF TELEGRAPHY

In practice, the keyed outputs from up to 18 sources are passed on a single system. Thus 18 channels of communication are available on the system and the arrangement is called a *multi-channel-voice-frequency* (m.c.v.f.) system.

The choice of frequencies used on m.c.v.f. systems is determined by the desire to use standard telephone bandwidth of 300 to 3,400 c/s. The frequencies used are all odd harmonics of a basic frequency. Table 1 shows tone frequencies based on odd harmonics of 60 c/s, starting at 420 c/s, with a channel separation of 120 c/s.

CHANNEL	TONE FREQUENCY	CHANNEL	TONE FREQUENCY
1	420 c/s	10	1500 c/s
2	540	11	1620
3	660	12	1740
4	780	13	1860
5	900	14	1980
6	1020	15	2100
7	1140	16	2220
8	1260	17	2340
9	1380	18	2460

TABLE. I. MCVF TONE FREQUENCIES

Notice that the sum of any two tone frequencies cannot produce an odd harmonic, e.g.

$$420 + 540 = 960; \quad \frac{960}{60} = \text{16th harmonic}$$

$$420 + 660 = 1080; \quad \frac{1080}{60} = \text{18th harmonic}$$

$$540 + 660 = 1200; \quad \frac{1200}{60} = \text{20th harmonic}$$

Thus interference due to interaction between tones is avoided.

Some installations use a different basic frequency (e.g., 85 c/s) but channel separation is again based on odd harmonics of this frequency.

Two-tone MCVF Telegraphy

In order to retain the full advantages of double current working, the single-tone m.c.v.f. system has been modified to use one tone for mark and one for space for *each channel*. Table 2 gives the most commonly used mark and space tones, space being 120 c/s higher than mark in each case.

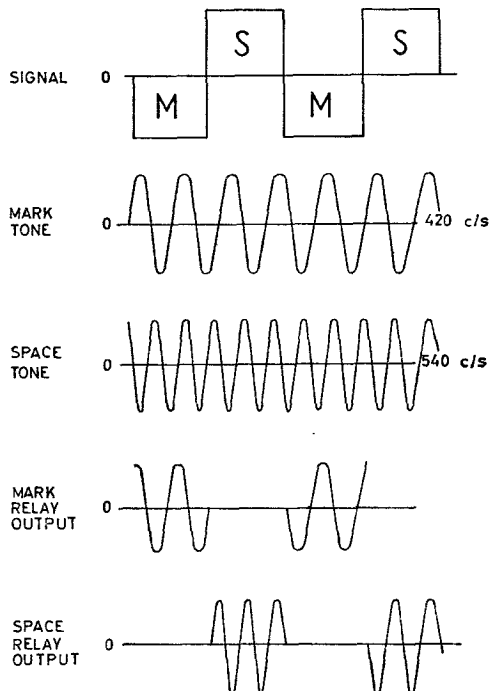| CHANNEL	MARK FREQUENCY	SPACE FREQUENCY
1	420 c/s	540 c/s
2	660 c/s	780 c/s
3	900 c/s	1,020 c/s
4	1,140 c/s	1,250 c/s
5	1,380 c/s	1,500 c/s
6	1,620 c/s	1,740 c/s

TABLE 2. TWO-TONE MCVF FREQUENCIES

FIG. 9. TWO-TONE VF SYSTEM (SEND SIDE)

FIG. 10. TWO-TONE VF WAVEFORMS

The send side of a single channel two-tone m.c.v.f. arrangement is outlined in Fig. 9.

The incoming signal from the teleprinter is a series of positive and negative pulses corresponding to the Murray code characters. Each negative element operates the mark relay to pass a 420 c/s output to a filter; each positive element operates the space relay to pass a 540 c/s output to a filter. The waveforms entering the filters are shown in Fig. 10.

The two send filters restrict the bandwidth of the modulated tone signals by cutting out the extreme sideband frequencies. This results in a rounding off of the output pulses, but the

179

resultant pulse is sufficiently well-shaped for transmission purposes. Fig. 11 shows the frequency spectrum of the signal before and after the mark filter for a 20 millisecond input pulse.
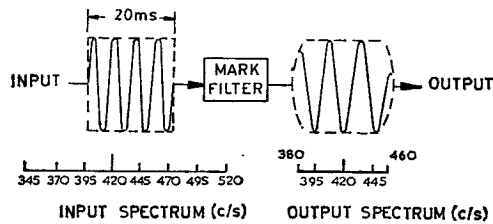

FIG. 11. ACTION OF THE SEND FILTERS

Attenuator Action

The variable attenuator in Fig. 9 is needed in order to restrict the power output to the permissible level regardless of the number of channels being used.

For example, assume the maximum permissible power level for a system is 120 watts and there are six channels available in the system; then if one channel only is being used, the power output from that channel can be maximum, i.e. 120 watts and the attenuator must not attenuate the signal. If all six channels are being used, the attenuator must reduce the power from each channel to 20 watts. If two channels are used, the attenuator must provide 3 db attenuation; and so on.

Receive Side

The receive side of a two-tone m.c.v.f. system is basically the same as the receive side of the single-tone system shown in Fig. 8. In the two-tone system, however, two receive filters are required for each channel, one for mark tones and the other for space tones. In fact, the receive side merely reverses the sequence shown in Fig. 9.
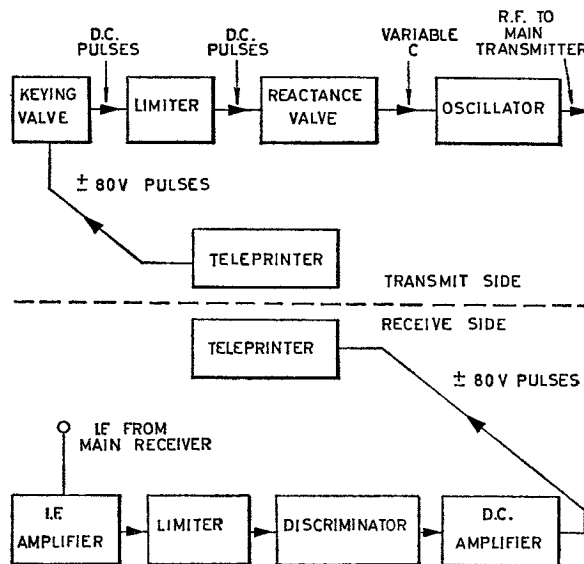

FIG. 12. FSK RADIO TELEPRINTER SYSTEM

180

Frequency Modulated Radio Telegraphy

An alternative radio teleprinter system employs the f.s.k. principles discussed in Sections 2 and 3 (p 77 and p 135). The ± 80V pulse output from the teleprinter, corresponding to the Murray code symbols, is used to frequency modulate the radio transmitter without the use of tones. Fig. 12 outlines the send and receive arrangements.

The ± 80V pulses from the teleprinter operate a keying valve the output from which is fed through a limiter to a reactance valve. This valve controls the carrier frequency generated by the oscillator. Service systems use a frequency deviation of 425 c/s; mark is 425 c/s above the normal carrier frequency and space is 425 c/s below it.

Frequency shift teleprinter systems are usually employed in tributary radio links where traffic is insufficient to warrant m.c.v.f. systems.

As a communication system, f.s.k. has all the advantages of f.m. without an inconveniently large bandwidth. As a telegraph system, it has the additional advantage of being effectively a double current system since there is not a "no current" condition during transmission.

CHAPTER 5

PULSE COMMUNICATION

Introduction

We have so far considered three methods of modulating a carrier in order to transmit voice or code messages. These methods are amplitude, frequency and phase modulation. It is also possible to transmit an audio waveform, for example, by producing pulses which contain information about the waveform and then applying these pulses to a landline, or using them to amplitude modulate a r.f. carrier.

A typical equipment could use pulses of 5 microseconds duration with a pulse recurrence frequency of 8,000 p.p.s. There is thus an interval of 120 μs between pulses. It is possible, therefore, to interleave other trains of pulses, representing other audio channels, in the spaces between the pulses of the first channel.

In Fig. 1, pulses representing channel B are interleaved with those of channel A. With the pulse length and p.r.f. given, even more channels can be interleaved and it is common to have equipments with 8 channels. This process of accommodating several channels by separating them into different time intervals is called *time division mutiplex* (t.d.m.).

Pulse amplitude Modulation

The simplest system of pulse modulation is *pulse amplitude modulation* (p.a.m.). Pulses are applied at regular intervals to the audio waveform to be transmitted and the amplitude of each pulse is made to correspond with the amplitude of the audio wave at that particular instant (Fig. 2).
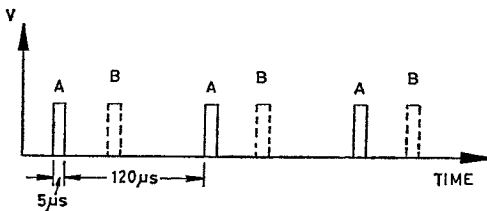

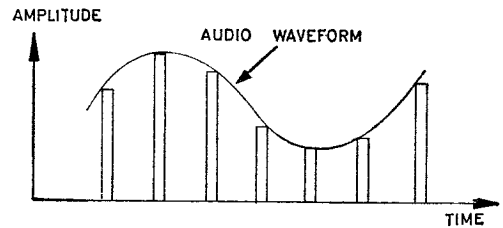FIG. 1. TIME-DIVISION MULTIPLEX

FIG. 2. PRINCIPLE OF PULSE AMPLITUDE MODULATION

The audio waveform is therefore *sampled* by the pulse train at regular intervals. The p.a.m. train can then be used to modulate a r.f. carrier for transmission.

At the receiver, the r.f. signal is demodulated in the usual way to extract the pulses from the r.f. wave and these pulses are then applied to a low-pass filter which extracts the audio signal. Although only a few instants are sampled per cycle, what the ear hears as a result is indistinguishable from the original sound, for the same reason that the eye cannot distinguish a rapid series of still pictures in cinematograph projection from true motion.

How PAM is Produced

The train of pulses can be produced by any of the methods of pulse generation described in Chapter 3, e.g. a blocking oscillator, or a multivibrator followed by a differentiating circuit. If necessary, the pulses are amplified so that they are comparable in amplitude with the audio waveform.

The pulse train and the audio waveform are then applied in series to the grid of a valve (Fig. 3).

The negative grid bias is sufficient to keep the valve cut off when the audio signal alone is applied, and the pulses applied are of sufficient amplitude to cause the valve to conduct slightly
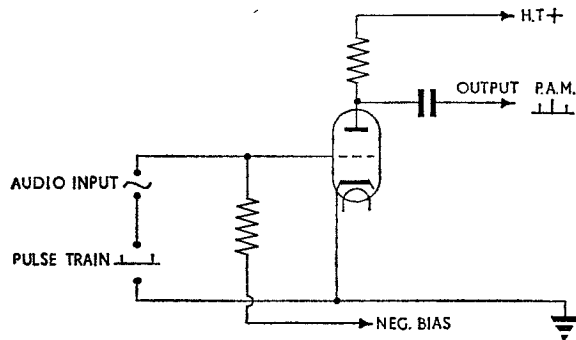
FIG. 3.   PRODUCTION OF PAM

when they are added to the lowest value of a.f. waveform.   The amount of anode current then depends upon the instantaneous value of the pulse plus that of the a.f. waveform.   As long as the valve works on the linear part of its characteristic curve, the output will be a train of pulses whose amplitudes are proportional to those of the audio waveform at the sampling instants. This is illustrated in Fig. 4.

FIG. 4.   WAVEFORMS PRODUCING PAM

The output voltage pulses would, in this case, be negative-going but they could be inverted by using another amplifier.

183

PAM Transmitter

The essential parts of a p.a.m. transmitter are shown in block form in Fig. 5.

FIG. 5. BASIC PAM TRANSMITTER
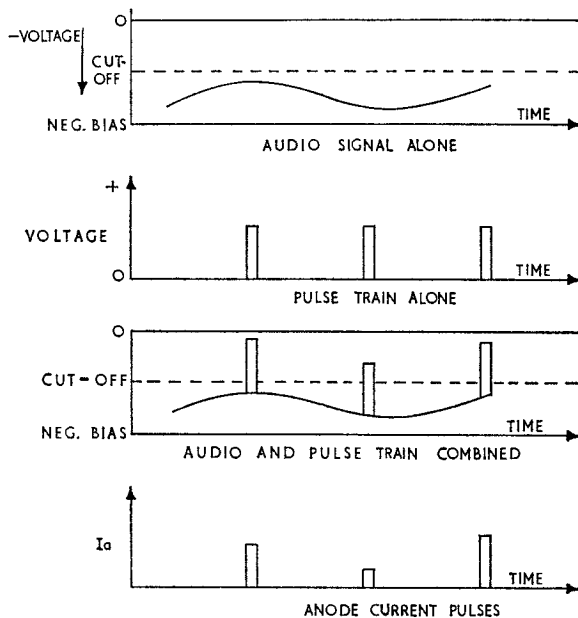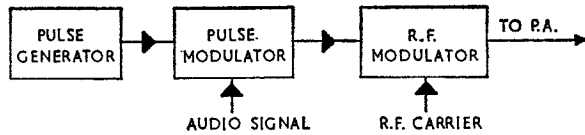
PAM Receiver

The p.a.m. receiver is a conventional superhet; the i.f. signal is demodulated to give the p.a.m. train of pulses.

Because the amplitude of each pulse varies, the mean signal level also varies and does so roughly according to the original audio waveform.   A low-pass filter therefore passes the mean level (audio) variation but rejects the higher frequency components of the pulse waveform (Fig. 6).
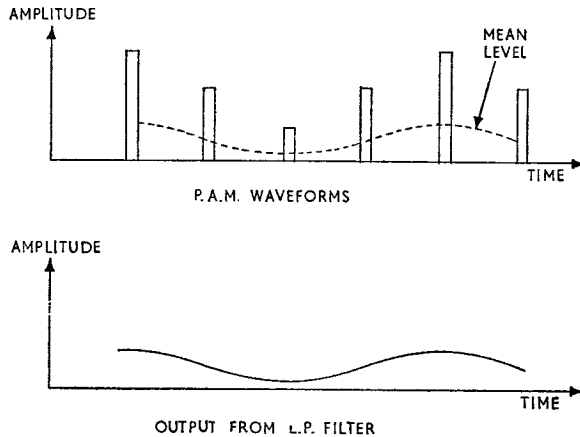
FIG. 6.   RECOVERY OF AUDIO WAVEFORM FROM PAM

The peak amplitude of the mean variations is small and considerable a.f. amplification is needed after recovery of the audio waveform.   A block diagram of a p.a.m. receiver is shown in Fig. 7.
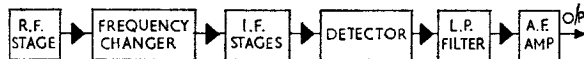
FIG. 7. BASIC PAM RECEIVER

Pulse-amplitude modulation is one of four methods of pulse modulation.   It is the simplest but unfortunately the least efficient method.   The three other methods in common use are *pulse duration modulation* (p.d.m.), *pulse position modulation* (p.p.m.) and *pulse code modulation* (p.c.m.).

Pulse Duration Modulation

In p.d.m. the duration of each pulse varies in accordance with the instantaneous *amplitude* of the audio waveform.   The method is illustrated in Fig. 8.   It is sometimes called *pulse-width* or *pulse-length* modulation.

184

A big advantage of p.d.m. over p.a.m. is that most of the noise present in the received signal can be eliminated. A "slice" of the received pulse is selected by a limiting circuit using diodes and amplified to remove most of the noise (Fig. 9). This method of noise reduction cannot be used with p.a.m. but can be applied to the other methods of pulse communication.
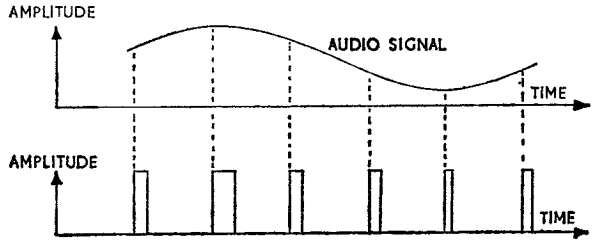
FIG. 8. PULSE DURATION MODULATION
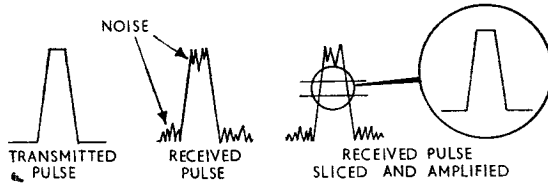
FIG. 9. REDUCTION OF NOISE

A circuit suitable for producing p.d.m. is shown in Fig. 10.
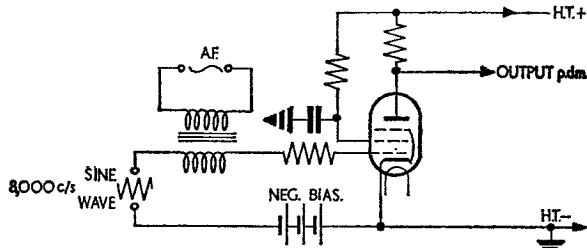
FIG. 10. SINE WAVE MODULATION FOR PDM
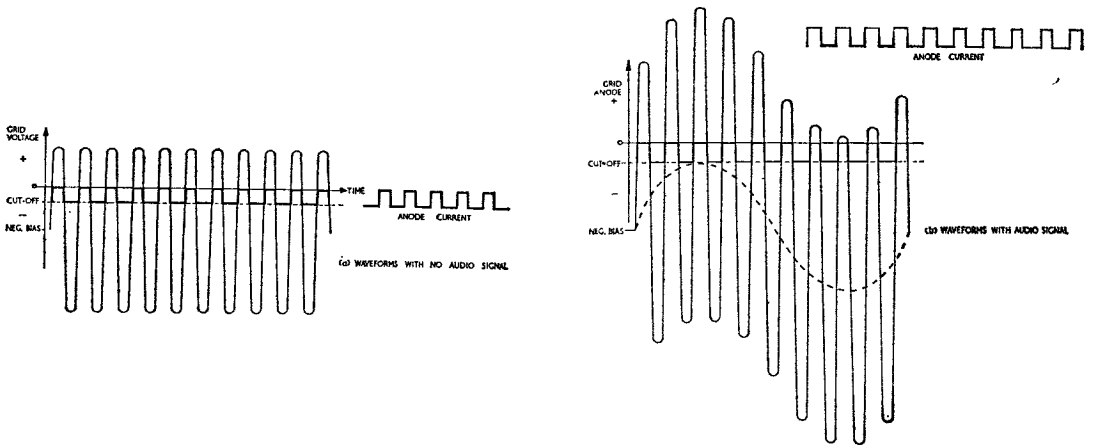
FIG. 11. PRODUCTION OF PDM WAVEFORMS

The valve is used as a squarer and so a pentode with a short grid base is chosen. Large amplitude 8,000 c/s sine waves are applied to the grid. The resistor in series with the grid limits the positive half-cycles of the 8,000 c/s wave and grid cut-off limits the negative half-cycles; this is normal pentode squarer action (Fig. 11a).

185

An a.f. signal whose maximum frequency is lower than 8,000 c/s is also fed to the grid; thus the portion of the 8,000 c/s wave that is limited will vary with the varying grid bias produced by the a.f. signal, and the resultant pulses will be duration-modulated. Linear modulation is possible only over a small range of audio amplitude but this is no disadvantage in pulse systems.

Another method of producing p.d.m. is by combining the audio waveform with the output of a sawtooth generator (Fig. 12). When the amplitude of the sawtooth waveform rises to that of the audio waveform a pulse generator is switched on. It is switched off when the sawtooth amplitude falls below that of the audio wave.

FIG. 12. SAWTOOTH MODULATOR WAVEFORMS

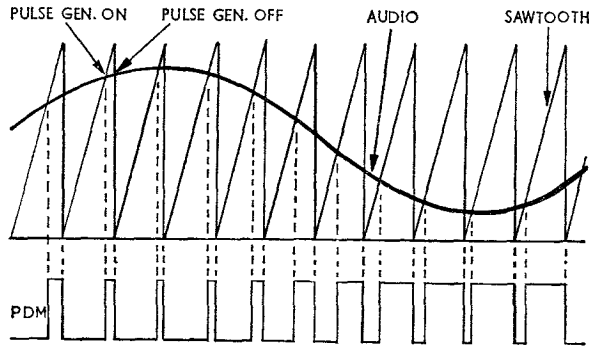The received p.d.m. signal contains a replica of the original audio signal, so it is merely necessary to pass the detected train of pulses through a low-pass filter to obtain the original audio waveform.

Synchronisation and Channel Separation

When time division multiplex transmission is used, synchronising pulses are transmitted as well as the channel pulses. These synchronising pulses are used in the receiver to trigger timing circuits so that transmission and reception are kept in step; and also to indicate the channels. If a synchronising pulse always occurs just before a channel A pulse, for example, it enables the receiver to identify all channel A pulses; other channels can then also be identified.

The synchronising pulses are distinguished from channel pulses by making them either much larger in amplitude, or of much longer duration. Sometimes two or more short pulses in rapid succession are used.

At the receiver, the various channels have to be separated after they have been identified. A *gating circuit* is often used to do this. All channel pulses and the gate pulses for the channel concerned are applied in series to the grid of a valve which is biased such that an output occurs only when both pulses are present.

Fig. 13 illustrates the process applied to p.d.m. with three channels only, although in practice eight or more channels are common. The gate pulses are arranged to be slightly longer than the maximum duration of the channel pulses. With three channels, three separate gating valves are needed; with eight channels, eight gating valves would be required.

Pulse Position Modulation

In p.d.m. all the necessary information concerning the audio waveform is contained in the position of the trailing edge of the pulse. If that position is transmitted then, in conjunction with
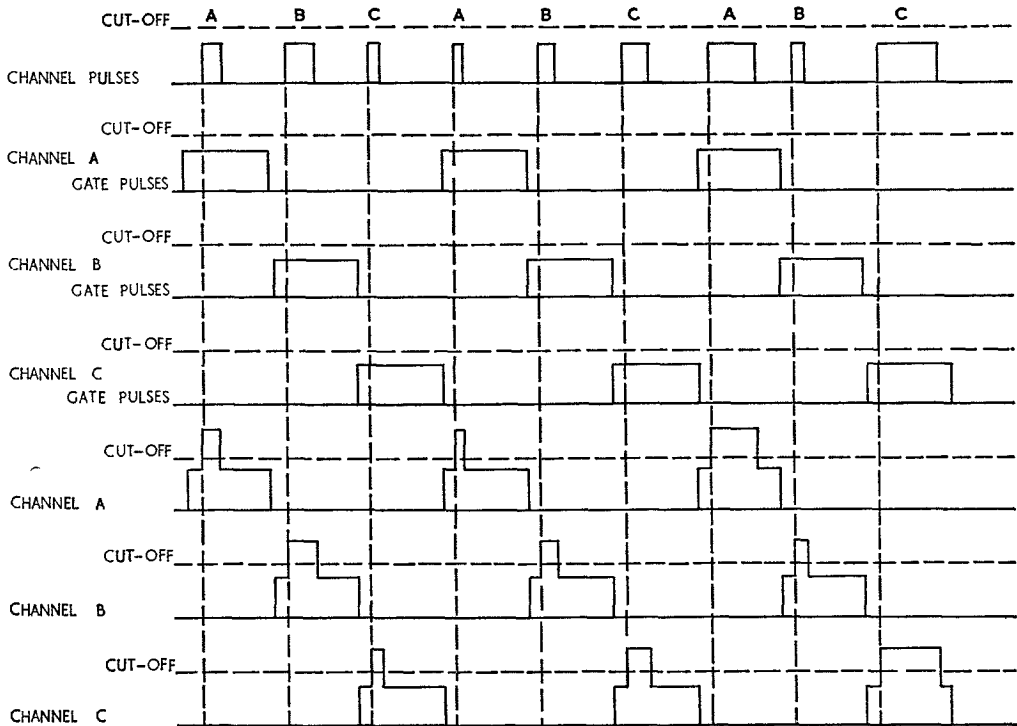
FIG. 13. CHANNEL SEPARATION

the synchronising pulses, reliable communication can be achieved. *Pulse position modulation* systems transmit narrow pulses (e.g. 0.5 microseconds) whose positions vary from the synchronising pulses (Fig. 14). The system is also known as *pulse phase modulation*.
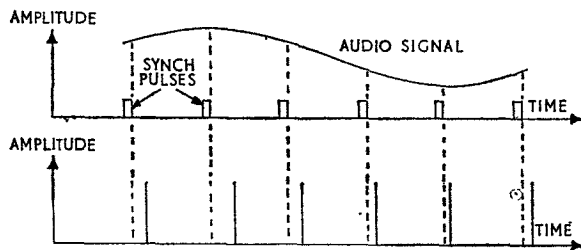
FIG. 14. PULSE POSITION MODULATION

Pulse position modulation has advantages over both p.a.m. and p.d.m. It can be used with "slicing" techniques to reduce noise and it also uses less average transmitter power because only narrow, constant-amplitude pulses are transmitted. The received signal is also less likely to suffer from distortion because all that the receiver needs to detect is the presence of a pulse at the correct time; its duration and amplitude do not matter.

Pulse position modulation can be produced from a p.d.m. train by differentiating each pulse with a short CR circuit and removing the unwanted pips with a limiter (Fig. 15).

187   (*A.L.* 3, *September,* 1964)
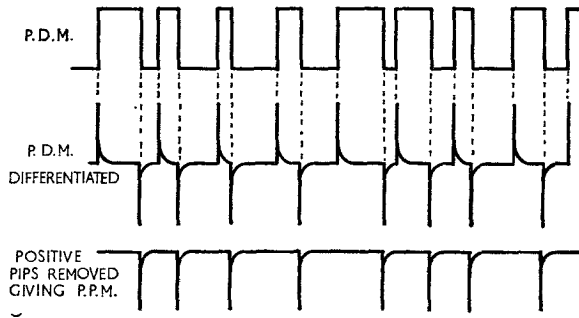
FIG. 15. PRODUCTION OF PPM FROM PDM

At the receiver, the p.p.m. signal must be converted to a p.d.m. form before the audio wave can be extracted using a low-pass filter. One method is to use the synchronising pulses to switch an Eccles-Jordan circuit "on" and the p.p.m. pulses to switch it "off". The output from the Eccles-Jordan circuit is thus a train of p.d.m. pulses corresponding to the original p.d.m. train used in the transmitter. (Fig. 16.)
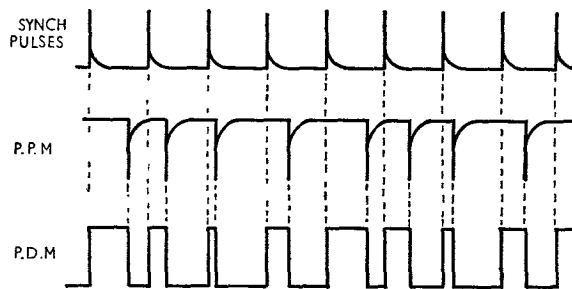
FIG. 16. PRODUCTION OF PDM FROM PPM

Pulse Code Modulation

Despite the many advantages of p.p.m. it is still subject to distortion because the received pulse may be distorted and the instant at which the leading edge occurs may not be clear. This is illustrated in Fig. 17 where the leading edge occurs at instant A, but the instant of receipt is registered as somewhere between B and C.
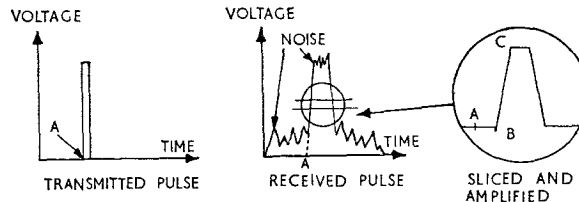
FIG. 17. PPM DISTORTION

Pulse code modulation depends for communication only on whether the receiver detects the presence of a pulse; the time that the pulse commences, its amplitude and its duration do not matter. Thus pulse distortion does not affect the received audio waveform.

In the p.c.m. system, the audio waveform is sampled by a pulse train and its instantaneous values are measured. However, instead of transmitting a p.a.m. signal, the values of the instantaneous amplitudes are measured to the nearest whole number and these values are transmitted as 5 or 7 unit code pulses. It is like reading the value of the Y ordinate on a graph to the nearest small square of graph paper and then transmitting the number. At the receiver, the numbers are transformed back into the original waveform.

A 5 unit code system is shown in Fig. 18. Each instantaneous value, A, B, C, etc. is given a number in 5 unit code and these numbers are transmitted. In Fig. 18 full lines represent pulses and dotted lines the absence of pulses.
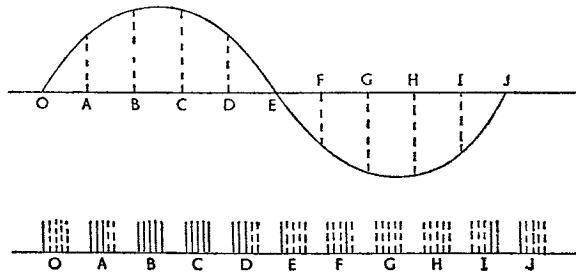
FIG. 18.   PULSE CODE MODULATION

With 5 unit code, 32 different whole numbers or standard values can be selected. With 7 unit code, 128 standard values are possible and the reconstituted waveform is very close indeed to the original.

Typical commercial p.c.m. systems use 0.5 μ sec pulses at a p.r.f. of 100,000 p.p.s. for the 7 unit code. This high p.r.f. enables time division multiplex to be used even though each sampled amplitude requires 7 pulses for coding.

Advantages of Pulse Communication

a. Time division multiplex can be employed to give multichannel communication. An advantage of t.d.m. itself is that there can be little interference between pulses since they occur at different times. Cross-talk between channels is therefore negligible.

b. The equipment required is simpler than that needed for frequency division multiplex systems such as m.c.v.f.

c. Noise can be reduced in all pulse systems, except p.a.m., by using the "slicing" technique.

Disadvantages of Pulse Communication

a. A rectangular pulse contains a large number of harmonics and in order to preserve the steep leading and trailing edges of the pulse, as many harmonics as possible must be transmitted. Therefore pulses occupy a fairly wide band of frequencies. For example, a train of 5 μsec pulses at 8,000 p.p.s. would require a bandwidth of at least 200 kc/s to enable the pulse to be transmitted and received with reasonable fidelity. Narrower pulses require an even wider bandwidth.

b. At the receiver, considerable pulse energy is wasted when the audio signal is extracted. This disadvantage can be partially overcome by good circuit design.

CHAPTER 6
CARRIER TELEPHONY AND MICROWAVE LINKS

Introduction

Carrier telephony is a communication system in which several speech channels may be sent simultaneously over one pair of landlines, or over one radio link, without mutual interference. We have already seen how this is done with telegraph systems (m.c.v.f.); carrier telephony works on the same principles but instead of using tone frequencies to carry the information, radio frequencies, modulated at a.f., are used.

In carrier telephony, transmitting and receiving ends may be linked by landlines or radio or by a combination of both, i.e. part by landline and, where it is more convenient, part by radio. In the radio link, the telephone signals modulate a radio transmitter which beams the signals to a radio receiver. The receiver demodulates and separates the signals and feeds them into the landline system. Radio links operate at frequencies above 900 Mc/s and are called *microwave links*. In this chapter we shall consider the principles of both carrier telephony and microwave links.

How Carrier Telephony Works

The range of speech frequencies covered by a normal telephone channel is 300 to 3,400 c/s, roughly taken as 0—4 kc/s. A pair of wires could carry a.f. currents of this frequency range and so provide communication on one channel.

A second channel may be carried on the same pair of wires by *translating* the speech frequencies so that they occupy a range from 4 to 8 kc/s. A third channel may be added covering 8 to 12 kc/s, and so on.

The process of frequency translation is done by injecting the speech frequencies, along with a suitable carrier frequency, into a non-linear device; a balanced modulator using metal rectifiers is normally used. For example, to translate speech frequencies to the range 8 to 12 kc/s, the speech frequencies would be mixed with an 8 kc/s carrier signal to produce sidebands of 4 to 8 kc/s and 8 to 12 kc/s. The lower sideband could then be removed by a filter leaving the speech channel occupying the band 8 to 12 kc/s.

At the receiving end, the 8 to 12 kc/s channel is separated from the other channels by means of a filter, mixed with a locally-generated 8 kc/s signal in a balanced modulator, and the 0 to 4 kc/s band selected by another filter. Each frequency-translated channel is treated similarly.

A combination of 12 frequency-translated channels is called a *group* and it occupies either the frequency range 12 to 60 kc/s or 60 to 108 kc/s, the latter being the more common.

The group is not generally transmitted directly as a range of frequencies from 60 to 108 kc/s but is used to modulate a carrier of much higher frequency and it is this modulated carrier which is transmitted. Before the channels can be separated at the receiver, the carrier must be demodulated.

This is the carrier telephony system and it has the advantage that landline and radio transmissions may be directly linked without difficulty.

If double sideband amplitude modulation is used, the bandwidth of the modulated carrier will be 2×108 kc/s $= 216$ kc/s, and so the carrier frequency must be high enough to accommodate this bandwidth without occupying a large part of the r.f. spectrum, i.e. the carrier frequency must be of the order of several megacycles per second at least. If even higher carrier frequencies are used, the acceptable bandwidth is much larger and more channels can be accommodated in one transmission. Modern equipment deals with blocks of 600 and 960 channels using a carrier frequency of about 4,000 Mc/s.

12 channels make one *group*; 5 groups, occupying the frequency range 312 to 552 kc/s, is a *supergroup;* and 16 supergroups (960 channels) occupy a *baseband* of 60 to 4028 kc/s.

The formation of groups, supergroups and the baseband is shown in Fig. 1 and Fig. 2. In these illustrations the lower sideband only is used, increasing still further the number of channels that can be handled.
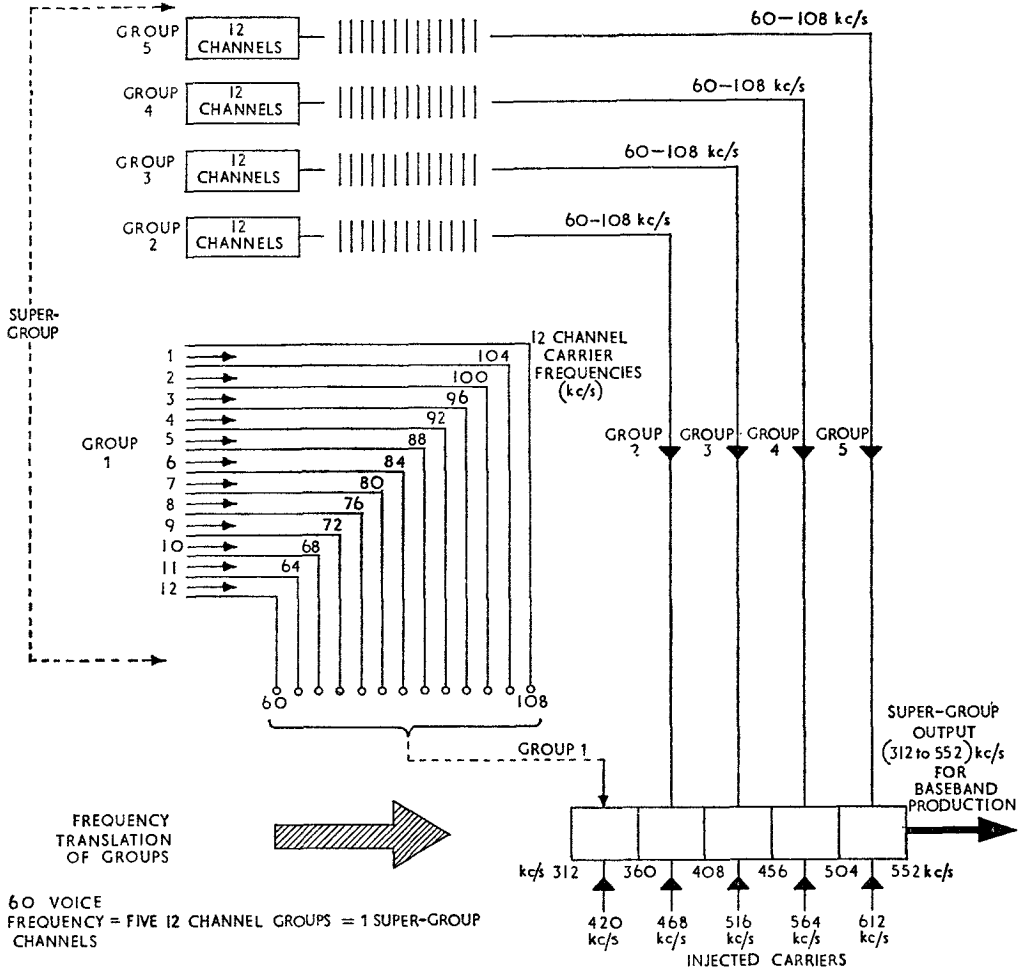
FIG. I.   GROUPS AND SUPERGROUPS

The baseband occupies roughly 0 to 4 Mc/s and hence the bandwidth to be passed by the transmission system is 8 Mc/s if double sideband amplitude modulation is used, or 4 Mc/s if s.s.b. is used. With a carrier frequency of 4,000 Mc/s, frequency modulation is normally used and the bandwidth is then much more than 8 Mc/s.

Some of the channels in the baseband are used for stand-by in case of an emergency failure of one or more channels. They are also used for routine testing and control.

FIG. 2. BASEBAND

192

Microwave Links

Microwave links in the communication network have the following advantages:—

a. They can carry signals across the sea and across country where it would be difficult to lay cables.

b. Many channels can be carried because of the wide bandwidth available.

c. The links can be made mobile—an important point for military applications.

d. At the very short wavelength employed, small, highly-directional aerials can be used with light, low-powered transmitters.

A block diagram of the main stations in a microwave link is shown in Fig. 3. Repeater stations are usually 20 to 50 miles apart and any number may be used. They receive the weak signal from the previous station, amplify it and re-transmit it to the next station. To avoid crosstalk, the received and re-transmitted signals are made to differ slightly in frequency by about 40 Mc/s.
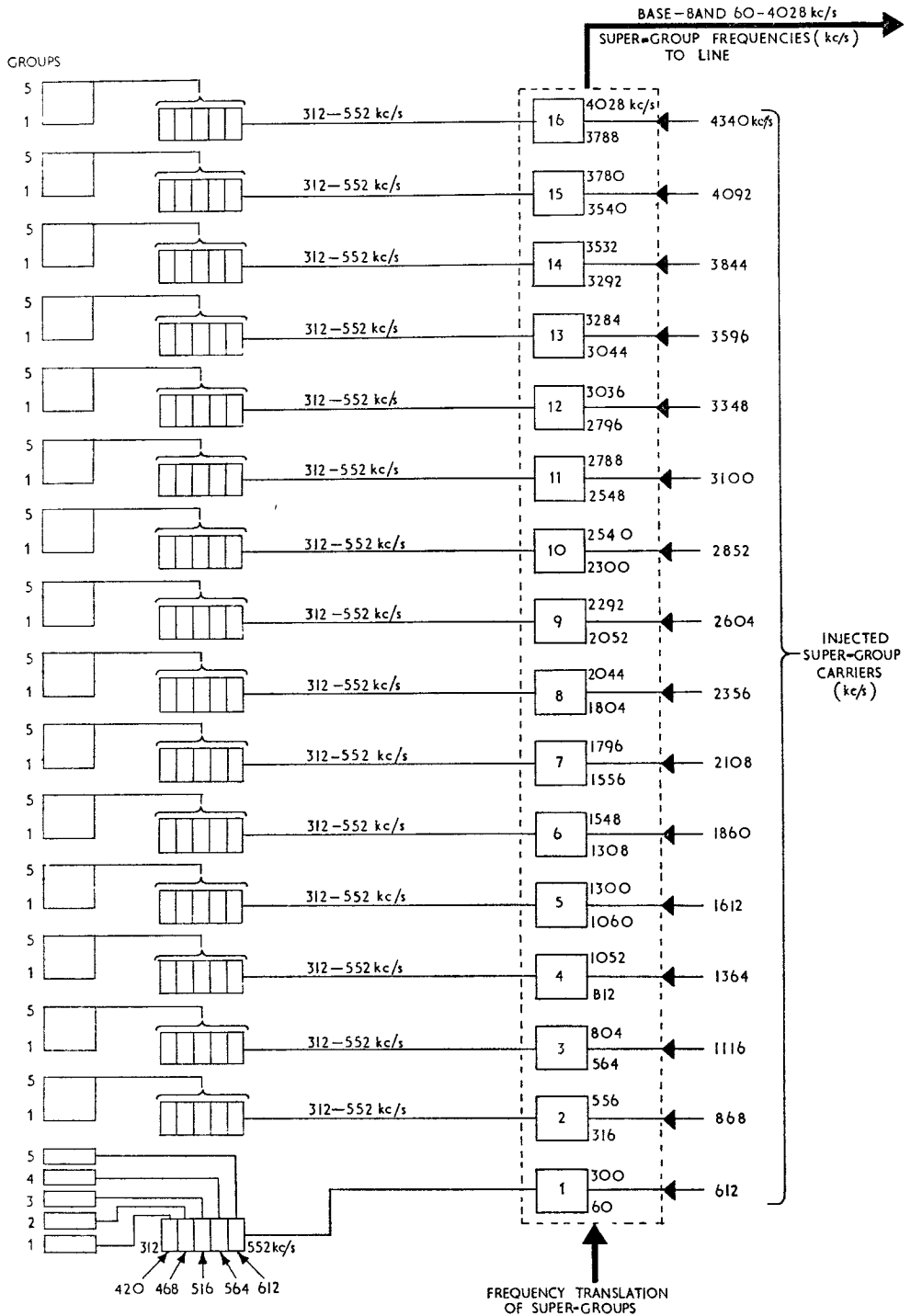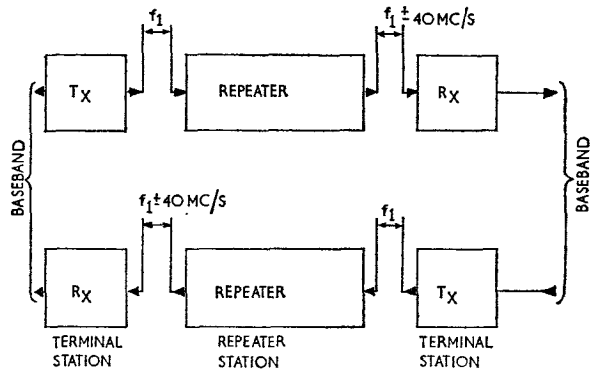

FIG. 3. MICROWAVE LINK

The Transmitter

The principal stages of one type of transmitter are shown in block form in Fig. 4. The modulation amplifier is a normal multistage video amplifier; the oscillator is a frequency modulated reflex klystron or travelling wave tube (t.w.t.); and the r.f. amplifier is another klystron or t.w.t. of higher power than the first.
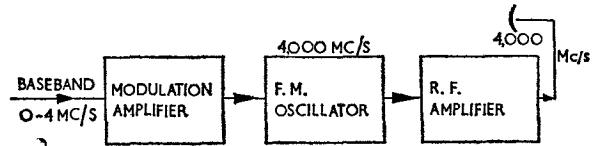
FIG. 4. TYPICAL MICROWAVE LINK TRANSMITTER

Another type of transmitter (Fig. 5) produces the f.m. output at an intermediate frequency (commonly 60 Mc/s). This is then changed to the final r.f. output. The f.m. oscillator is reactance modulated and the frequency changer is a t.w.t.
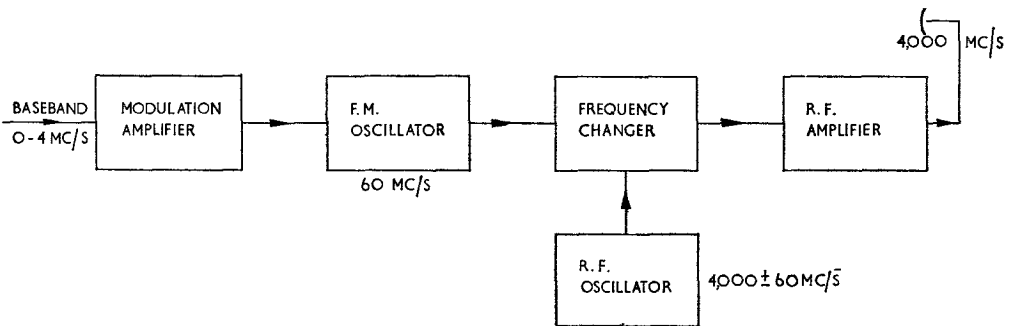
FIG. 5. ALTERNATIVE MICROWAVE LINK TRANSMITTER

193

The Receiver

The final receiver is a superhet f.m. receiver; a block diagram is shown in Fig. 6.

The frequency changer is normally a crystal diode and the local oscillator a klystron or coaxial line oscillator. The rest of the receiver follows standard f.m. practice.

| FREQUENCY CHANGER | I. F. AMPLIFIER | LIMITER | DISCRIMINATOR | VIDEO AMPLIFIER | BASEBAND |

LOCAL OSCILLATOR
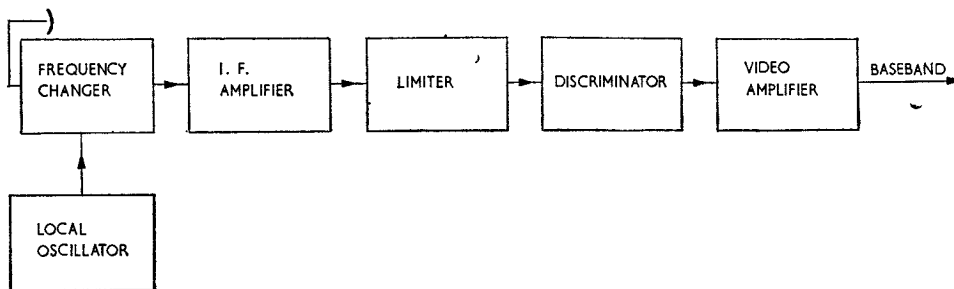
FIG. 6. A MICROWAVE LINK RECEIVER

The Repeater

Because of the difficulties of amplifying microwave signals, earlier practice was to change the incoming microwave signal to a lower frequency. Amplification was carried out at this lower frequency and the amplified signal then changed to the original input frequency $\pm$ 40 Mc/s, and re-radiated. A block diagram of this type of repeater, using an i.f. of 60 Mc/s, is shown in Fig. 7.
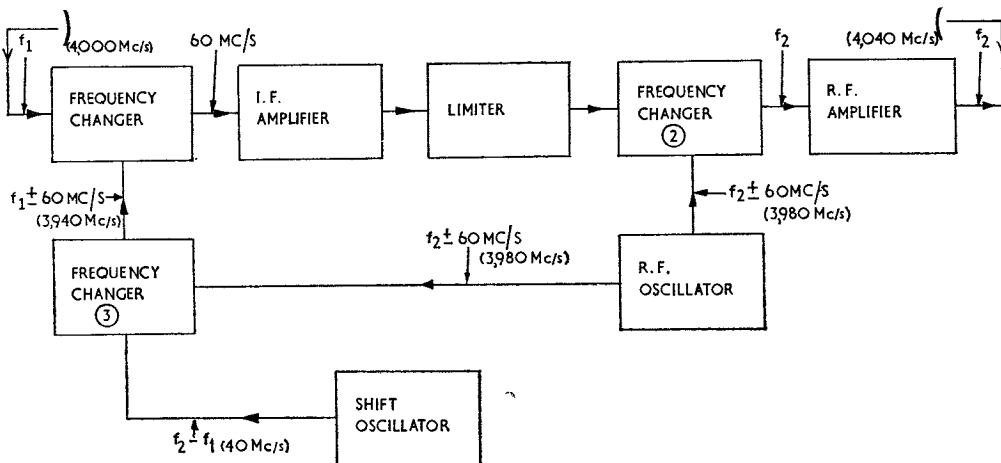
FIG. 7. LINK REPEATER UNIT

The frequency changers are crystal diodes or t.w.t's, depending on the power level; the i.f. amplifier is a normal multistage wideband amplifier using pentodes; the r.f. amplifier is a klystron or t.w.t.

The development of t.w.t's has enabled amplification of microwaves to be carried out without introducing excessive noise into the early stages of the repeater. Thus it is modern practice to use t.w.t's to amplify the incoming microwave signal and it is not necessary to reduce the signal to an i.f. of 60 Mc/s. To change the frequency of the received signal by $\pm$ 40 Mc/s, a t.w.t.

frequency changer and a shift oscillator are used. A block diagram of an all-t.w.t. repeater is shown in Fig. 8.
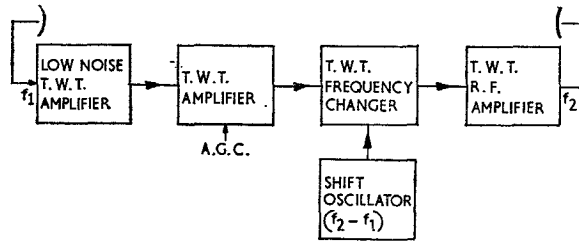
FIG. 8. A TWT REPEATER

The shift oscillator in both types of repeater would be a crystal controlled oscillator with frequency multiplication.

Repeater stations are usually unattended. If the main power supply fails, diesel-electric auxiliary supply sets immediately start up and provide power until the main supply is available again. The auxiliary supply operates automatically when a pilot signal in one of the stand-by channels falls below a certain level.

The Elgin-Wick Microwave Link

In order to bring out some other features of microwave links, the Elgin-Wick link across the Moray Firth will now be described.

The link carries 240 channels over a single 55 mile hop and links with the Wick-Kirkwall (Orkneys) microwave link.

In order to ensure that a good signal is received at all times despite abnormal propagation conditions, a combination of space and frequency diversity is employed.

The frequency diversity is achieved by using two transmitters at each terminal, working on the frequencies shown in Fig. 9.
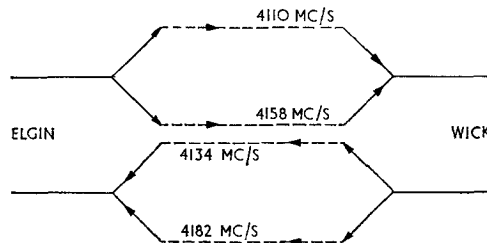
FIG. 9. TRANSMITTING AND RECEIVING FREQUENCIES

Space diversity is obtained by having two receiving aerials at different heights, Fig. 10 *a* shows the aerial system used at Elgin.

The different path lengths of direct and reflected rays may cause out-of-phase conditions at the upper receiving aerial (Fig. 10 *b*), but by having a lower aerial so placed that the path length difference is half a wavelength with respect to the upper aerial, reliable signals can always be obtained whatever the propagation conditions.

The comparatively small difference between the two frequencies employed enables the aerials to receive signals on either frequency.

195   (*A.L.* 3, *September*, 1964)
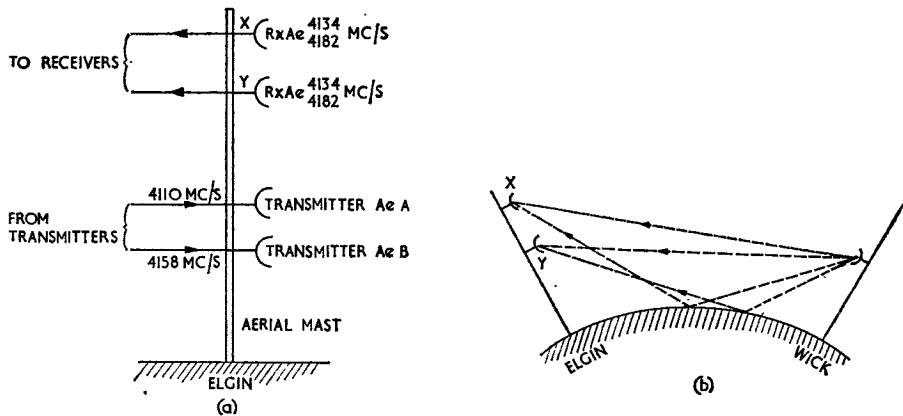
FIG. 10.   SPACE DIVERSITY

The aerial reflectors are 7 foot diameter paraboloids with a beam width of 3°.   The radiated power is 1 watt and the attenuation over the transmission path is 71 db.

Each terminal has four receivers, two for each frequency of reception but connected to different aerials.   The best signal of the four received is automatically selected by switches.

A block diagram of the transmitter and receiver arrangements is shown in Fig. 11.   At each terminal station there are two transmitters working on slightly different frequencies.   Each transmission can be received by two aerials at different heights, and each aerial feeds into two
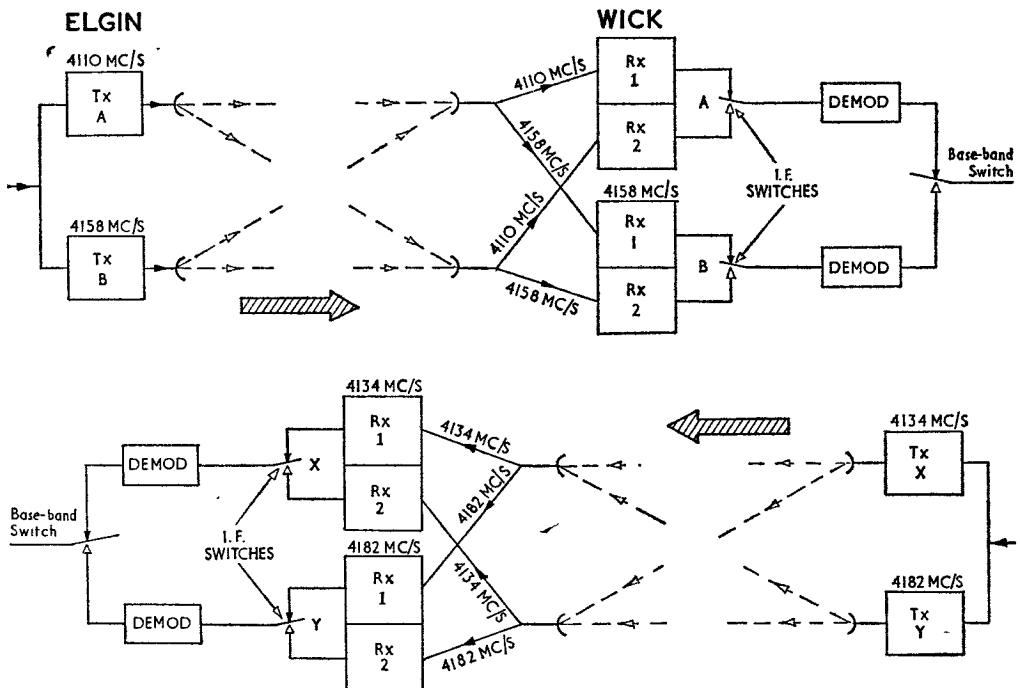
FIG. 11.   SIMPLE BLOCK DIAGRAM OF COMPLETE LINK

196

receivers tuned to the two transmission frequencies. Thus there is a choice of four receiver outputs. The i.f. switch selects the louder signal from receivers 1 and 2 for each frequency and the baseband switch selects the signal from A or B (or from X or Y), whichever has the better signal-to-noise ratio.

Transmitter

A simplified block diagram of one transmitter is shown in Fig. 12. The double line paths represent waveguides; single lines represent coaxial cable. After amplification the baseband
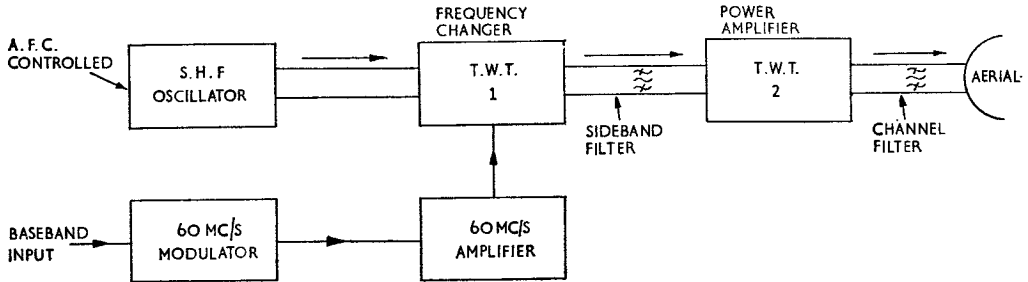
FIG. 12.   BLOCK DIAGRAM OF TRANSMITTER

signal is applied to the 60 Mc/s modulator. This is a frequency-doubled 30 Mc/s oscillator, using conventional valves, which is frequency modulated by the baseband.

The 60 Mc/s amplifier accepts the f.m. signal at a level of 0.5V and gives an output of 20 to 25 volts for use in phase-modulating the first t.w.t.

The s.h.f. oscillator is a reflex klystron frequency stabilised to $\pm$ 100 kc/s and producing an r.f. output at about 4,000 Mc/s. The f.m. signal is injected in series with the beam voltage of the first t.w.t. and the input waveguide receives the microwave s.h.f. carrier. This method of modulation is rather different from any described previously, but the result is to phase-modulate the microwave carrier, producing sidebands displaced from the carrier by multiples of 60 Mc/s. The filter selects the first upper sideband. The first t.w.t. has a beam current of 5 mA with a d.c. voltage of 1,400 V; the second t.w.t. has a 15 mA beam current and a d.c. voltage of 3,000 V.

Receiver

A simplified block diagram of one receiver is shown in Fig. 13. The inputs from the two receiving aerials are separated so that f_1 is passed to receiver A (or X) and f_2 to receiver B (or Y).
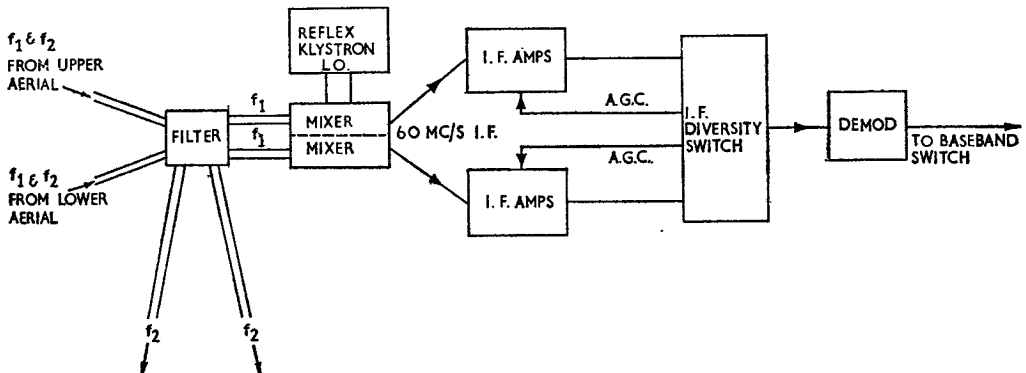
FIG. 13.   BLOCK DIAGRAM OF RECEIVER "A"

197

Each receiver (A or B) consists of two receivers, one handling the lower aerial signal and the other dealing with the signal picked up by the upper aerial. The system is duplicated for the other receiver frequency, so Fig. 13 shows only half of the complete receiving equipment.

The signal is mixed with the output from a r.f. local oscillator in a crystal mixer, the local oscillator being a low-power reflex klystron. The i.f. of 60 Mc/s is amplified by a succession of i.f. amplifier stages using pentode valves and the output from both i.f. strips is taken to the i.f. diversity switch which selects the larger amplitude output. The selected signal is then demodulated and applied to the baseband switch along with the demodulated output from receiver B where the signal with the better signal-to-noise ratio is finally selected.

CHAPTER 7

FACSIMILE

Introduction

Facsimile is a system of communication in which information is conveyed over long distances by landline or radio, in the form of pictures. By this means black and white charts, type-written messages, sketches, etc., can be automatically transmitted and received. In this chapter we shall consider the principles of facsimile transmission and reception and the outline of a typical facsimile transmitter and receiver.

Principle

The "eye" of a facsimile transmitter is an electron-multiplier photo-electric cell. As described in Part 1 of these notes, this is a device with an unheated cathode made of light-sensitive material. The number of electrons emitted by the cathode, and hence the magnitude of the output current, depends upon the amount of light striking it. In order to obtain the required resolution, the document is "scanned" by the photo-electric cell in thin spiral lines.

The document to be transmitted is wrapped round a drum and rotated at a constant speed of between 200 and 360 r.p.m. A small area of the document is brightly illuminated and the image of this area is focused through a lens and mirror on to a small white screen. An aperture in the screen allows light from the image to fall on an electron-multiplier photo-electric cell. As the drum and document rotate, the lamp and cell move along the drum axis. In this way the document is scanned line by line in the form of a closed spiral.

The amount of light reflected from the document will depend on whether the element under the aperture is black or white and so the output current from the photo-electric cell will vary with the shade of the document. This current is used to modulate a carrier wave which transmits the intelligence over line or radio link to the receiver.

At the receiver, the modulated signals are amplified, demodulated and then used to reproduce the black and white shades of the original document. The paper on which the document is reproduced may be either light-sensitive or electro-sensitive, depending on the type of receiver used. With a light-sensitive system (a photo-receiver) the demodulated signals are used to produce variations in the light intensity of a scanning lamp in the receiver. Light from this lamp, applied through lenses, exposes photographic paper secured to the receiver drum. The receiver scanning system is synchronised with that of the transmitter by means of transmitter synchronising pulses. After reception, the photographic paper is processed to produce copies of the original document.

In a receiver which uses electro-sensitive paper, the paper is drawn slowly between two electrodes called the *writing edge* and the *helix*. The helix rotates in step with the distant transmitter drum and the demodulated signal is applied between the helix and the writing edge. Thus current passes through the paper at the point of contact, causing a chemical action which stains the paper. The density of the mark produced depends upon the magnitude of the current; this is controlled by the transmitted signal. The rotation of the helix and the movement of the paper cause the point of contact to cross the paper in a series of lines, each line corresponding to a scanning line at the transmitter.

In order to bring out the main features of a facsimile system, we shall now consider the functions of essential circuits in a typical facsimile transmitter and receiver.

Facsimile Transmitter

A simplified block diagram of a facsimile transmitter is shown in Fig. 1. The output can be fed either to landlines linking the transmitter and receiver or to a radio transmitter which provides a modulated r.f. carrier.
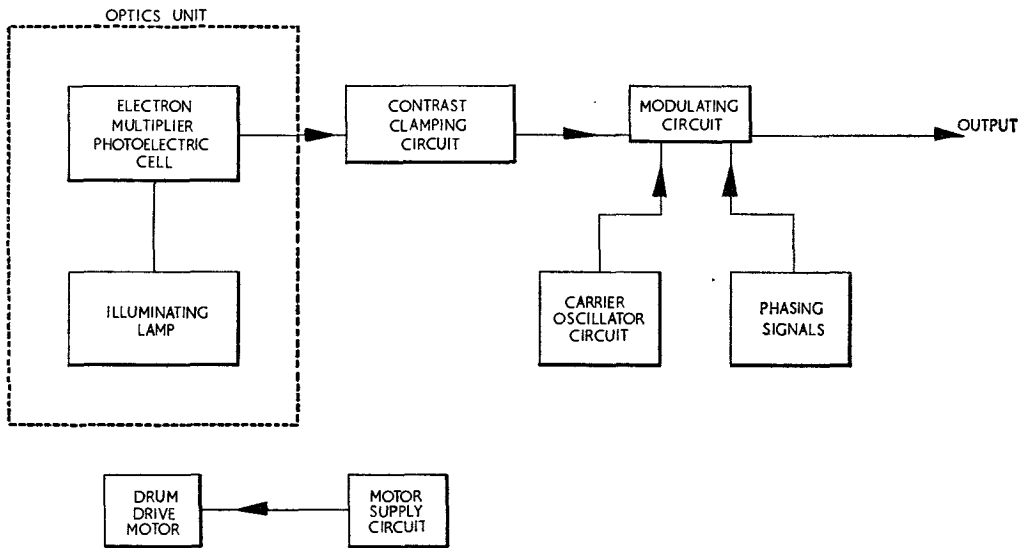
FIG. I. SIMPLIFIED BLOCK DIAGRAM OF A FACSIMILE TRANSMITTER

Optics Unit. An electron-multiplier photo-electric cell and an illuminating lamp are mounted on a sliding bar and are moved along the bar by a leadscrew. As the optics unit moves along the axis of the rotating drum, the light falling on the photo-cathode of the photo-electric cell varies with the shade of the document elements. This light controls the current flowing through the cell, an increase in light causing an increase in current. Hence the voltage developed across a load resistor in the final anode of the cell varies, being large for white elements and small for black.

The cathode of the photo-electric cell is held at an e.h.t. voltage of about $-1,000$ V and the final anode is taken through the load resistor to earth (Fig. 2). The output voltage varies from a negative value (about -5 V) for white to 0 V for black. This is fed to the contrast clamping circuit.

Contrast clamping circuit. This circuit provides one output voltage for white signals (0 V) and one for black signals $(+3$ V), for application to the modulator. Two clamping (or limiting) diodes are connected as shown in Fig. 3.

When an area of black passes under the scanning aperture of the optic unit, the voltage across R_L rises to about earth potential. This represents a positive rise of about 5 V which is limited by the action of rectifier 2 to $+3$ V. When the voltage across R_L falls to -5 V (white), rectifier 1 conducts and the output voltage is 0 V. Thus the voltage fed to the modulator is either 0 V (for white) or $+3$ V (for black). This offsets any slight variation in background whiteness and improves the appearance of the received copy.

Phasing signals. It is essential that the receiver helix rotates in step with the transmitter drum. To ensure this, a synchronising signal is transmitted once per revolution of the drum.
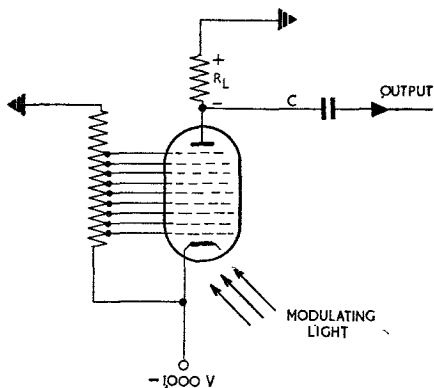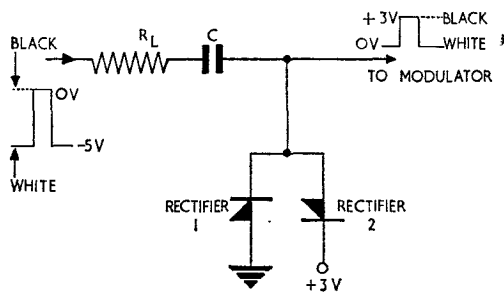
FIG. 2. CIRCUIT OF PHOTOELECTRIC CELL

FIG. 3. CONTRAST CLAMPING CIRCUIT

For 5 seconds before the document signal is transmitted, a black signal is sent, interrupted by a white signal once per drum revolution. These *phasing* signals are used to prepare the receiver phasing circuits for reception of the document.

Carrier oscillator circuit. This circuit may be any conventional type of oscillator with good frequency and amplitude stability. When the transmitter and receiver are connected by landline the output frequency is audio. When the link is radio, the carrier frequency is r.f., generated in the radio transmitter.

Modulator circuit. The outputs from the carrier oscillator circuit and from the contrast clamping circuit are fed into the modulator. Here the carrier is amplitude modulated by the document signal and then passed to the output line circuits. If a radio link is employed, the modulation can be a.m. or f.m.

Drum drive circuits. The motor which rotates the transmitter drum is a synchronous motor supplied with an input voltage obtained from the mains at 50 c/s, or from a 1,000 c/s "fork" oscillator. The frequency generated by this type of oscillator is controlled by the mechanical vibrations of a tuning fork and hence has a high degree of stability. The output from the fork oscillator is fed to a power amplifying circuit where sufficient power is obtained to drive the drum motor.

Facsimile Receiver

The facsimile receiver (or recorder) changes the modulated input signals into a reproduction of the original transmitted document. The input is amplified, demodulated and applied between a rotating helix and a metal writing edge. Damp electro-sensitive paper is unrolled slowly between these electrodes and the current passing through the paper reproduces the black and white shades of the original document. The essential circuits of a facsimile receiver for use with a landline link are shown in Fig. 4. If a radio link is used, a superheterodyne receiver is required to receive and amplify the radio signals; the output from this receiver is fed to the marking power amplifier and signal detector circuits.

Before the document is transmitted, a 5 seconds period of black signal and white phasing signal is sent. These signals are used to prepare the receiver for reception of the document. When the signal detector circuits respond to the carrier signal, relays are energised and the following occurs:—

a. The helix drive motor is started.

b. After a delay of about 2 seconds, the phasing circuits operate the helix drive clutch, the helix starts to rotate and the writing edge makes contact with the recording paper.
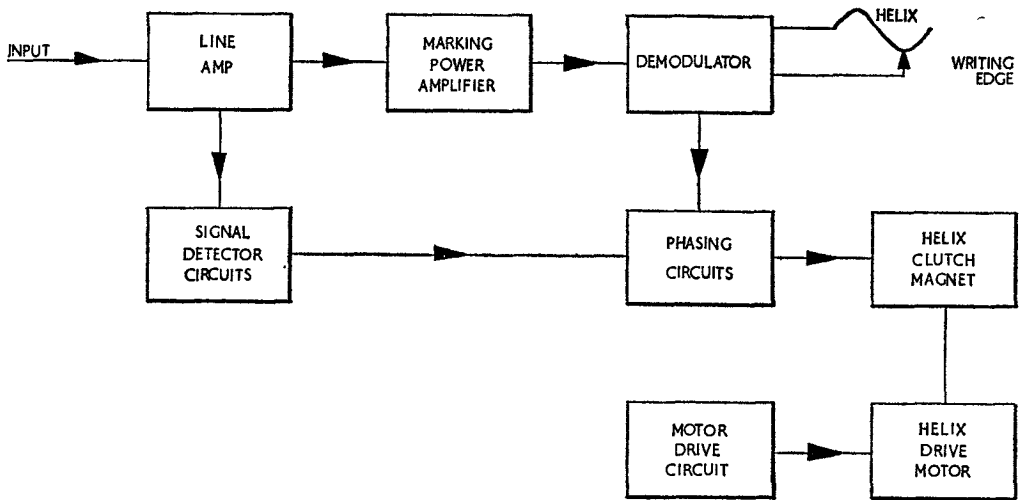
```
INPUT →  [LINE      ] → [MARKING    ] → [DEMODULATOR]          HELIX
         [AMP       ]   [POWER      ]                                    WRITING
                        [AMPLIFIER  ]                                    EDGE
              ↓              ↓              ↓
         [SIGNAL    ] →              → [PHASING    ] → [HELIX      ]
         [DETECTOR  ]                  [CIRCUITS   ]   [CLUTCH     ]
         [CIRCUITS  ]                                  [MAGNET     ]

                                       [MOTOR      ] → [HELIX      ]
                                       [DRIVE      ]   [DRIVE      ]
                                       [CIRCUIT    ]   [MOTOR      ]
```

FIG. 4.   SIMPLIFIED BLOCK DIAGRAM OF A FACSIMILE RECEIVER

c. After a further 3 seconds delay, the demodulated output from the marker power amplifier is applied between the writing edge and the helix and the receiver is ready to record the document.

On receipt of the document signal, recording takes place and the above conditions are maintained until, at the end of transmission, the receiver returns to the stand-by condition.

Line amplifier and signal detector circuits. The received signal is applied to a voltage amplifier circuit which has two outputs, one to the marking power amplifier and the other to the signal detector circuits. When a signal of the correct carrier frequency is received, relays are operated and remain energised throughout the transmission. These relays prepare the phasing circuits and the holding circuit of the helix motor clutch, and start the helix motor.

Marking power amplifier and demodulator circuits. The input from the line amplifier is applied to a power amplifying circuit, the output from which is rectified. The rectified output is connected to the phasing circuits during the initial 5 seconds phasing period, and when the phasing is completed the output is transferred to the helix and writing edge.

Phasing circuits. These circuits consist of a thyratron and a cold-cathode gas-filled trigger valve with their associated components. The output from the signal detector during the initial 5 seconds preparatory period consists of synchronising pulses. These pulses are differentiated and after a delay of about 2 seconds, a differentiated pulse fires the thyratron. The large amplitude output from the thyratron energises the helix motor clutch solenoid which pulls the clutch in. The helix then starts rotating in step with the remote transmitter drum. The writing edge is also moved into contact with the paper and helix.

After a further 3 seconds delay, the trigger valve strikes and the demodulated power output is connected to the helix and writing edge.

After completion of a transmission, the thyratron is extinguished and the receiver returns to stand-by.

Helix drive circuit. As with the transmitter drum, the receiver helix is driven by a synchronous motor. The motor input is derived from the mains or from a stable fork oscillator, the output of which is amplified to give enough power to drive the helix.

SECTION 5

AERIALS AND PROPAGATION

CHAPTER 1

AERIAL ARRAYS

Introduction

A simple aerial radiates e.m. energy in all directions except along its length. That is, a *vertical* aerial radiates all round itself in the *horizontal* plane and a *horizontal* aerial radiates all round itself in the *vertical* plane (Fig. 1).
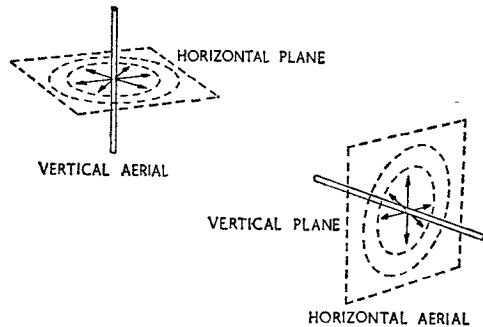

FIG. I. RADIATION FROM A SIMPLE AERIAL

However, there are many occasions when we require a *directional* aerial to form the energy radiated into a beam, rather like the beam of light from a searchlight. Directional aerials are formed by combining several simple driven aerials into an aerial array or by using undriven reflectors. The basic types of directional aerial were described in Part 1 of these notes. Further details of directional aerials will now be considered.

Beam Width

A convenient way of measuring the directional properties of an aerial array is by measuring the angle between the aerial and the two points on the radiation diagram where the radiated power has fallen to half its maximum value. As radiation diagrams are usually plotted in field strength units which are proportional to the square root of power units, the half-power points correspond to the points where the field strength has fallen to 0.707 of maximum (Fig. 2). The

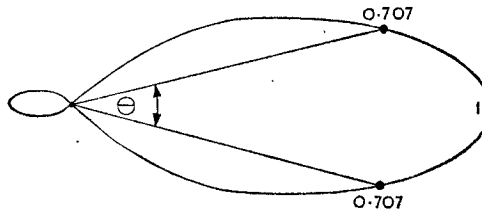FIG. 2. BEAM WIDTH

angle θ, measured in degrees, is the *beam width* of the array. As the array is made more directional, the beam width decreases and thus θ gives us a good means of measuring the directional properties of the array.

Aerial Gain

Another method of stating the effectiveness of a directional aerial array is by comparing the maximum power radiated in the main direction of the beam with the power radiated by a non-directional aerial supplied with the same total power. The gain of an aerial array is thus defined as:—

$$\frac{\text{Power radiated in direction of maximum radiation}}{\text{Power radiated from a non-directional aerial}}$$ both aerials being fed with the same value of current. This ratio increases as the beam width decreases.

Producing a Narrow Beam

A common aerial array used to produce a narrow beam is the stacked array discussed in Part 1. This array consists of a number of driven dipoles spaced a half-wavelength apart and fed with *in-phase* currents. If current in one of the dipoles is out of phase with that in the others the width of the main lobe is increased. Also, with currents out of phase, the aerial radiates in a direction other than at right-angles to the plane of the array. The array is said to have a "squint". An array which radiates maximum energy at right angles to the plane of the array is called a broadside array.
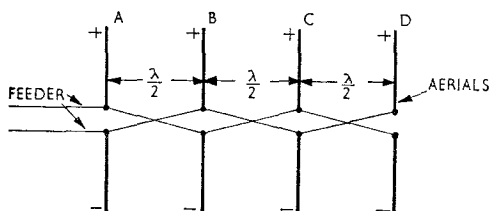

FIG. 3. FEEDING A BROADSIDE ARRAY

Method of Feeding Broadside Arrays

The method of feeding a broadside array with in-phase currents is shown in Fig. 3.

Aerial A is connected to the feeder at a certain point. Half a wavelength from this point the r.f. energy in the feeder reverses in phase; thus the feeder connections to aerial B are reversed in order to feed aerial B in the same phase as aerial A. This method of connection is used throughout the array, thus ensuring that all aerials radiate in phase.

The input impedance of a centre-fed half-wave dipole is about 75 ohms. Since the aerials in a broadside array are in effect in parallel across the feeder, the combined impedance of, say, eight dipoles would be small. The total impedance would be much too low to match normal feeder lines and, for this reason, some broadside arrays use aerials which are longer than half a wavelength in order that the combined aerial impedance matches the characteristic impedance of the feeder.
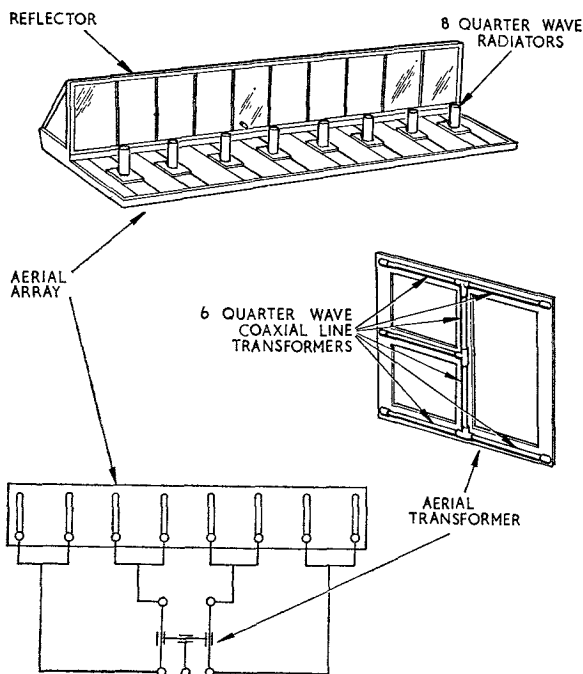

FIG. 4. A BROADSIDE ARRAY

Typical Broadside Array

Some low-power broadside arrays are fed from a 46 ohm coaxial line and

an arrangement of quarter-wave transformers is required to match the aerials to the feeder. A typical example is shown in Fig. 4.

End-fire Array

The direction of maximum radiation from an end-fire array is along the line of the individual aerials. To obtain this, the individual aerials are fed, not in phase as in the broadside array but in such a way that the phase difference between adjacent aerials is the same, in wavelengths, as the aerial spacing.

This is shown in Fig. 5 where two dipoles spaced by a quarter-wavelength radiate 90° out of phase, the radiation from aerial 1 leading that from aerial 2 by 90°.

Radiation from aerial 2 in the direction A is thus in antiphase with the radiation from aerial 1 and resultant radiation in direction A is therefore zero. However, radiation from aerial 1 in the direction B arrives at aerial 2 in phase with the radiation from that aerial and increases radiation in direction B. Maximum radiation is in one direction only, in line with the aerials.
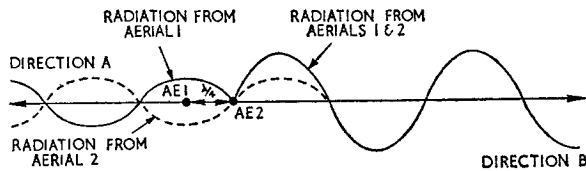

FIG. 5. RADIATION FROM END-FIRE ARRAY

Radiation Diagrams

As with the broadside array, the beam width of the radiation from an end-fire array becomes narrower as the number of driven aerials is increased (Fig. 6).
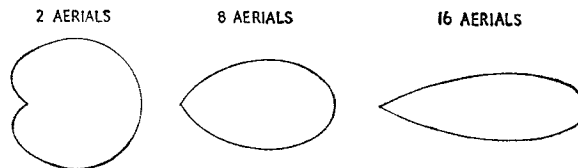
FIG. 6. END-FIRE RADIATION DIAGRAMS

Method of Feeding End-Fire Arrays

The phase difference between the input to each aerial in an end-fire array must be equal to the spacing between the aerials. The method of feeding the aerials is therefore as shown in Fig. 7. The aerials are spaced at intervals of a quarter-wavelength along the feeder. Current in aerial B lags that in aerial A by 90° and current in aerial C lags that in aerial B by 90°. Thus phasing conditions are satisfied for radiation in the direction shown in Fig. 7.

FIG. 7. FEEDING AN END-FIRE ARRAY

205 *(A.L. 3, September, 1964)*

Beam Swinging

An array of dipoles fed in phase produces maximum radiation at 90° or broadside to the line of the array and an array in which the phasing between each element differs by 90° produces a beam in line with the array. By making the dipoles radiate out of ph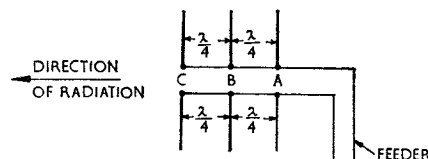ase by different phase angles, an array can be made to have its direction of maximum radiation at any desired angle. This might be done, for example, to beam radiation towards a particular part of the ionosphere. This technique is sometimes known as *beam-slewing*.
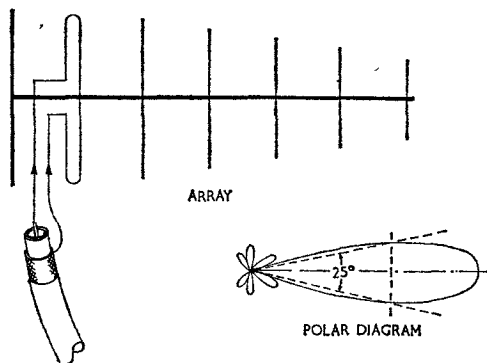
ARRAY

POLAR DIAGRAM

FIG. 8. YAGI ARRAY WITH FOLDED DIPOLE

Yagi Array

Another array which produces directional radiation is the Yagi array described in Part 1. In this array one element is driven and the other elements (directors and reflector) are parasitic. One effect of the parasitic elements is that they decrease the input impedance of the driven dipole. This can easily be seen, since the currents in the parasitic elements must be supplied from the energy radiated by the dipole, causing more current to be drawn from the transmission line, so effectively reducing the input impedance to the aerial.

It is therefore difficult to match a Yagi array, using a conventional dipole, to its coaxial feeder. Because of this, a folded dipole with several parasitic elements is often used in a Yagi array. The folded dipole has an input impedance of about 300 ohms but when it is used in conjunction with parasitic elements its input impedance is reduced to less than 100 ohms and matching to a coaxial feeder is simplified.

Fig. 8 shows a Yagi with a folded dipole. It will be recognised as the familiar television aerial, and has an input impedance of about 70 ohms.

CHAPTER 2

TRAVELLING WAVE AERIALS

Introduction

Directional aerial arrays for use in the v.h.f. and u.h.f. bands are fairly simple to construct since each aerial is relatively short in length. However, systems operating at high and medium frequencies, where one wavelength may be several hundred feet long, often employ a *travelling-wave* aerial to obtain directivity.

Standing-Wave and Travelling-Wave Aerials

When an alternating current flows in a conductor, electromagnetic waves are always radiated from that conductor. Standing-wave aerials such as the half-wave dipole and quarter-wave Marconi act as a resonant circuit (Fig. 1a). The travelling-wave aerial, however, is not a resonant circuit and standing waves are not produced. Instead, a travelling wave moves from the generator, along the wire, to the terminating resistor, in one direction only (Fig. 1b).
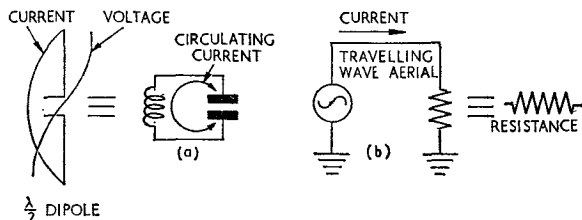
FIG. I. STANDING-WAVE AND TRAVELLING-WAVE AERIALS

Principle of the Travelling-Wave Aerial

We already know that in a twin-wire transmission line some radiation from the lines occurs; to reduce this to a minimum, the two lines are mounted close together (less than one tenth of a wavelength apart). Thus the E and H fields produced by the current in one wire almost completely cancel those produced by current in the other wire and the resultant radiation is negligible.

A travelling-wave aerial is a transmission line used to radiate e.m. waves. Radiation is achieved by using the earth as one wire and mounting the other wire parallel to, and high above, the earth. The usual length is between 3 and 8 wavelengths and, unlike a resonant aerial, the length is not critical. The efficiency of a travelling-wave aerial depends on the magnitude of the current travelling towards the termination. If reflection from the termination occurs, the amplitude of the current is decreased (Fig. 2). If the aerial is terminated by a resistance equal in value to the characteristic impedance of the line, no reflection occurs and the current is therefore uni-directional and of maximum value. Maximum radiation from the aerial then occurs.
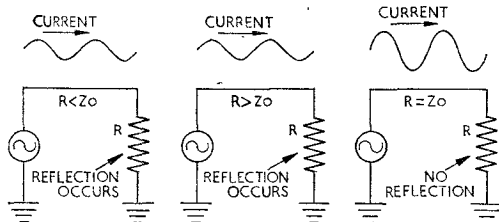
FIG. 2. REFLECTION IN A TRAVELLING-WAVE AERIAL

207

When a travelling wave moves along a conductor in one direction only, radiation from the conductor occurs in the direction of travel as shown in Fig. 3. The radiated e.m. energy from each part of the wire is in the form of a cone. The amount of energy radiated depends upon the amplitude of the travelling wave and upon the frequency.

Radiation Patterns

Fig. 4 shows radiation patterns for three travelling-wave aerials of the same physical length but fed with current at different frequencies, the frequency of *c* being twice that of *a*. In each case two major conical lobes are formed, the angle between the lobes depending to some extent on the frequency, but this angle varies by only 10° when the frequency is doubled. This means that one aerial of fixed physical length will provide satisfactory communication over a fairly wide band of frequencies.
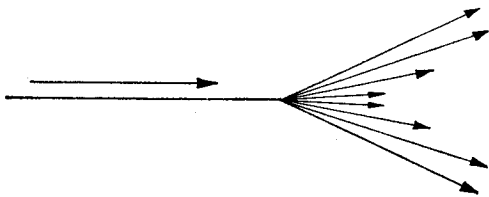
FIG. 3. RADIATION FROM A TRAVELLING-
WAVE AERIAL
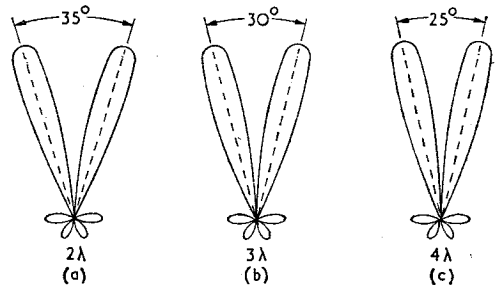
FIG. 4. RADIATION PATTERNS OF
TRAVELLING-WAVE AERIALS

The Inverted V Aerial

This is a travelling-wave aerial used at h.f. It consists of a wire inclined at 30° to the horizontal as shown in Fig. 5. Each half of the wire is between 2 and 4 wavelengths long. The wire is earthed via a resistor equal in value to the characteristic impedance of the feeder.
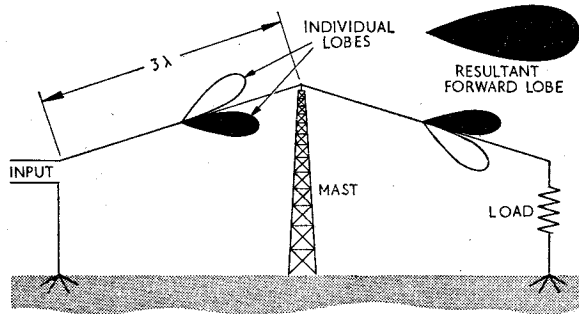
FIG. 5. THE INVERTED V AERIAL

As shown in Fig. 5 two lobes lie parallel to the ground and give a concentrated forward lobe in the direction shown. The inverted V aerial has good directional properties and gives good communication over a 3 : 1 frequency range.

The Rhombic Aerial

The rhombic aerial is a travelling-wave aerial used extensively for long-range h.f. communication. As shown in the plan view of Fig. 6, it consists of four wires forming a diamond or rhombus

208

shape in the horizontal plane. It is fed by a 600 ohm line and terminated with a 600 ohm resistor to prevent reflections.
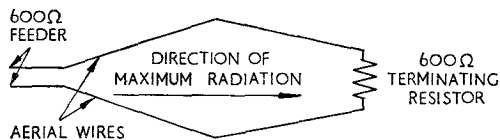
FIG. 6. PLAN VIEW OF A RHOMBIC AERIAL

The cone of radiation from each of the four sides is shown in Fig. 7 and if the angle θ and side length L are correctly chosen, the shaded lobes combine to give increased directivity in the forward direction. As with the inverted V aerial, the characteristics of the rhombic are not greatly affected by changes in frequency and a frequency range of 2 : 1 can easily be covered.

Variation of Z_o

The characteristic impedance, Z_o, of a transmission line depends on the diameter of the conductors and their distance apart. Since the distance apart of the wires in a rhombic aerial varies, the Z_o varies, being greatest where the wires are far apart. This means that reflections from the termination occur, reducing the amount of energy radiated by the aerial. This can be a serious disadvantage in an aerial handling high power and to overcome it the *multi-wire* rhombic is often used (Fig. 8).

In this arrangement each leg of the rhombic consists of several wires tapered as shown. Thus as the spacing increases, the effective diameter of the wire increases and the Z_o is kept reasonably constant.
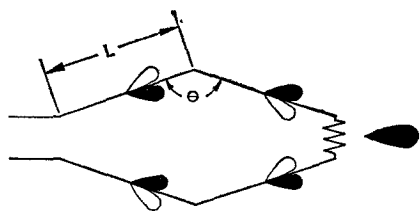
FIG. 7. RADIATION FROM A RHOMBIC AERIAL IN THE HORIZONTAL PLANE
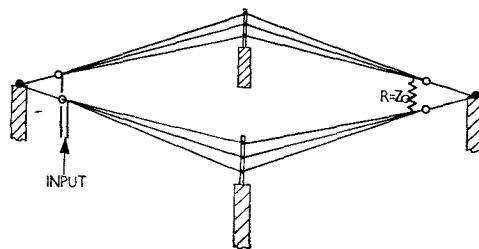
FIG. 8. MULTI-WIRE RHOMBIC

Termination of Rhombic Aerials

The terminating resistor of a rhombic aerial used with a high power transmitter must be capable of dissipating between 30% and 50% of the input power. The resistors are sometimes made of large blocks of carbon compound occupying several cubic feet.

An alternative method of terminating a rhombic aerial is to use an *absorber line* made of iron wires (Fig. 9). The dimensions of the wires and their spacing gives the correct terminating resistance and the heat developed in the line is dissipated in the air. The absorber line, which may be 100 yards long, is terminated in a smaller resistor.

Effect of Ground Reflection

The free space horizontal and vertical radiation patterns of a rhombic aerial are very similar, as shown in Fig. 10a. A practical rhombic array has to be mounted close to the ground and it

209
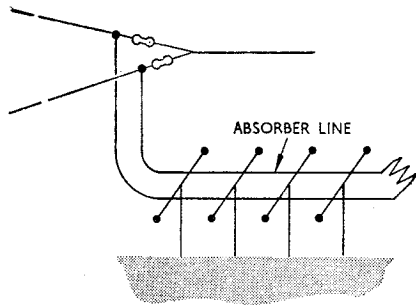
FIG. 9. ABSORBER LINE TERMINATION
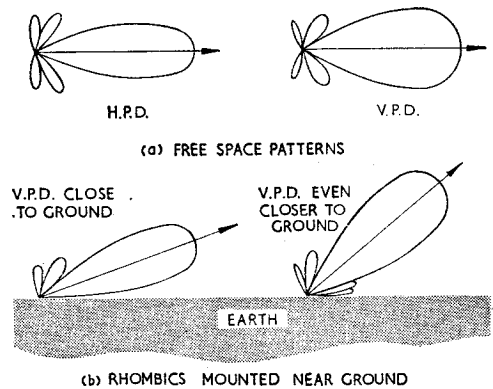
FIG. 10. EFFECT OF GROUND ON RADIATION PATTERN

can be seen that all radiation below the horizontal will strike the ground and be reflected. The result will be to tilt the main lobe upwards (Fig. 10b); the nearer the aerial is to the ground the greater will be the upward tilt.

This effect is used in long range h.f. point-to-point communication systems to beam the radiation at a desired angle to the ionosphere. By mounting the aerial at the correct height above the ground the required angle of elevation can be obtained and optimum reflection from the ionosphere results.
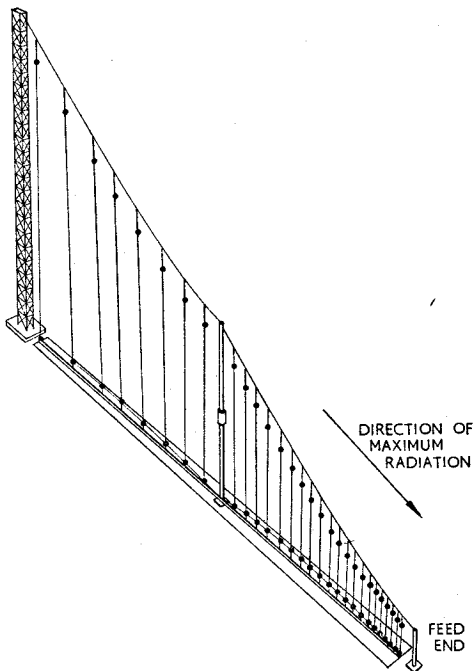

FIG. 11. VERTICAL LOG PERIODIC AERIAL

The Log Periodic Aerial

The main disadvantage of a rhombic aerial is that the installation requires a large ground area. Also, although the multi-wire rhombic has a fairly wide bandwidth its gain and directivity do vary with changes in frequency. An aerial which largely overcomes these disadvantages is the *log periodic* aerial. This aerial can be designed to cover the whole h.f. (or m.f.) band and has the important property that its characteristics remain constant for any frequency within that band. Horizontally or vertically polarised aerials can be built and the ground area covered is less than one third of that required for a rhombic. Fig. 11 shows a vertical log periodic aerial.

The main characteristics of log periodic aerials are summarised as follows:—

a. The direction of maximum radiation is towards the feed input.

b. The beam width in the vertical plane is about 50° and that in the horizontal plane about 60°.

210

c. The power gain may be between 8 and 25 compared with an omni-directional aerial, but for any one installation the gain remains constant over the designed frequency range.

d. The electrical distance (i.e., distance measured in wavelengths) of the effective radiating portion of the aerial from the feed point and from the ground is fairly constant.   This means that the radiating elements of the aerial move from the feed point as the frequency is lowered.

e. The highest and lowest frequency which the aerial can accept without changing its characteristics are decided by the shortest and longest elements respectively.

The mast height for a log periodic aerial designed for a 3 to 30 Mc/s frequency range is about 220 feet and the overall length of the aerial about 490 feet.

CHAPTER 3

CHOICE OF FREQUENCY AND AERIAL

Ground and Sky Waves

The propagation of radio waves has been dealt with in Part I of these notes. It will be remembered that an aerial radiates *ground waves* and *sky waves*. The ground wave is propagated parallel to the earth's surface and is attenuated as it moves from the transmitter aerial. The decrease in field strength over a given distance depends on the frequency of the wave, i.e. as frequency *increases* field strength *decreases*. The sky wave is propagated away from the earth's surface and may or may not be reflected from the ionosphere. In general, frequencies below about 30 Mc/s are reflected while higher frequencies are not reflected.

Communication at VLF and LF

Since attentuation increases as the frequency increases, low frequencies are employed for long-distance broadcast communication using the ground wave. The sky wave radiated from aerials working in the v.l.f. and l.f. bands is small and is reflected from the ionosphere, as shown in Fig. 1. Since the reflected wave is weak compared to the ground wave it does not greatly affect the resultant signal induced in a receiving aerial and fading is not noticeable.
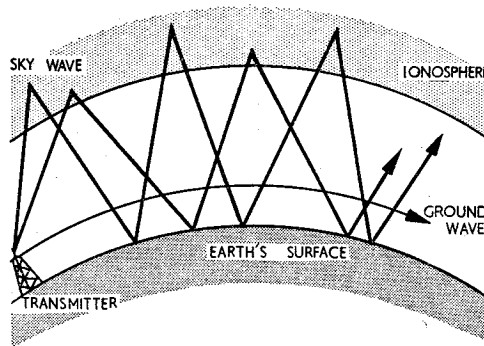
FIG. 1. GROUND AND SKY WAVES AT VLF AND LF

To obtain a reasonable radiation efficiency, v.l.f. and l.f. aerials must be comparable in length to the operating wavelength; they are therefore very large structures and in most cases this prohibits the use of directional arrays. Because of their non-directivity they are mainly used in world-wide broadcast systems, e.g., to shipping. A vertically polarised ground wave is not attentuated as much as a horizontally polarised wave and therefore most v.l.f. and l.f. aerials are vertical.

Often towers with a flat top are used as aerials (Fig. 2a). If wires are used, the aerials are suspended between towers as shown in Fig. 2b. At low frequencies the earth connection is very important and counterpoises are used to improve the earth's conductivity.

Communication at MF

The m.f. band is generally used for medium range broadcasting. The ground wave attenuation at m.f. is higher than at l.f. but an adequate coverage for normal broadcast purposes is obtained, the actual range obtained depending on the transmitter power.

FIG. 2. VLF AND LF AERIALS

The attenuation experienced by the sky wave depends on the state of the ionosphere and varies throughout the day and night. The greatest variation occurs at sunrise and sunset because the attenuation depends upon the degree of ionisation and this in turn depends upon the sun's activity. At m.f. most sky wave attenuation occurs during the day; this is the reason why some continental broadcasting stations are only received in this country during the hours of darkness.
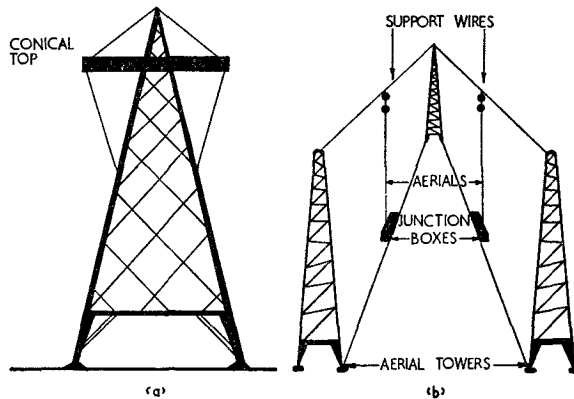
In places where both ground and sky waves are received simultaneously, fading is sometimes experienced because of the variation in phase between the two waves. In some cases the amount of fading is too great for the receiver automatic gain control to provide sufficient compensation. When this is the case, two or three widely spaced aerial systems can be connected via relays to the receiver. In this way a fairly constant input can be obtained by selecting the aerial which is giving the maximum output. This system is called *space diversity reception* and is used on some networks where reliable communication is essential throughout the 24 hours.

Aerials for MF and HF

The aerials used in m.f. broadcasting are fairly large structures and are usually omni-directional, although in some cases slightly directional aerials are used. For m.f. communication, towers supporting L type and T type aerials are usual.

At h.f. it is possible to use directional aerial arrays such as broadside, stacked and rhombic. Towards the higher end of the h.f. band folded dipoles, which have wider bandwidths than ordinary dipoles, are used on ground installations. Airborne h.f. installations may use a general purpose suppressed aerial of the helmet type or a notch aerial.

Point-to-point HF Communication

Radio systems operating at h.f. are often required to provide communication between two fixed stations and are thus termed point-to-point systems. The system may be long or short range. Because of the high attenuation of the ground wave at h.f. the ground wave is used only for short range communication; long range point-to-point working relies on the reflection of the sky wave from the ionosphere.

We already know that reflection from the ionosphere depends upon frequency and that above about 30 Mc/s the wave penetrates the ionosphere and is not reflected. Reflection also depends upon the degree of ionisation and therefore upon the time of day or night and the season of the year.

213 *(A.L. 3, September, 1964)*

Reflection of a radio wave from the ionosphere is not a straight-forward "bouncing effect" (Fig. 3a). The wave travels into the ionised layer and is gradually bent or *refracted* so that it
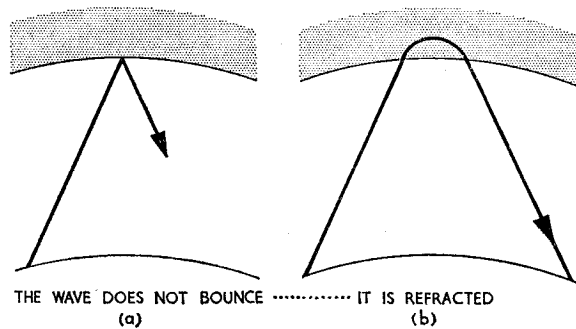
THE WAVE DOES NOT BOUNCE ············ IT IS REFRACTED
(a)   (b)

FIG. 3.   REFRACTION

eventually turns towards the earth (Fig. 3b).   The amount by which the wave is refracted depends upon :—

a. Frequency.   The higher the frequency of the wave the further it penetrates into the ionosphere before it turns towards the earth.

b. Degree of ionisation.   The more free electrons there are in the ionised layer the greater is the angle of refraction.

c. Angle of incidence.   The greater the angle of incidence at which the wave strikes the layer (Fig. 4) the less refraction is required before it returns earthwards.

From these three facts we see that when it is required to communicate between two fixed stations, the frequency used and the angle at which the sky wave should radiate from the aerial depend upon the degree of ionisation, i.e. upon the time of day or night and upon the season.

Day and Night Frequencies

If communication between stations is to be reliable throughout 24 hours, the working frequency must be changed at certain times to compensate for the changing skip distances caused by the varying degrees of ionisation in the ionosphere.

The radio wave is attenuated in the ionosphere. This attenuation varies with frequency and becomes less as frequency increases. Hence as high a frequency as possible must be used.
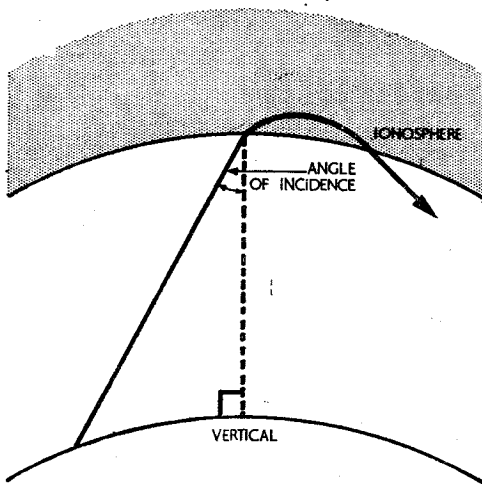
FIG. 4.   ANGLE OF INCIDENCE

A frequency which is satisfactory by day is too high for night use for the wave could then pass right through the ionosphere. Thus the communication frequency must be reduced at night and in practice 3 or 4 changes may have to be made throughout the 24 hours. A station operating on 6 Mc/s during the day might use 2 Mc/s at night.

214

Maximum Usable Frequency (MUF)

Point-to-point communication using the sky wave is possible only if the wave is reflected by the ionosphere. If the frequency is too high for reflection the sky wave is not returned to the earth's surface. The highest frequency which will be reflected by the ionosphere when the wave is directed vertically upwards is called the *critical frequency*. At greater angles of incidence higher frequencies would be reflected but for each angle there is a *maximum usable frequency* (m.u.f.). Thus the range for long distance point-to-point communication cannot be increased indefinitely merely by increasing the operating frequency and so increasing the skip distance. As shown in Fig. 5 an increase in frequency from f_1 to f_2 results in an increase in range from D_1 to D_2; but a further increase in frequency achieves no further increase in range because the sky wave is not returned to earth.

Since the ionosphere is constantly changing, the critical frequency and the m.u.f. also vary. The changes do not follow a constant pattern —the critical frequency for 1200 hours on 25th March, 1963 might well differ from that at 1200 hours on 25th March, 1964, for example. Prediction charts giving the most suitable m.u.f. for various paths over the earth's surface are pub-

FIG. 5. MAXIMUM USABLE FREQUENCY

lished throughout the year. In practice, the frequency used is about 80% of the m.u.f. and is called the *optimum working frequency* (o.w.f.). This slightly lower frequency allows a margin for ionospheric fluctuations.
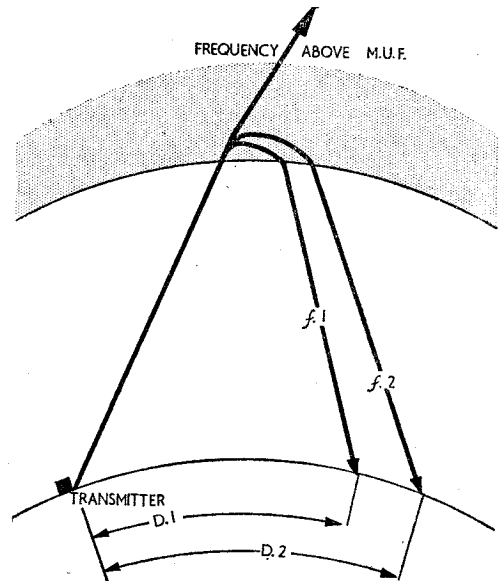
Aerials for HF Communication

Directional aerial arrays can be constructed for point-to-point communication in the h.f. band. Travelling wave aerials such as the rhombic are widely used because these aerials have a wide bandwidth and only a few aerials are required to cater for changes in frequency due to changes in the ionosphere. For short distance point-to-point working an omni-directional Marconi aerial is often used.

Suppressed aerials for airborne h.f. communication take the form of the helmet concealed aerial or the Y shaped notch aerial. The latter is cut in the leading edge of the tail fin and a matching unit is provided so that efficient radiation is obtained throughout the band.

Communication at VHF, UHF and SHF

Because v.h.f., u.h.f. and s.h.f. ground waves are severely attenuated and the sky waves penetrate the ionosphere, communication at these frequencies can, in general, take place over only short distances.

This does not mean that the v.h.f., u.h.f. and s.h.f. bands are unimportant for at these frequencies highly efficient directional aerials with beam widths as narrow as $\frac{1}{2}°$ can be designed.

Communication systems which require large sideband frequencies also need a high carrier frequency. Hence television, f.m. broadcasting, pulse transmission systems and frequency

division multiplex systems use v.h.f., u.h.f. and s.h.f. carrier frequencies. Also, the use of these frequencies enables many more stations to be accommodated without causing adjacent channel interference.

Because of the short ground wave range, ground-to-ground communication is limited to 'line-of-sight' distances as shown in Fig. 6. Since the height of an aircraft aerial can be much greater than that of a ground station, a greater uninterrupted line-of-sight range is possible for an
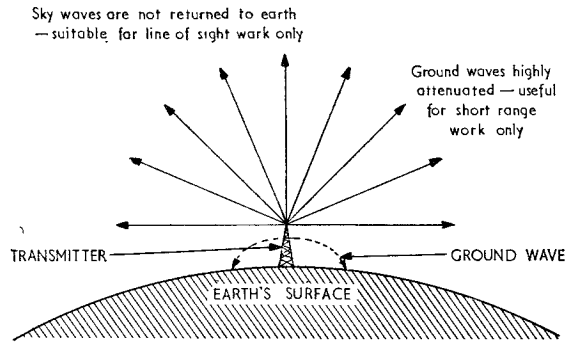

FIG. 6. RADIATION FROM A VHF AERIAL

aircraft. In practice, most air-to-ground communication systems use frequencies in the 100–156 Mc/s (v.h.f.) band and in the 220–400 Mc/s (mainly u.h.f.) band. However, the air-to-ground range depends upon the altitude of the aircraft.

Aerials for VHF, UHF and SHF

Half-wave dipoles are commonly used at v.h.f. and u.h.f. but when a signal with a wide bandwidth is to be handled, special wide-band aerials are used. A common wide-band v.h.f. ground installation aerial consists of several dipoles in parallel (Fig. 7a). Slow-speed aircraft may employ 'blade' or 'whip' aerials as shown in Fig. 7b and c). High speed aircraft use suppressed aerials such as the v.h.f. canopy type, the notch type or slot aerials discussed in Part 1 of these notes.

Two u.h.f. aerials used for ground installations are the discone and biconical aerials shown in Fig. 8. These aerials have a wide bandwidth and the discone aerial can be made of collapsible rods which make it suitable for mobile ground installations.

Yagi arrays make good directional v.h.f. and u.h.f. radiators but at s.h.f. narrower beams can be produced by using solid reflectors. These reflect radio waves at s.h.f. in the same way
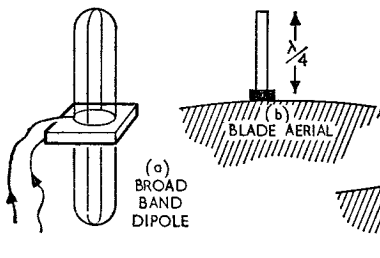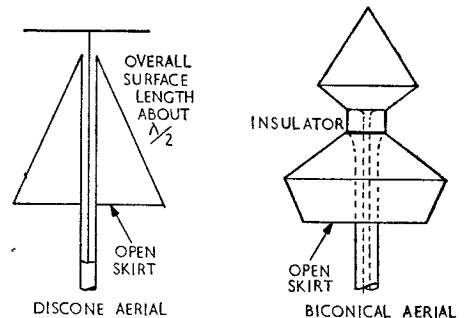

FIG. 7. VHF AERIALS

FIG. 8. UHF AERIALS

that the reflector of a car headlight reflects light waves. Two such reflectors are shown in Fig. 9; they are known as *parabolic* reflectors. The larger the aperture of the reflector the narrower is the radiated beam. On some large ground installations the reflectors are made of wire mesh to reduce wind resistance; to the radio waves they appear solid. The parabolic

(a) PARABOLIC DISH

(b) PARABOLIC CYLINDER

SOLID END

OUTGOING RADIATION
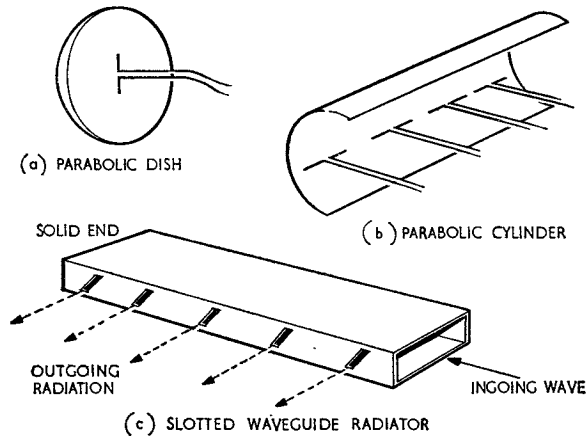
INGOING WAVE

(c) SLOTTED WAVEGUIDE RADIATOR

FIG. 9. SHF AERIALS

cylinder of Fig. 9*b* is shown fed by a row of dipoles. At frequencies above about 3,000 Mc/s it becomes more efficient to feed the aerial from a waveguide than from coaxial transmission line.

If slots are cut in one side of a length of waveguide, the slots will radiate e.m. energy in the same manner as a row of dipoles (Fig. 9*c*). A waveguide radiator can be used to feed a parabolic cylinder reflector thus producing a narrow beam.

Slotted waveguide arrays can be either resonant or non-resonant. Non-resonant slot arrays have a wide bandwidth but the array has a "squint", i.e. the beam is radiated at a small angle (about 5°) to the normal to the waveguide (Fig. 10*b*).

ANGLE OF SQUINT

MAJOR LOBE

SLOTS IN THIS FACE

RESONANT ARRAY
(a)

MAJOR LOBE

SLOTS IN THIS FACE

NON-RESONANT ARRAY
(b)
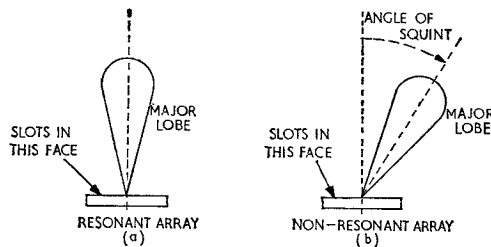
FIG. 10. SQUINT IN A NON-RESONANT SLOTTED ARRAY

The wavelength of the energy being propagated down the waveguide is called the guide wavelength (λ_g) and so that the internal reflections from the slots do not set up standing waves inside the waveguide, the slots are spaced slightly more than 0.5 λ_g apart. In order to avoid reflections from the end of the waveguide a dummy load is inserted to absorb any residual energy.

As the wave travels down the guide it becomes progressively weaker due to the energy being radiated by the slots. To enable the slots to radiate equal energy, it is arranged that the

217

slots at the transmitter end of the guide are loosely coupled to the wave and that slots near the dummy load end are tightly coupled. This is achieved by altering the tilt of the slots as shown in Fig. 11.

FIG. 11. VARIATION OF SLOT TILT

The angle of tilt is increased as the end of the guide is approached thus increasing the coupling. Alternate left-hand and right-hand tilts ensure that the slots radiate almost in phase by allowing for the reversal of phase of the wave every half-wavelength.
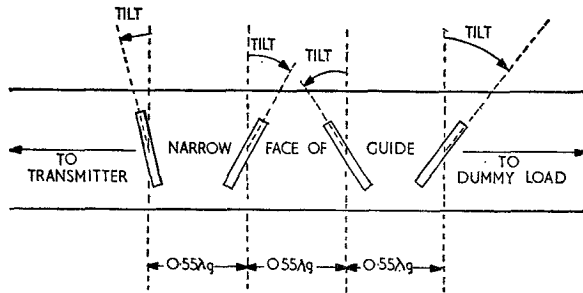
CHAPTER 4

SCATTER PROPAGATION

Introduction

The simple theory of the propagation of radio waves explains how h.f. waves are reflected by the ionosphere to render long-distance communication possible. Ionospheric reflection is used a great deal for point-to-point communication with low-power transmitters. The frequency range employed is about 6 to 30 Mc/s and frequency may have to be changed three or four times during 24 hours to deal with varying ionospheric conditions. Waves above about 30 Mc/s pass through the ionosphere and communication at these frequencies is limited to 'line-of-sight' distances.

It has been found, however, that signals are received at frequencies above 30 Mc/s at distances which are far greater than the line-of-sight distance. These signals, although weak, are thousands of times stronger than the elementary theory envisages and they are now considered to be due to *scattering* in the ionosphere or troposphere. The signals are always present although they fluctuate fairly rapidly and they do not disappear during ionospheric storms as h.f. signals do; they are therefore able to give continuous communication.

Types of Scatter

There are two main types of scatter, as follows:—

a. Back scatter. This is not used for communication but has been employed to make propagation measurements.

b. Forward scatter. This type of scatter is used for communication. It may be divided into:—

    i. Ionospheric scatter, a development of which is the Janet system.

    ii. Tropospheric scatter in which we may include stratospheric scatter for completeness.

Scattering Mechanism

When a radio wave strikes an object which is comparable in size with the wavelength of the wave, some of the wave energy is reflected in various directions away from the main beam. Fig. 1 shows a beam passing through a number of small objects.

Most of the energy passes straight on, but Fig. 1 shows how small amounts of energy are reflected in a number of directions, including backwards.
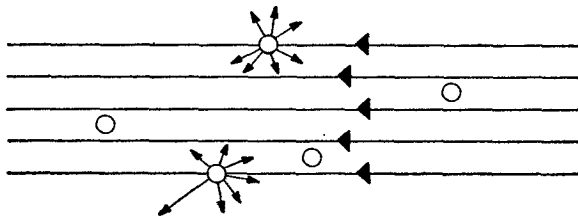

FIG. I. HOW SCATTERING OCCURS

This phenomenon is scattering. With light waves it accounts for the blue colour of the sky, due to scattering by dust particles and water vapour molecules; for the blue colour of tobacco smoke, and for other phenomena. It is emphasised that the particles must be comparable with

the wavelength; if they are too small they do not affect the wave noticeably; if they are too large ordinary reflection and refraction occurs.

Although particles have been specified, only the presence of something with a different refractive index is necessary for scattering to occur. For ionospheric scattering it is thought that the scattering elements are 'blobs' of different ionisation occurring in the ionosphere. These irregularities are superimposed on the normal ionisation.

For tropospheric and stratospheric scattering the blobs are of different refractive index due to variations in water vapour content and air density.

Back Scatter

Back scatter can occur with h.f. waves below the critical frequency; here the scattering is produced by irregularities on the earth's surface (Fig. 2).

Some of the energy travels back to the transmitter, as shown by the dotted path, and can cause interference.
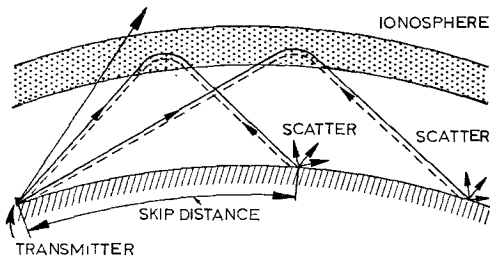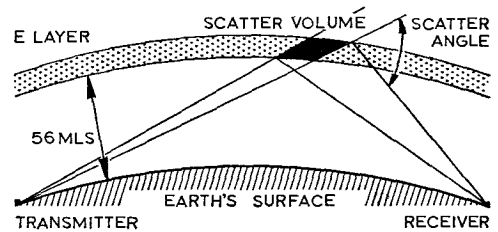

FIG. 2. BACK SCATTER

FIG. 3. SCATTER VOLUME

Back scatter may be used as a basis for making propagation measurements. Pulses are radiated from a directional aerial and the echoes are shown on a c.r.t.; the frequency is increased to the m.u.f. and there is then no echo. The corresponding distance is the skip distance. By using a rotating aerial and a p.p.i. display, a picture of large scale ionospheric movements is obtained.

Ionospheric Forward Scatter

Forward scatter is used for communication. A highly directional transmitter aerial intersects the line from the receiver in the scatter volume in the ionosphere (Fig. 3).

The scattering blobs of ionisation in the scatter volume are caused by solar ultraviolet and corpuscular radiation, and by meteors. Most of the energy from the transmitter passes through the ionosphere but a small amount is scattered and picked up by the receiver. High transmitter powers (40kW) are necessary, with directional aerials at transmitter and receiver, and the best results are obtained when the scatter angle is small, 2° being a typical value. The maximum range occurs when the transmitter beam is directed at the horizon and is about 1,250 miles; the minimum range occurs when the scattered signal becomes too weak to be useful and is about 600 miles.

The strength of the signal decreases as frequency increases and hence the useful range of frequency is restricted from 30 to 60 Mc/s.

Since scattering is due to blobs of ionisation, signal strength at the receiver is enhanced when ionospheric storms occur, and communication can be maintained when normal h.f. stations are 'blanked out'. The received signal depends on the average amount of ionisation and hence varies with the usual diurnal, seasonal and annual solar activity. In particular, there is a marked

minimum each day at about 2000 hours, due to cessation of solar radiation, and a slow increase during the night, due to increasing meteoric activity.

The received signal is the aggregate of a large number of components of random phase emanating from different parts of the scatter volume and thus the signal fluctuates fairly rapidly (between 0.2 and 5 times per second). This fluctuation is combated by a.g.c. and diversity reception. Meteors produce strong ionisation and signals are reflected rather than scattered from meteor trails. These reflections interfere with the normal scattered signals and cause variation in signal strength, often with Doppler frequency shifts, producing heterodyne whistles. This can be combated by using s.s.b. transmission.

One form of diversity reception used by the GPO is 'wave angle' diversity (Fig. 4). Two receiving aerials point a little to one side of the direct transmission path and collect from different parts of the scatter volume. Yagi aerials are used to give increased directivity.

FIG. 4. WAVE ANGLE DIVERSITY

Janet System

One of the contributions to ionospheric scatter is from ionised meteor trails at 40 to 62 miles in height, which are at the right point to cause reflection. The contribution is sporadic but signals due to meteor trails are always strong and sometimes last several seconds. If transmission occurs only when suitable meteors are present, a low-power transmitter can be used. This is the basis of the Janet system. The path length of the Janet system is 600 miles and the frequency 40 Mc/s.

Calculations show that communication by meteors should be possible on an average of 5% of the time, and the system was designed to operate a duplex 60 words-per-minute radio teleprinter channel. As transmission occurs only when suitable meteors are present, to achieve an average of 60 w.p.m. a sending rate of 1,200 w.p.m. is necessary.

The incoming teleprinter message is converted to punched paper tape and stored to await transmission. When the signal strength is high enough the tape is read at 120 characters per second by a photo-electric reader, and transmitted. At the same time, the incoming message is detected and written on 5 channels on magnetic tape which travels at 5 inches per second. The tape is stored and read at 60 w.p.m. to produce a teleprinter line signal.

The transmitter radiates continuously and the a.g.c. voltage level in the receivers is compared with the noise level. When the signal/noise ratio is favourable (12 db at least), communication is started by allowing the main receiver detectors to operate; at the same time, the photo-electric reader is brought into operation.

The information is transmitted in code, five digits being used for each character, in the form of two-position pulse position modulation. A pulse in the first position represents 'space', in the second position 'mark'. The pulses are cosine-squared shaped (Fig. 5).
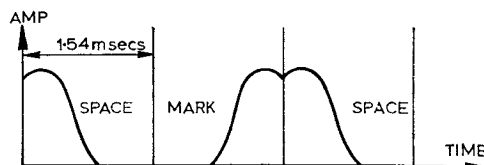

FIG. 5. JANET CODE PULSES

221

The fundamental frequency of the pulse is 650 c/s and so the total double-sideband transmission bandwidth is 2,600 c/s. A block diagram of the Janet equipment is shown in Fig. 6.

The magnetic tape unit contains 24 digits to the inch and will store 40 feet of tape, corresponding to 1.5 minutes transmission time. Special design has produced acceleration times for the tape drum from stop to full speed in 3 milliseconds.
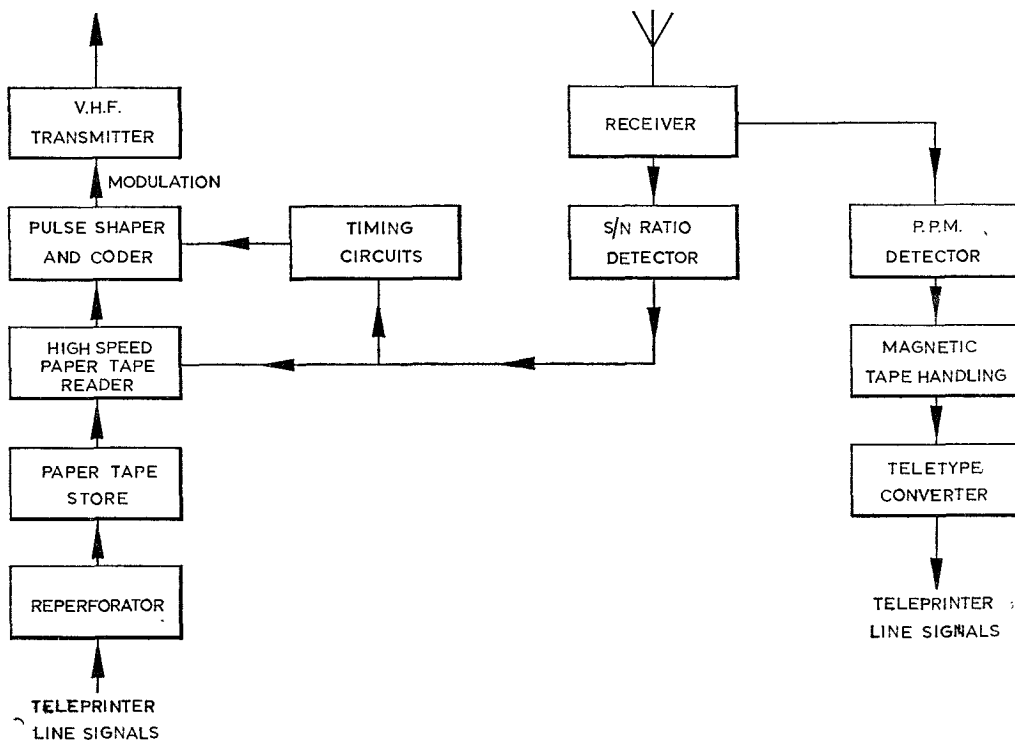

FIG. 6. BLOCK DIAGRAM OF JANET EQUIPMENT

Since communication takes place in short bursts of average time 1 second, speed in switching on and off is essential and the whole equipment takes less than 7 milliseconds to start or stop.

The aerial system is a Yagi type and the transmitter power is 500 watts.

Tropospheric Forward Scatter

With tropospheric scatter the scattering mechanism is 'blobs' in the troposphere, caused by winds and earth heating. The scattering communication process is similar to that shown in Fig. 3 except that the scatter volume is now at a maximum height of about 40,000 feet. The range is therefore less, typical values being between 75 and 400 miles.

High transmitter power (about 40 kW) and directional transmitter and receiver aerials are necessary as in the case of ionospheric scatter. The frequencies used, however, are very different and range from about 100 Mc/s to 10,000 Mc/s, the best frequencies being between 300 and 2,000 Mc/s.

Bandwidths of 3 to 4 Mc/s can be transmitted up to 200 miles if very large, narrow-beam aerials are used with high transmitter powers. The useful bandwidth decreases considerably as range increases.

The best results are obtained when the scatter angle is small and in practice beams are directed at the horizon. A site free of obstructions is necessary and advantage is gained if an elevated site is used with beams directed slightly downwards.

Reflections from aircraft are troublesome, and Doppler beat frequencies are produced similar to those from meteor trails in ionospheric scatter.

Since the 'blobs' causing scattering are caused by meteorological factors, weather conditions affect the transmission. For example, signals are enhanced in fog or mist but are reduced in heavy rain or snow.

Tropospheric scatter systems are in use in various parts of the world. They are not as good as microwave links from the point of view of cost, ease of insertion of channels, or quality of signal but tropospheric scatter communication is useful for over-sea paths and over inaccessible land. Scatter systems cause interference because of their long range and t.v. services already suffer.

The high frequencies associated with tropospheric scatter enable parabolic aerials to be used to give narrow beams. Common aerial diameters are 60 feet and 30 feet. Aerials with 120 feet diameters have been used but do not give greatly improved performance, since if the beam is too narrow the scatter volume becomes rather small. The aerials must be accurately made and must be able to withstand high winds; their cost may represent a large part of the total cost of the installation. With lower frequencies it is possible to build the paraboloid into the main building by shaping one wall. The same aerials are used for transmission and reception and TR switches, which prevent the transmitted signal entering the receiver and the received signal entering the transmitter, are used. Coaxial lines or waveguides are used as feeders, according to the frequency.

To give reliable communication, diversity reception is usual, using two aerials spaced about one hundred feet apart. A block diagram of a dual-diversity installation is shown in Fig. 7.

In the system shown, frequency modulation is used. This is common, but s.s.b. may also be used. The transmitted signal frequency, f_1, is slightly different from the received signal, f_2.
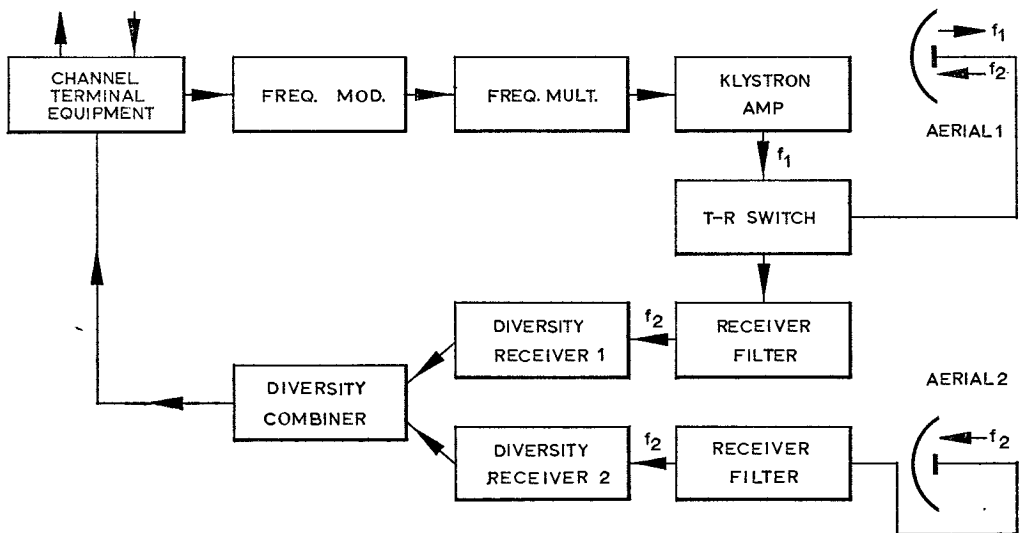

FIG. 7. DUAL DIVERSITY SYSTEM

The diversity combiner compares the signals from the two diversity receivers and selects the one with the better signal/noise ratio.

The transmitter klystron amplifier uses large four-cavity water-cooled klystrons as r.f. power amplifiers. The receivers are carefully designed for sensitivity and low noise factor.

A further increase in reliability of communication is obtained by using *quadruple diversity*. Such a system is shown in block form in Fig. 8.
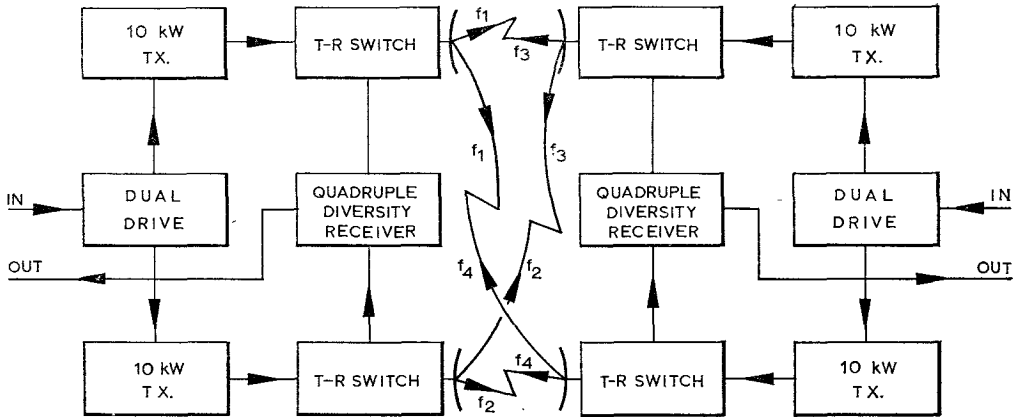

FIG. 8. QUADRUPLE DIVERSITY

There are two transmitters at each installation radiating from spaced aerials on slightly different frequencies. Each aerial receives both frequencies transmitted by the other station and the four received signals are then routed to four separate receivers, comprising the 'quadruple diversity receiver'. As before, the signal with the best signal/noise ratio is selected. It is found that the polarisation of the scattered signal is not affected by the scattering process and so it is possible to use only two frequencies with polarisation at right angles, giving polarisation diversity as well as spaced-aerial diversity.

Without diversity reception a satisfactory signal/noise ratio can be obtained for 95% of the time; with dual diversity, 98%; and with quadruple diversity 99%.

Stratospheric Scatter

Occasionally, signals have been received at distances up to 600 miles at frequencies associated with tropospheric scatter. It is thought that these may be due to scattering from meteorological 'blobs' in the stratosphere but there is little information available as yet.

Summary

The main advantage of scatter communication over h.f. multi-hop communication is its greater reliability. When ionospheric storms, sunspot activity and solar flares upset the normal state of the ionosphere, normal long-distance h.f. communication is frequently blanked out. Under the same conditions communication employing scatter propagation is unaffected. The most serious disadvantages are the restricted range and the high cost. Tropospheric and Janet systems seem to have the greatest possibilities.

SECTION 6

RADIO NAVIGATIONAL AIDS

CHAPTER 1

PRINCIPLES OF DIRECTION FINDING

Introduction

Direction finding (d.f.) by means of radio waves is a very valuable and widely used navigational aid. By using special d.f. aerials, an aircraft or marine navigator can determine the bearing of his moving station with respect to a fixed transmitter: with a further bearing taken on another fixed station he can 'fix' his position on a map. By turning his aircraft so that its bearing corresponds to the bearing from a transmitter, the navigator can 'home' to the transmitter. Alternatively, a ground station fitted with d.f. aerials can give an aircraft bearings taken from the aircraft's transmissions and two or more such ground stations can be used to fix the position of an aircraft.

In this chapter we shall consider briefly the principles on which these radio navigational aids are based. In the following chapters of this section we shall see how these and other principles are applied in radio navigational equipments.

The Loop Aerial

A simple loop aerial consists of a number of turns of wire wound round an insulating frame. An alternating voltage induced in the loop causes circulating currents which in turn cause a voltage at the receiver input (Fig. 1). This voltage is amplified by the receiver in the normal manner and the receiver output voltage is applied to a device such as a telephone receiver, a meter or a c.r.t. which indicates the amplitude of voltage across the loop terminals.
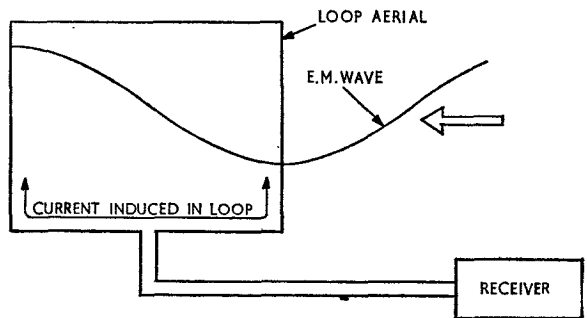

FIG. 1. SIMPLE DF AERIAL

The voltage across the loop terminals depends upon the *relative phase* of the voltages induced in each upright limb. If the plane of the loop is in line with the direction of the received e.m. wave as shown in Fig. 2a there will be a maximum phase difference between the voltage induced in AB and that induced in CD. These two voltages act in opposition around the loop and the resultant voltage across the loop terminals is the *difference* between them. In this position of the loop the resultant voltage is a maximum. As the plane of the loop is moved through 90°, the
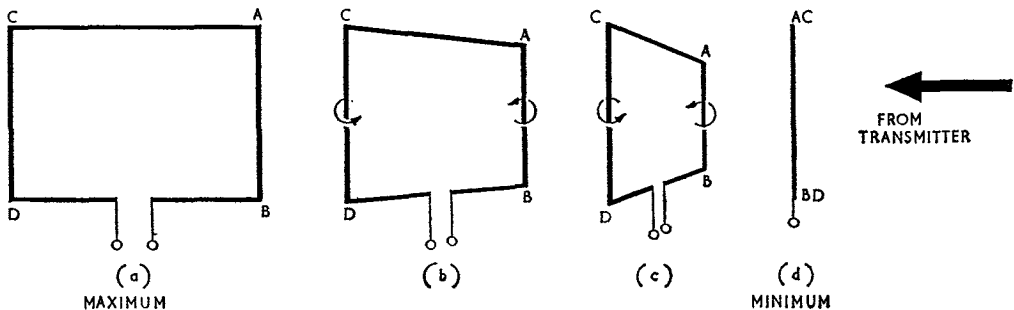

FIG. 2. VOLTAGE ACROSS LOOP TERMINALS

phase difference in AB and CD decreases as does the voltage across the loop terminals until, when the plane of the loop is broadside on to the transmitter, the voltage across the loop terminals is zero.

Loop Polar Diagram

Since the voltage across the loop terminals varies from a maximum to a minimum as the loop is rotated through 90° as described, the polar diagram of the loop will have *two* maxima and *two* minima as shown in Fig. 3. To find the direction in which a transmitter lies, the loop may be rotated until either a maximum or a minimum voltage is indicated. In practice the direction of the transmitter is found by rotating the loop until a *minimum* voltage (or 'null') is indicated because a null is easier to detect and is more sharply defined than a maximum voltage.

Sense-finding

Since there are two maximum and two minimum positions for the loop, the transmitter may possibly lie in one of two opposite directions. This ambiguity is resolved by using a simple omni-directional aerial, known as a *sense* aerial, placed between the two vertical limbs of the loop as shown in Fig. 4.
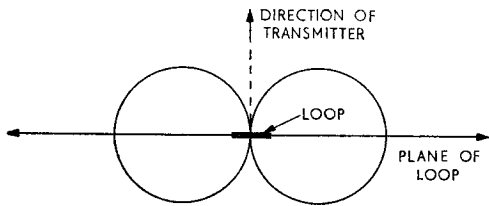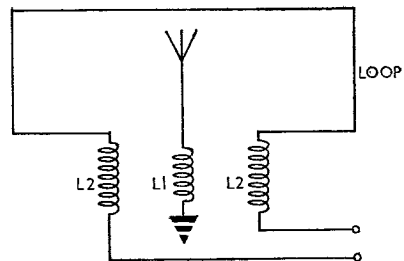

FIG. 3. LOOP POLAR DIAGRAM

FIG. 4. DETERMINATION OF SENSE

The voltage induced in the omni-directional aerial is combined with the loop voltage via the r.f. transformer L_1L_2. The phase relationships are such that the omni-directional aerial signal increases the output from the loop in the direction of one maximum and reduces it in the direction of the other maximum. If the dimensions of the omni-directional aerial and the coupling of the r.f. transformer are correctly chosen, one maximum is completely cancelled and the resultant polar diagram is as shown in Fig. 5. With this heart-shaped polar diagram (known as a *cardioid*) only one minimum is present and ambiguity is avoided.
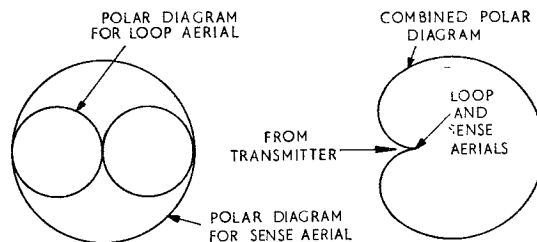

FIG. 5. COMBINED POLAR DIAGRAM OF DF AND SENSE AERIALS

As the minimum obtained when the sense aerial is used is less pronounced than the minimum from the loop alone, the sense aerial is used only to find the general direction of the transmitter *after* an accurate bearing has been obtained using the loop alone.

Thus the method of finding the direction of a transmitter is to rotate the loop, with the sense aerial switched out, until a zero signal is obtained. The sense aerial is then switched in and the loop rotated 90° in a certain direction. If the signal falls to zero again the transmitter lies in one direction; if it increases in strength the transmitter is in the opposite direction.

Aircraft Loop Aerials

The type of loop aerial used on aircraft depends on the band of frequencies over which the loop is to operate. At m.f., the d.f. aerial usually consists of a rigid framework of insulating material over which is wound 15 to 20 turns of wire so that a usable signal is picked up, the loop terminals being connected to the receiver through slip rings. The loop is usually encased in a plastic streamlined shell for protection.

For h.f. operation the number of turns of wire required is much smaller. The reason for this is that because the wavelength is shorter the loop dimensions are a greater fraction of the wavelength and hence fewer turns are necessary to produce a workable signal. Also, by using fewer turns, the interturns capacity of the loop is reduced, resulting in smaller losses. The loop is encased in, but insulated from, a metal tube which acts as a shield and reduces precipitation noise caused by rain and dust particles. The tube is earthed, thereby equalising stray capacitance between each vertical limb of the loop and earth, so increasing the accuracy of bearings. If the tube completely surrounded the loop, the loop would be completely screened from the signal. This is avoided by making part of the casing of an insulating material. Typical m.f. and h.f. aircraft loops are shown in Fig. 6.
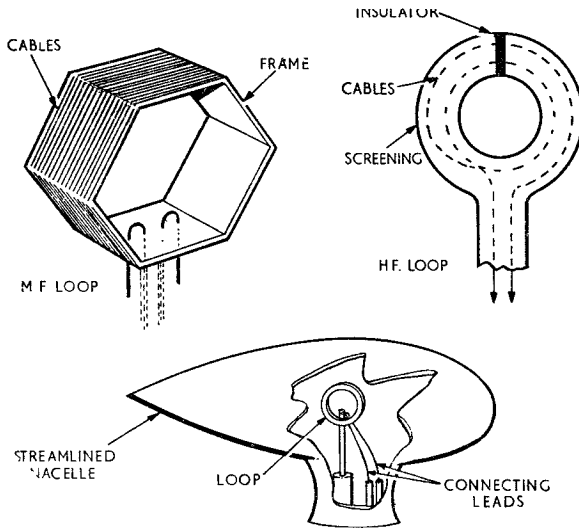

FIG. 6. AIRCRAFT MF AND HF DF AERIALS
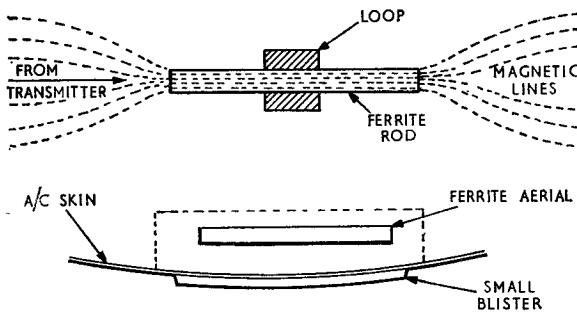
FIG. 7. FERRITE-CORED LOOP

The Ferrite-cored Loop

On modern high-speed aircraft the loops shown in Fig. 6 would considerably affect the aerodynamic properties of the aircraft. A much smaller ferrite-cored loop can be mounted in a 'blister' projecting only an inch or so beyond the aircraft skin as shown in Fig. 7. The ferrite core concentrates the magnetic field into a very small area, thus making a few turns of wire wound around a small rod of ferrite much more effective than a large wire aerial. The ferrite-cored loop is directional because maximum voltage is induced into the loop when the ferrite rod is in line with the lines of the magnetic field.

Ground Station DF Aerials

Ground station m.f. and h.f. direction finding aerials are fairly large structures and would be difficult to rotate. Thus, instead of a loop aerial, the

(*A.L. 3, September,* 1964)

installation usually consists of fixed vertical aerials as shown in Fig. 8. Each pair of aerials is connected by screened underground cables to two fixed coils L_1 and L_2 mounted at right angles to each other. A small coil L_3, known as the *search coil*, is coupled to L_1 and L_2 and is free to rotate through 360°. The ends of the search coil are connected to a receiver. This system of coils is called a *radio goniometer* and the aerial system is known as an *Adcock aerial.*
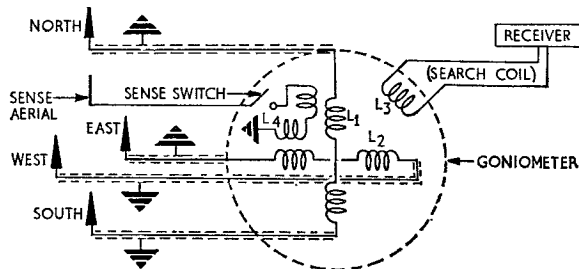

FIG. 8. HF DF AERIAL SYSTEM

The two pairs of fixed aerials are equivalent to the vertical limbs of two loop aerials mounted at 90° to each other. In Fig. 8 the north-south aerials are connected to coil L_1 and the east-west aerials to coil L_2. The e.m.f. across L_1 depends upon the resultant voltage induced in the north-south aerials and that across L_2 depends upon the resultant voltage of the east-west aerials. These voltages in turn depend upon the direction of the transmitter. Thus the strength of the magnetic fields in L_1 and L_2 depends upon the angle the transmitter makes with the aerial assembly. These two fields combine to produce a resultant field, the axis of which makes the same angle with L_1 and L_2 as the direction the transmitter makes with the aerials. By aligning the search coil with this magnetic field and incorporating a calibrated scale, the search coil behaves as a rotating loop and will indicate two maxima and two minima.

A sense aerial mounted equidistant from the d.f. aerials is coupled, via a switch and a further coil (L_4), to L_1 and L_2. Thus the voltage from the sense aerial produces a magnetic field which opposes or adds to the field due to L_1 and L_2. With the sense aerial switched in the aerial assembly has a polar diagram with a single minimum.

VHF DF Ground Station Aerials

The aerial system of a v.h.f. d.f. station is a fairly small structure, easy to rotate. As with m.f. and h.f. systems, a v.h.f. d.f. station can provide "homing" and bearing facilities and three such stations provide a "fixer" service. The frequency range covered is 100 to 156 Mc/s.

A typical v.h.f. d.f. aerial system is shown in Fig. 9. It consists of two vertical dipoles
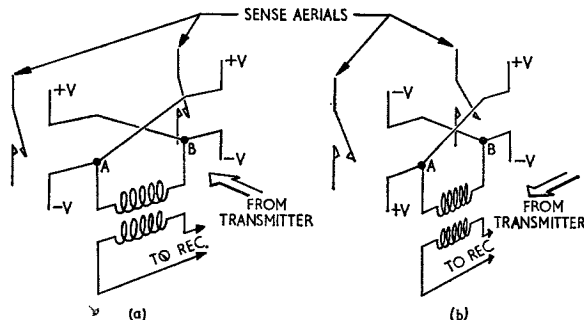

FIG. 9. VHF DF AERIAL SYSTEM

228

spaced a half-wavelength apart and mounted so that the assembly can be rotated. The upper half of one dipole is connected to the lower half of the other. Coupling to the receiver input is via a r.f. transformer. Two reflectors, which can be switched into or out of use, are mounted behind the dipoles.

When the transmitter lies in a direction at right angles to the plane of the aerials, the dipole voltages are equal (Fig. 9a). The polarities are such that current leaving point A equals current leaving point B, and as they flow in opposite directions through the transformer primary the voltage in the secondary is zero.

When the transmitter lies in the plane of the array (Fig. 9b), the dipole voltages are no longer in phase and maximum voltage is produced at the receiver input.

The two reflectors, one behind each dipole, are used to determine the sense of the bearing. If the dipoles are nearer the transmitter than are the reflectors, the aerial gain and thus the receiver input, is increased when the reflectors are switched in (Fig. 10a). The opposite happens if the reflectors are in front of the dipoles as in Fig. 10b.

Fig. 11 shows the difference between the polar diagrams when the reflectors are switched in and out.
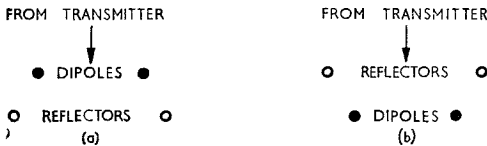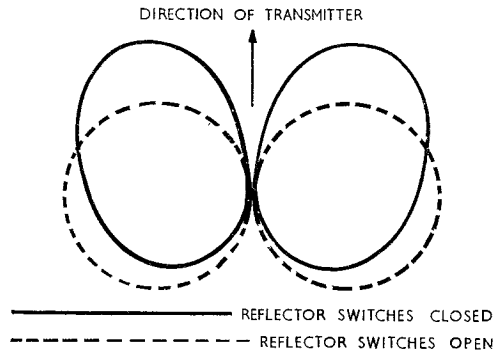

FIG. 10. SENSE DETERMINATION

FIG. 11. POLAR DIAGRAMS

DF Errors

Bearings taken by the d.f. systems so far discussed are subject to certain errors. Some of these errors can be corrected and others cannot.

Quadrantal Error

When a radio wave strikes a metal object near a d.f. aerial the wave is re-radiated. This re-radiated wave induces small voltages in the d.f. aerial which result in incorrect bearings. This type of error is termed *quadrantal error*. The error is maximum when the signal arrives from a direction along the middle of each quadrant (Fig. 12a).

To compensate for quadrantal error, d.f. stations are calibrated on installation and supplied with a correction chart similar to that shown in Fig. 12b. By means of this chart the bearing indicated at the station can be corrected to give the true bearing of the transmitter. With an automatic d.f. system such as an airborne radio compass, pre-set compensation is provided and the bearing indication is automatically adjusted to compensate for quadrantal error.
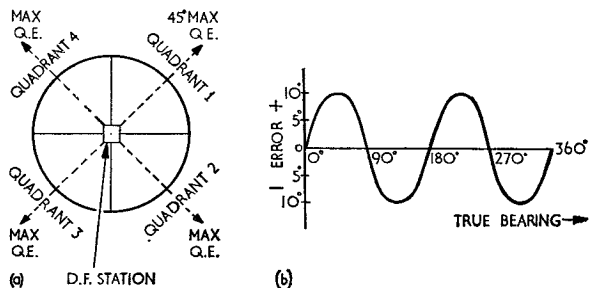

FIG. 12. QUADRANTAL ERROR

Night Effect

It has been assumed that a radio wave which strikes a loop aerial is vertically polarised and travels parallel to the earth's surface. Under certain conditions, however, the wave arriving at the d.f. aerial consists of both ground wave and sky wave. The ground wave which is parallel to the earth's surface induces voltages in the *vertical* limbs of the loop. The sky wave, because of the angle at which it strikes the loop, induces voltages in the *horizontal* parts of the loop and this gives rise to incorrect readings. Sky waves at m.f. are more prevalent at night; hence the term *night effect* is used to describe this error. Radiation from an aircraft to a ground d.f. station has the same effect as a wave reflected from the ionosphere and so an alternative name for the error caused is *aeroplane effect.*

The effect is cured in the Adcock aerial system by removing the top horizontal part of the loop and cancelling the effect of the lower horizontal parts by screening them or by running them together so that the currents cancel. There are various forms of Adcock aerial but they are all designed to cancel the effects of the horizontal members. In airborne loop d.f. systems a "flat" loop 2 or 3 inches high minimises the effect and by selecting frequencies which are not reflected from the ionosphere the effect is further reduced. This means using m.f. stations for long range d.f. and v.h.f. or u.h.f. stations for short range.

Vertical Effect

Unequal stray capacitances to earth between the sides of the loop and nearby earthed conductors produce an output voltage across the loop terminals when the loop is in the null position. This results in either an indistinct or an incorrect minimum. This effect can be reduced in the following ways:—

a. By adjusting a differential capacitor connected across the loop as shown in Fig. 13.

b. By enclosing the loop in an earthed shield as previously described (p. 227). It should be noted that the loop is not screened from the radio wave because the shield is broken by an insulator but if the shield is earthed opposite the gap, capacitances between the loop and the earthed shield are equalised.

Coastal Refraction

Radio waves are bent (or refracted) as they travel across a coastline and this gives rise to a bearing error as shown in Fig. 14.
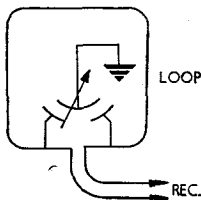
FIG. 13. BALANCING THE
LOOP TO EARTH

FIG. 14. COASTAL REFRACTION
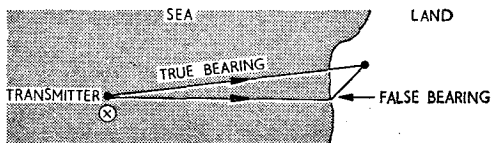
If a radio wave is transmitted from an aircraft at X, the wave is refracted at the coastline giving rise to a bearing error. In the diagram the error is exaggerated. The error is maximum when the radio wave crosses the coastline at an acute angle. There is no remedy for this error and any bearings taken when a radio wave is known to have crossed a coastline must be regarded as suspect.

CHAPTER 2
AIRBORNE NAVIGATIONAL AIDS

Introduction

Navigational aids intended for use by aircraft can be divided into three categories. There is the type which employs an omni-directional ground transmitter with special d.f. equipment carried in the aircraft. An example in this category of airborne equipment is the radio compass or a.d.f. equipment.

In the second category is the d.f. system in which the special d.f. equipment is a ground installation and a non-directional receiver and transmitter are carried in the aircraft. C.a.d.f. and c.r.d.f. come into this category.

Thirdly there are systems, such as ILS, in which both ground and airborne directional equipments are employed.

The Radio Compass

Manually operated d.f. systems employing a loop aerial are subject to interpretation and manipulation errors. The time required to take a bearing by manually rotating a loop to the null position is too great for navigation in high-speed aircraft. Most aircraft are nowadays fitted with an *automatic direction finding* (a.d.f.) equipment, often called a *radio compass*. The frequency coverage is usually from 150 to 2,000 kc/s. With a radio compass the bearing indication can be continuous, i.e. once a bearing is indicated any change in bearing due to movement of the aircraft or transmitter is automatically shown by the indicator.

In the radio compass the loop is turned by an electric motor when there is a loop signal. When the loop is in the null position the loop signal is zero and the motor stops. The direction of rotation of the motor and hence of the loop is controlled by the relative phasing of loop and sense aerial signals.

With a.d.f. there is no ambiguity as sensing is an integral part of the a.d.f. process. After suitable treatment the loop and sense aerial signals are applied to the electric motor in such a way that if the transmitter is to the right of the correct null position the loop turns right, and vice versa. Thus the operator has only to tune the receiver to the frequency of the required transmitter, switch over to a.d.f. and read off the bearing. Quadrantal error is automatically corrected.

The essential circuits of an a.d.f. equipment are shown in Fig. 1. For sensing purposes it is necessary to phase shift either the loop aerial or sense aerial voltage so that they are either in phase or in antiphase, depending on the loop position relative to the null position. In a.d.f. systems it is usual to phase shift the loop voltage by 90°.
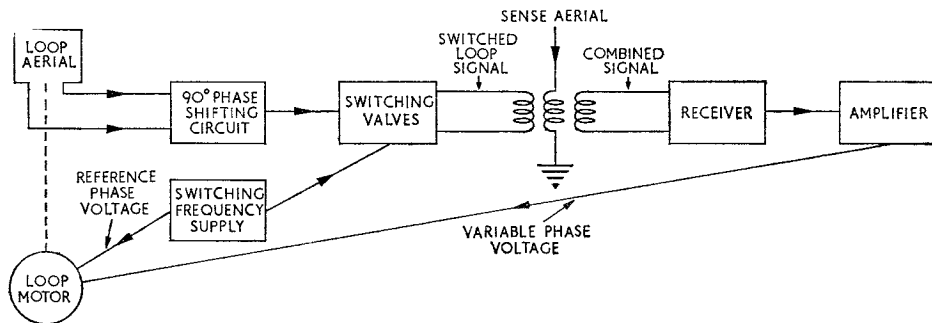

FIG. 1. OUTLINE OF ADF CIRCUITS

In Fig. 1 the loop aerial output is applied to an amplifier which increases the amplitude of the loop signal and also shifts its phase by 90°. This voltage is then applied to the grids of two switching valves which are alternately cut on and off by a low frequency square wave switching voltage obtained from a switching frequency supply circuit. The output from the switching valves is a signal which is alternately in phase with and in antiphase with the sense aerial voltage, the phase changing at each half cycle of the switching voltage (Fig. 2).
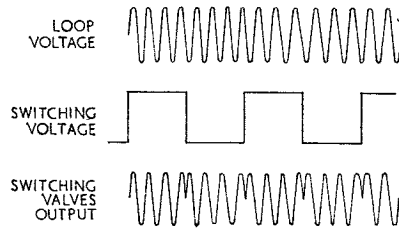

FIG. 2. LOOP SWITCHING ACTION

This voltage combines with the sense aerial voltage in the r.f. transformer. The resultant voltage applied to the receiver input is thus amplitude modulated at the switching frequency. The receiver extracts the switching frequency and it is fed via an amplifier to the loop motor. This is a two phase a.c. motor requiring two inputs of the same frequency but 90° out of phase with each other to make it turn. Both square wave and sinusoidal inputs will operate the motor.

In the a.d.f. equipment one of the motor inputs is obtained from the amplifier output and is called the *variable phase* input. The other input, called the *reference phase* input, is obtained directly from the switching frequency supply circuit. Depending on the position of the loop relative to the null position, the variable phase input either leads or lags the reference phase input by 90°. If there is a 90° lead the motor turns in one direction; if a 90° lag it turns in the opposite direction. If there is no variable phase input the motor does not move.

Fig. 3a shows the waveforms at various stages in the a.d.f. equipment for a transmitter which is to the right of the loop null position and Fig. 3b shows the waveforms when the transmitter is to the left. We have seen how waveform (*iii*) is produced from the loop signal and the switching voltage. Waveform (*v*) is produced by combining the switching valves' output with the sense

(I) LOOP INPUT TO SWITCHING VALVES

(II) SWITCHING VOLTAGE

(III) SWITCHING VALVES OUTPUT

(IV) SENSE AERIAL SIGNAL

(V) RESULTANT INPUT TO RECEIVER (III + IV)

(VI) RECEIVER OUTPUT

(VII) VARIABLE PHASE FROM AMPLIFIER TO MOTOR

(VIII) REFERENCE PHASE TO MOTOR

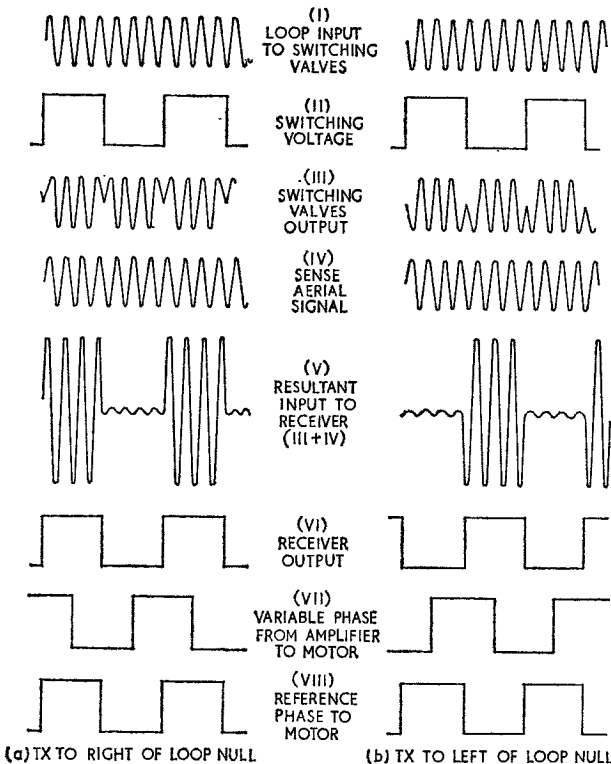(a) TX TO RIGHT OF LOOP NULL   (b) TX TO LEFT OF LOOP NULL

FIG. 3. ADF WAVEFORMS

aerial signal. When this is demodulated in the receiver, waveform (*vi*) is obtained. This is applied to an amplifier which shifts the phase by 90° as in waveform (*vii*). Notice that waveform (*a–vii*) and (*b–vii*) are 180° out of phase. Thus if waveform (*a–vii*) leads the reference waveform by 90°, waveform (*b–vii*) lags it by 90°. Thus, depending on the position of the transmitter relative to the loop null position, the motor will turn clockwise or anticlockwise, moving the loop towards the null position. When this position is reached the loop output is zero and as the variable phase input to the motor is thus zero, the motor stops in the null position.

The relative bearing of the transmitter as indicated by the plane of the loop with respect to the fore and aft axis of the aircraft is transmitted via Dessyn transmission systems to bearing indicators mounted in convenient positions in the aircraft.

Phase Shift DF System

This is a method of direction finding used to enable an aircraft to "home" on to a transmitter and employs frequencies in the v.h.f. and u.h.f. bands. Two fixed aerials, one on the starboard and the other on the port side of the aircraft, are used as the vertical limbs of a loop aerial. The phase difference between voltages induced into these aerials is used to indicate the direction of the transmitter. A sense aerial is not used. The two aerials are connected to a *phase shift network* which causes a signal passing through it to be phase retarded by a constant angle $\theta°$ (Fig. 4*a*).
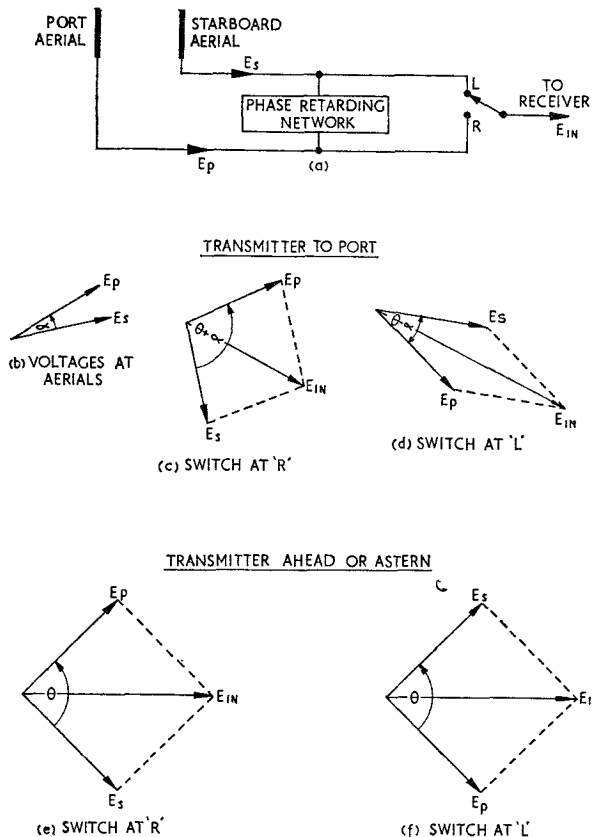

FIG. 4. PHASE SHIFT PRINCIPLE

A L/R switch enables either aerial to be connected directly to the receiver input while the other aerial is connected through the phase retarding network to the receiver input.

When the transmitter is to port of the aircraft there is a phase difference (say $\alpha°$) between voltages in the two aerials (Fig. 4b). With the switch in the 'R' position, the voltage at the switch contact is the vector sum of E_p and E_s and the phase angle between these voltages is $(\theta + \alpha)°$. The voltage applied to the receiver (E_{in}) is as shown in Fig. 4c. If the switch is now put to the 'L' position the phase angle between the two voltages at the switch contact is $(\theta - \alpha)°$ and E_{in} is *larger* than it was when the switch was in the 'R' position (Fig. 4d).

If the transmitter is to starboard the larger input voltage will be received with the switch in the 'R' position.

When the transmitter is dead ahead or astern of the aircraft the phase angle α is zero and E_{in} is of the same amplitude in either switch position (Fig. 4e and f). If the pilot turns the aircraft to port a few degrees and the stronger signal is then received with the switch at 'R' the transmitter is ahead and the ambiguity is resolved.

The resultant signal from the two aerials can be used to operate an indicator which shows in which position of the L/R switch the larger amplitude signal is being received, i.e. on which side of the aircraft the transmitter is situated. To "home" to the transmitter, the pilot turns the aircraft left or right as indicated.

In practice the L/R switch is a multivibrator and no manual operation of the switch is needed. By switching out the port aerial and substituting an aerial mounted under the fuselage, elevation indications can be obtained. A simplified block diagram of a typical phase shift system is shown in Fig. 5. The L/R switch of the basic circuit (Fig. 4a) is replaced by switching valves V_1 and V_2 which are alternately cut on and off by a switching voltage obtained from the multivibrator.
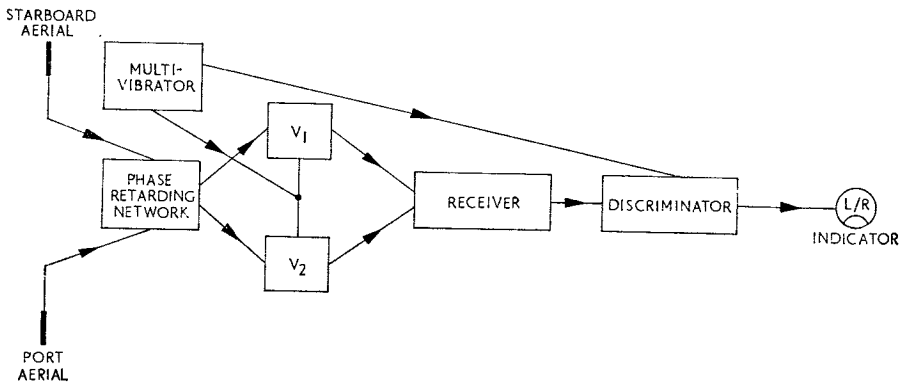

FIG. 5. SIMPLIFIED PHASE SHIFT DF SYSTEM

When V_2 is conducting we have the same conditions as when the L/R switch of the basic circuit is in the 'L' position; when V_1 conducts, the switch is in the 'R' position.

If a transmitter is to port of the aircraft the input to the receiver is as shown in Fig. 6b. The receiver amplifies this input and extracts the square wave modulation (Fig. 6c). This waveform is applied to the discriminator, along with an output from the multivibrator. The discriminator output is a direct voltage the polarity of which depends upon the *phase* of the voltage input to the receiver.

Thus in Fig. 6d (transmitter to port) the discriminator output is a positive d.c. while in Fig. 6h (transmitter to starboard) the output is a negative d.c. This voltage is fed to a centre zero

meter. When the transmitter is to port of the aircraft, the polarity of the applied voltage is such that the meter needle moves to the left; when starboard, the needle moves to the right; and when dead ahead or astern the needle remains central.
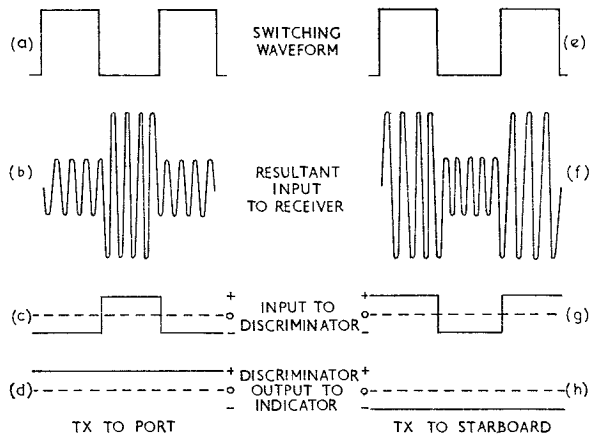

FIG. 6. PHASE SHIFT WAVEFORMS

The Radio Altimeter

An instrument used in an aircraft to indicate height is called an *altimeter*. For normal flight a barometric altimeter, which depends for its action on the decrease in atmospheric pressure with height, is used.   For low-level flying (between 0 and 5,000 feet), an accurate indication of the aircraft's height above the ground is provided by a *radio altimeter*.   There are several types of radio altimeter in use but we shall consider the *frequency modulated* radio altimeter.

In this type of radio altimeter the frequency of an airborne transmitter is made to vary at a regular rate.   This frequency modulated radio wave is transmitted downwards from the aircraft. The wave reflected from the earth's surface is received in the aircraft on a separate aerial and compared with the frequency then being transmitted.   The difference in frequency between these two signals is proportional to the time taken by the wave to reach the ground and return and hence proportional to the distance of the aircraft from the ground.

In order to obtain a continuous indication of height the transmitter frequency is repeatedly varied over a certain frequency band as shown in Fig. 7a.   If the sweep time is correctly chosen the frequency difference between transmitted and received waves can be measured and used to indicate the height of the aircraft above the ground.
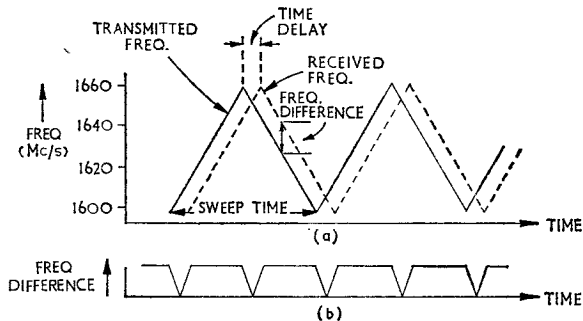

FIG. 7. PRINCIPLE OF THE FREQUENCY MODULATED RADIO ALTIMETER

Fig. 8 shows a simplified diagram of a typical f.m. altimeter which operates in the frequency band 4,200 Mc/s to 4,400 Mc/s. The transmitter consists of a resonant cavity excited by a coaxial line oscillator. This oscillator is a velocity-modulated device and its action is similar to that of a double cavity klystron but the frequency output can be swept over a wide range. The output is frequency modulated by mechanical rotation of a paddle wheel inside the cavity. The motor which rotates the paddle wheel is supplied from a motor control circuit. The transmitter output is fed to a horn aerial and a low-level output is applied to the receiver for use as a reference signal.
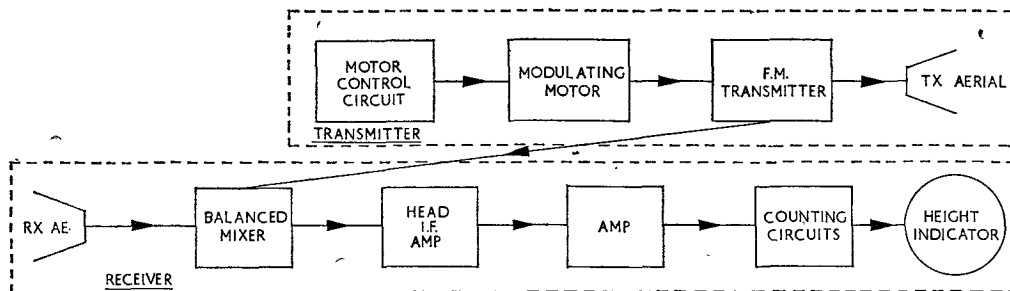

FIG. 8. SIMPLIFIED DIAGRAM OF A FREQUENCY MODULATED RADIO ALTIMETER

The received signal is fed to a balanced mixer, the local oscillator voltage being the reference signal injected from the transmitter. In order to reduce weight, the mixer is made of stripline instead of waveguide. The output beat note from the mixer is amplified in a low noise head amplifier before being passed to the main amplifier where it receives sufficient amplification to operate the counting circuits.

The beat note produced by mixing the transmitted and reflected waves will vary in frequency during the modulation period if the modulation is not linear, but the number of cycles in the beat note during this period depends only on the swept band. Thus by counting the number of cycles occurring in a fixed time an accurate indication of aircraft height is obtained. The counting circuits consist of an Eccles-Jordan bistable circuit which converts the beat note signal to square waves which are then differentiated and limited to give a d.c. output proportional to the height of the aircraft above the ground.

This radio altimeter has two ranges, 0–500 feet and 500 to 5,000 feet, the swept band being 100 Mc/s in the lower range and 10 Mc/s in the higher range. It has an accuracy at heights below 50 feet of 1 foot and can be used as the radio altimeter in automatic landing systems.

CHAPTER 3
GROUND-BASED NAVIGATIONAL AIDS

Introduction

This chapter deals with the outline of some typical ground-based radio navigational aids which are made use of by aircraft. The *airborne* equipment used to obtain navigational information from these aids consists of a normal communication v.h.f. or u.h.f. transmitter and receiver; no special d.f. circuits are employed. All the special d.f. circuits are contained in the ground equipment.

Cathode-ray Direction Finders (CRDF)

CRDF is a ground station direction-finding system using the v.h.f. band of 100 to 156 Mc/s and providing a bearing which is automatically displayed on a c.r.t. An Adcock aerial system is used, consisting of north-south and east-west pairs of aerials (Fig. 1). The voltages induced in each pair of aerials are fed to amplifiers the outputs of which are applied to the X and Y plates of a c.r.t.

With no voltage in the aerials the spot on the c.r.t. will be in the centre of the screen. When a voltage is present across the east-west aerials an amplified version of this voltage is applied to the X plates and the spot moves backwards and forwards across the screen as shown in Fig. 2.
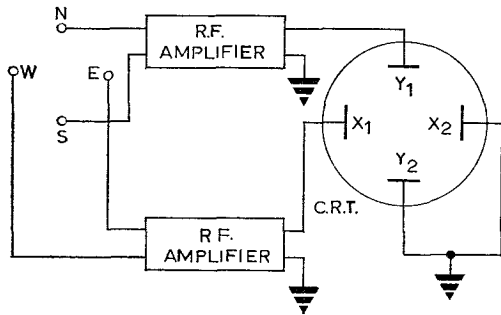

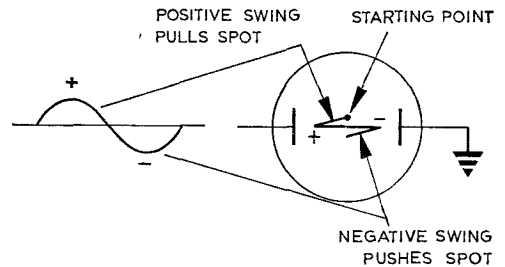FIG. I.   OUTLINE OF A SIMPLE CRDF

FIG. 2.   PRODUCTION OF CRT TRACE

When voltages appear in both aerials, due to an incoming signal, the spot will be deflected both horizontally and vertically. For example, if the signal comes from the north-west (point X in Fig. 3) equal voltages will appear across each pair of aerials and equal voltages are applied

AERIAL VOLTAGE 1 & 2
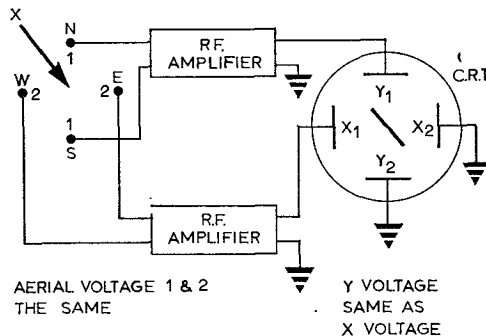THE SAME

Y VOLTAGE
SAME AS
X VOLTAGE

FIG. 3.   CRT DISPLAY

to the X and Y plates. Thus the c.r.t. spot is deflected mid-way between the X and Y plates and the trace lies diagonally across the screen as shown, i.e. the trace automatically shows the angle at which the signal strikes the aerial and therefore indicates the direction of the transmitter.

In this simple system the voltages from the north-south and east-west aerials are amplified in separate receivers and to avoid errors the receivers must have identical phase and amplification characteristics. This is very difficult to obtain.
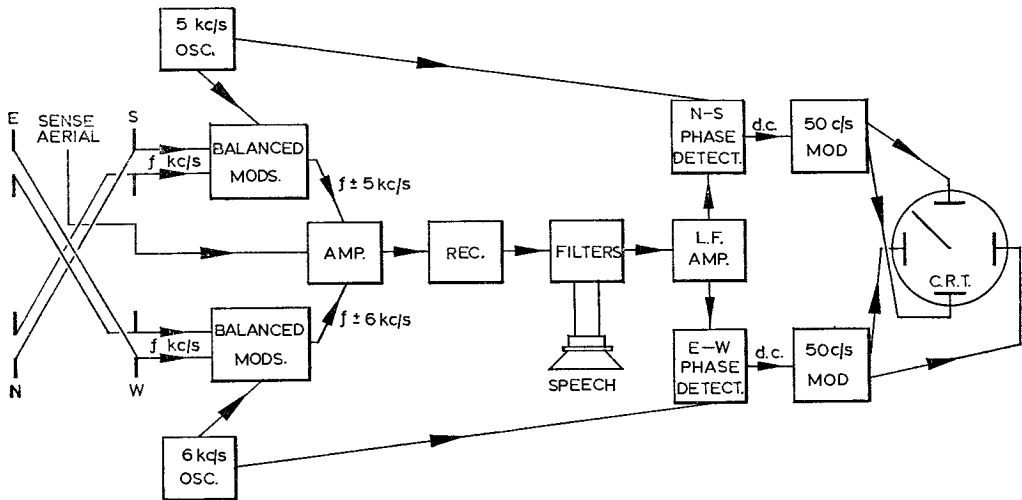

FIG. 4. SIMPLIFIED BLOCK DIAGRAM OF CRDF

In the c.r.d.f. equipment shown in block form in Fig. 4 this difficulty is overcome by modulating the N–S aerial voltage with a 5 kc/s signal and the E–W aerial voltage with a 6 kc/s signal. Thus both pairs of aerial voltages are identifiable and can be amplified in the same receiver. Two balanced modulators are used for this purpose and provide outputs of f $\pm$ 5 kc/s and f$\pm$ 6 kc/s as shown, the carrier component f being suppressed. The carrier transmitted by the aircraft may be speech-modulated to enable communication between the aircraft and the d.f. station. This voice modulated carrier signal is obtained from the central sense aerial and combines with the sidebands from the balanced modulators to give a composite signal which is amplitude modulated at 5 and 6 kc/s. The ratio of the depths of modulation at these two frequencies is proportional to the relative direction of the transmitter.

This composite signal is amplified in the receiver and the 5 kc/s, 6 kc/s and voice signals are separated by filters. The speech signals are then applied to a loudspeaker and the 5 and 6 kc/s signals are fed through an amplifier to phase-sensitive detectors where they are compared with reference signals from the modulation sources.

The outputs from the phase detectors consist of direct voltages, proportional to the amplitudes of the signals in the corresponding aerials, and of polarity depending upon the sense. In order to produce a trace on the c.r.t. these direct voltages are interrupted by a 50 c/s switching circuit.

Usually the aerial system is sited away from the airfield and clear of large obstructions which may affect the accuracy of bearings. Two separate receivers are provided so that simultaneous bearings from two aircraft can be obtained. Pre-set channels are provided in the receivers which are remotely controlled from the airfield control tower where a remote indicator unit is installed.

Commutated Aerial Direction Finder (CADF)

This system of direction finding is used in ground stations working in the frequency band 200 to 400 Mc/s. The principle employed is common to many other d.f. systems: the *phase* of an incoming signal is sampled and used to indicate the direction from which the signal arrives.

Fig. 5 shows an aerial (A) being moved at a constant speed round the circumference of a circle. A signal is shown arriving at an angle θ with respect to north. Because the aerial is moving in a circle, the phase of the induced signal voltage will be continuously changing in a cyclic manner, i.e. the aerial voltage will be *phase modulated* at the frequency of rotation of the aerial. The phase of the modulation will depend upon the direction from which the signals arrive and by comparison with a reference voltage obtained from the rotating mechanism an unambiguous indication of the direction of the transmitter is obtained.

Since it is not practicable to rotate an aerial at the required speed, a number of evenly-spaced fixed aerials are mounted in a circle and by means of a commutator each aerial is connected successively to the input of a receiver (Fig. 6). The commutation is done electronically using diode switches. In this way we get the same effect as a single rotating aerial. The phase
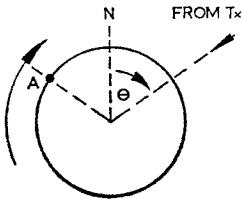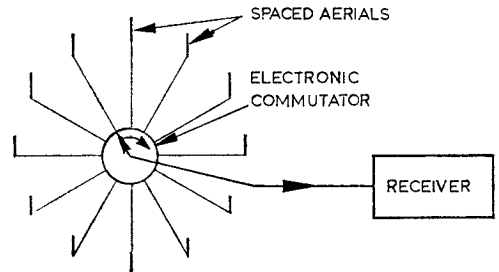

FIG. 5. PRINCIPLE OF CADF

FIG. 6. COMMUTATED AERIAL

modulated signal is applied to a phase-sensitive detector, the output from which is a steady voltage with a value proportional to the phase of the input. This voltage is used to give an indication of the transmitter direction on a c.r.t.

This basic system has certain practical disadvantages. If the transmitted signal is even slightly phase or frequency modulated the bearing obtained is inaccurate. Another disadvantage of the basic system is that a phase-sensitive detector cannot measure a phase difference of more than ± 180° without ambiguity. In a practical c.a.d.f. system this is overcome by spacing the aerials less than 180 electrical degrees apart and measuring the phase difference between two consecutive aerials.

A simplified block diagram of a practical c.a.d.f. equipment is shown in Fig. 7. Eighteen aerials are spaced round the circumference of a circle and are effectively rotated by electronic commutation, each aerial being connected to the receiver for 1 millisecond. The commutation is achieved by exciting cold cathode diode valves with voltage pulses obtained from a crystal controlled pulse generator.

The phase modulated output from the commutated aerial is applied to a receiver where its frequency is changed to an i.f. of 2 Mc/s. In order to cancel the effects of phase or frequency changes in the transmitted signal and to provide an audio output, an auxiliary aerial is connected to a second receiver. One output from this receiver operates the loudspeaker to provide communication with the aircraft and another output is applied to a mixer stage where it beats with a 130 kc/s signal obtained from a crystal oscillator. The 1.87 Mc/s component in the output of this mixer stage is selected and mixed with the phase modulated 2 Mc/s output from receiver 1 to
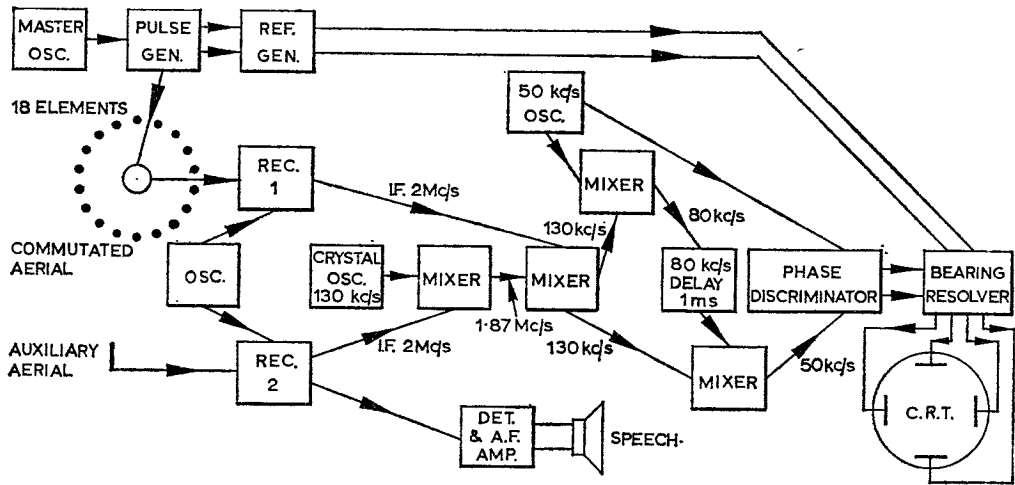
FIG. 7.   SIMPLIFIED DIAGRAM OF CADF

produce a signal at 130 kc/s.   This signal has the stability of the crystal oscillator and carries the phase modulation imposed by the commutated aerial but any phase modulation present in the original transmitted signal (and therefore common to both auxiliary and commutated aerials) has been removed.

At this point the signal can have a maximum imposed phase modulation of about $\pm$ 360°. In order to apply this to the phase discriminator which can only measure up to $\pm$ 180° without ambiguity, the phase deviation is reduced in the following manner.

The 130 kc/s signal is mixed with the output from a 50 kc/s oscillator and the 80 kc/s difference signal is selected and applied to a one millisecond delay network.   The delayed output is then mixed with the 130 kc/s undelayed signal and the signal at the difference frequency of 50 kc/s is selected.   This signal carries the original phase modulation but the deviation is the difference between the 80 kc/s delayed and the 130 kc/s undelayed signals, i.e. the difference in phase between two adjacent aerial voltages.   Thus the maximum deviation must be less than 180° (see Fig. 8).

This signal is applied to the phase discriminator along with a reference signal from the 50 kc/s oscillator.   The output from the discriminator is fed to the bearing resolver circuits along with voltages obtained from the aerial commutation reference generator.   The resultant balanced voltages are applied to the deflection system of a c.r.t. and a visual indication of the bearing is obtained.

The c.a.d.f. system gives very accurate bearings and a wide frequency range in the u.h.f. band can be covered by one aerial installation.   No additional airborne equipment is required other than the usual u.h.f. transmitter and receiver.

Automatic UHF Direction Finder

This ground d.f. system operates in the u.h.f. band of 225 to 400 Mc/s.   The signal received from an aircraft is phase modulated by mechanical rotation of a directional receiving aerial array. The phase modulation imposed on the signal can be used to indicate the relative bearing of the transmitter.   The bearing is displayed on a c.r.t. which can be remotely sited from the d.f. equipment.
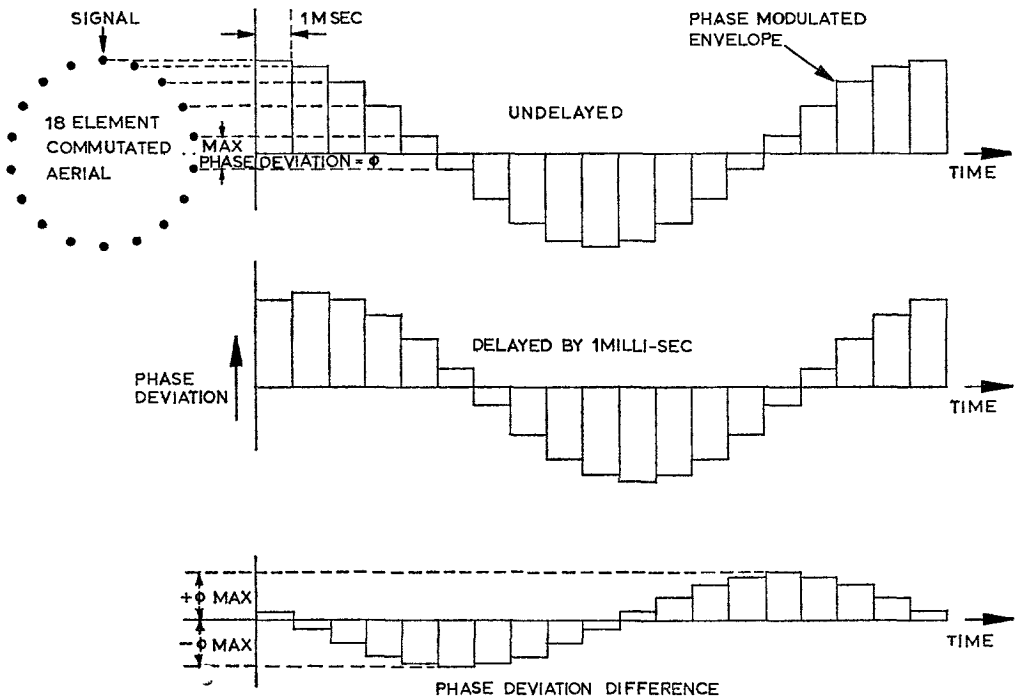
240

FIG. 8. ACTION OF DELAY NETWORK

The aerial consists of a wide-band vertical dipole mounted inside a fibre-glass cylinder on the surface of which are five vertical strip reflectors (Fig. 9). The horizontal radiation pattern has a cardioid shape. The cylinder is rotated round the dipole at a regulated speed of 2,520 rev/min and this imposes a modulation at 42 c/s on voltages induced in the dipole. The phase difference between the 42 c/s modulation envelope of the received signal and a 42 c/s voltage obtained from the aerial rotating mechanism gives an indication of the bearing of the transmitter.
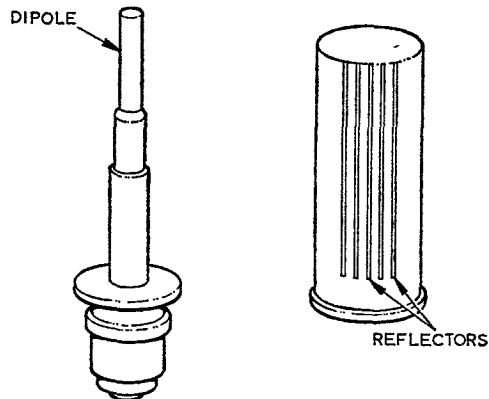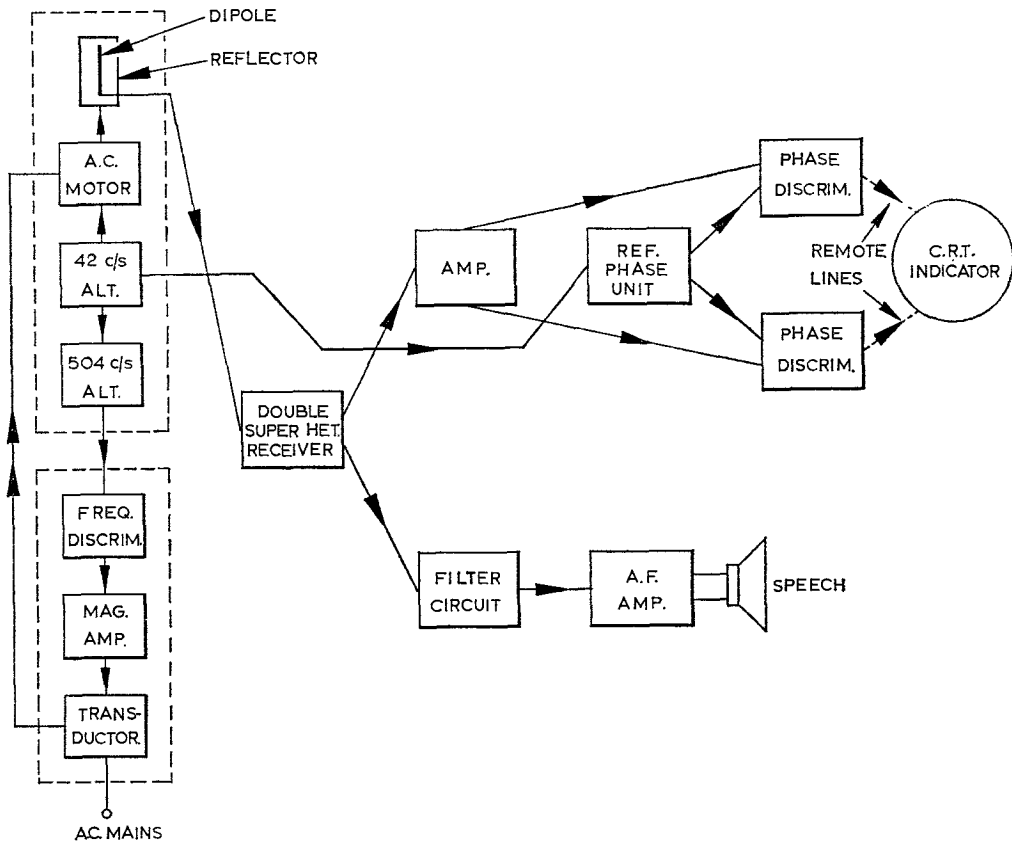

FIG. 9. UHF DF AERIAL

241

FIG. 10. SIMPLIFIED DIAGRAM OF UHF DF

A simplified block diagram of the aerial assembly and d.f. equipment is shown in Fig. 10. The aerial reflector is rotated by an induction motor. Two alternators are connected directly to the motor shaft, one providing a two phase output at the frequency of rotation of the reflector and the other gives an output 12 times the motor rotation frequency, i.e. at 504 c/s. The output from the first alternator provides a reference phase voltage and that from the second alternator is used to control the speed of the motor. This output at 504 c/s is fed to a frequency discriminator circuit which produces a positive or negative d.c. error signal if the motor is running too fast or too slow. The error voltage is amplified in a magnetic amplifier and fed via a transductor to the motor to maintain its speed constant at 42 c/s.

The receiver is a double superhet with two stages of r.f. amplification followed by a first i.f amplifier at 24 Mc/s and a second at 1,975 kc/s. After demodulation the speech modulated signal is fed through a circuit which removes the 42 c/s modulation imposed by the aerial to an a.f. amplifier which drives a loudspeaker.

A second output from the receiver carrying the 42 c/s modulation envelope is passed through an amplifier to two phase discriminators where they combine with a 42 c/s signal from the reference alternator. One discriminator provides an output proportional to sin θ and the other an output proportional to cos θ, where θ is the direction of arrival of the received signal with respect to north. These outputs are fed via remote lines to the deflection system of a c.r.t. housed in the Air Traffic Control tower.

242

AIR/GROUND NAVIGATIONAL SYSTEMS

Introduction

Some radio navigational aids require special installations both in the aircraft and on the ground. For efficient maintenance of either installation a knowledge of its counterpart is necessary. In this chapter we shall consider both aspects of some typical systems.

Instrument Landing System (ILS)

So far we have considered direction finding systems which enable the pilot of an aircraft to obtain his bearings from a known point, to be able to 'home' to a known point, and to 'fix' his position. Another important requirement in aviation is for a system providing information to enable a pilot to approach the runway in the correct manner to make a landing in conditions of poor visibility. One such system, operating on frequencies in the v.h.f. band, is the *instrument landing system* (ILS).

ILS provides information which enables a pilot to line his aircraft up accurately with the runway and to approach the touchdown point at the correct angle, i.e. on a correct *glide path*. The ground installation, which may be either mobile or fixed, consists of a *localiser* transmitter which indicates the projected centreline of the runway, a *glide path* transmitter to indicate the correct angle of approach and three marker beacons at fixed distances from the touchdown point (Fig. 1).
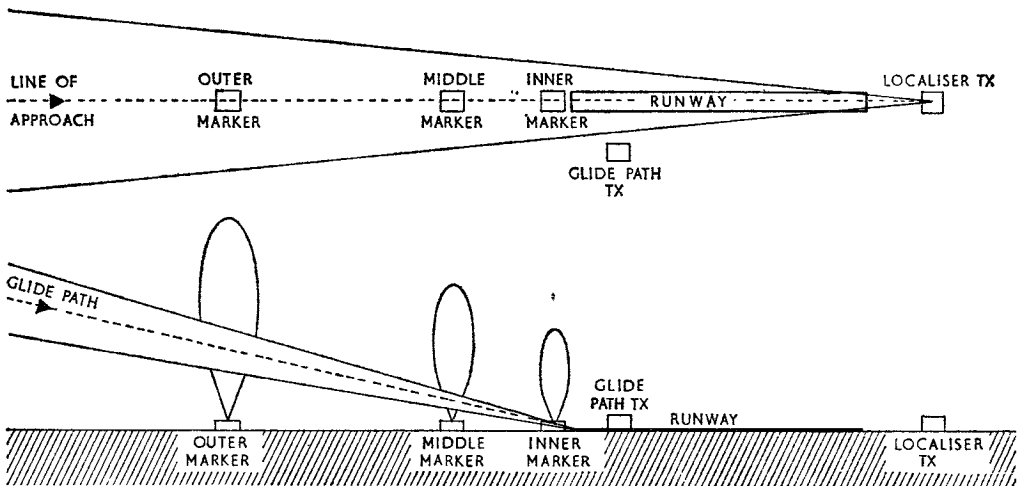

FIG. I. INSTRUMENT LANDING SYSTEM

The airborne equipment consists of a localiser receiver, a glide path receiver, a marker beacon receiver, a cross-pointer indicator and a marker lamp indicator. In addition, three separate aerials are required, one for each receiver.

The cross-pointer indicator is shown in Fig. 2. The output from the localiser receiver is fed to the vertical movement of the dual movement meter and the glide path receiver output is applied to the horizontal movement. When the two pointers cross within the small central circle

the aircraft is on the correct approach path. In Fig. 2a the indicator shows that the aircraft must be flown down and to the left to gain the correct approach path.

There are two additional movements in the indicator which operate vertical and horizontal alarm flags marked OFF when the localiser or glide path systems are not operating correctly (Fig. 2b). Under correct operating conditions the alarm flags are not visible.
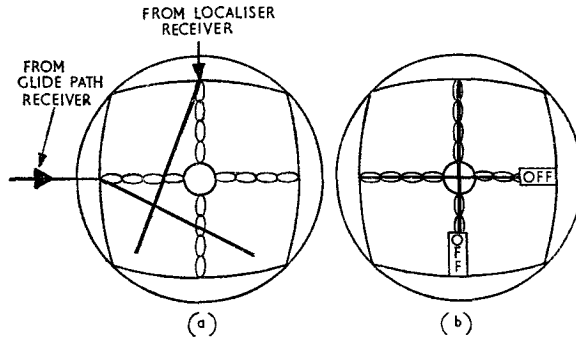

FIG. 2. ILS VISUAL INDICATOR

The accuracy of ILS is such that it can guide the pilot of an aircraft to within a few hundred feet of the touchdown point in conditions of poor visibility.

The ground *localiser* transmitter, which provides left-right information to the pilot, works on a frequency in the v.h.f. band between 108 and 112 Mc/s. An equisignal beam down the centreline of the runway is produced by overlapping the lobes of two directional aerials. The carrier frequency of each lobe is the same but one lobe is modulated at 150 c/s and the other at 90 c/s (Fig. 3). In the area of overlap along the projected centreline of the runway the 150 c/s and 90 c/s modulated signals are equal. If the aircraft flies to the left of the centreline the 90 c/s modulated signal increases and if it moves to the right the 150 c/s modulated signal predominates.
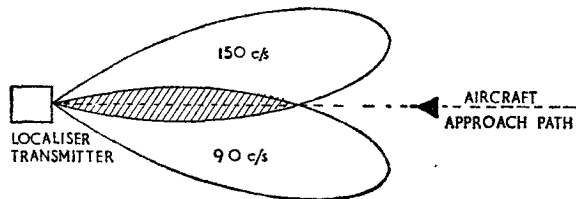

FIG. 3. LOCALISER RADIATION PATTERN

A simplified block diagram of the localiser transmitter is shown in Fig. 4. The 90 c/s and 150 c/s modulation frequencies are obtained by frequency multiplication from a common 30 c/s resistance-capacitance oscillator. The two outputs are passed to the 150 c/s and 90 c/s modulator units where they are amplified to the level required to modulate the output units. The station identification and speech signals are also fed to both modulator units.

The output from a temperature controlled crystal oscillator is frequency multiplied by a factor of 18 to provide the v.h.f. carrier. This output is fed to the 90 c/s and 150 c/s output units which include an anode modulated p.a. stage giving a power output of about 30 watts. Each output unit feeds its appropriate aerial via the phasing unit which ensures that the carriers are in the correct phase.
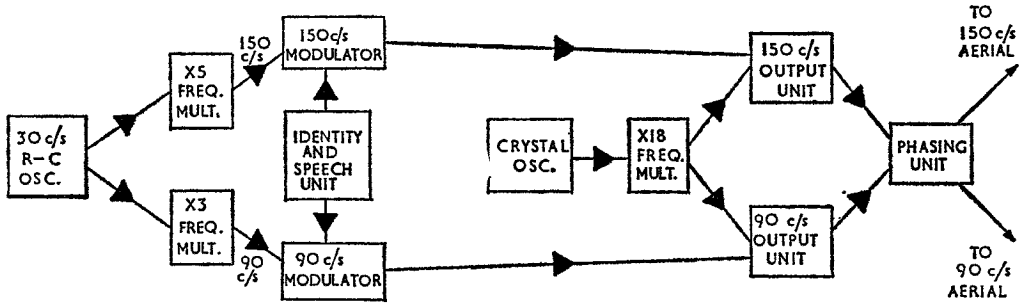
FIG. 4. SIMPLIFIED DIAGRAM OF A LOCALISER TRANSMITTER

A monitoring system is employed to ensure that the radiated beam is correct in all respects. If the beam is unsafe the monitor system sets off alarm signals and the beam is automatically switched off.

The general principles of the *glide path* transmitter are similar to those of the localiser transmitter. The glide path transmitter feeds two directional aerials mounted one above the other so that their radiation patterns produce an equisignal beam inclined to the horizontal (Fig. 5). The signals from these two aerials have a common carrier frequency in the band 328 to 336 Mc/s, but the lower lobe is modulated at 150 c/s and the upper lobe at 90 c/s. The equisignal overlap indicates the glide path.

Three marker beacons are installed along the centreline of the localiser beam and provide actual ranges from the touchdown point. They operate on a common frequency of 75 Mc/s and radiate a narrow vertical beam. Keying of different modulation frequencies identifies each marker beacon.

The interconnection of the main components of the airborne ILS installation is shown in Fig. 6. The channel selector control unit has a 12 channel selector switch which enables the pilot to select the localiser and glide path frequencies for different airfields by switching in appropriate crystals.

The localiser receiver amplifies the signals of the localiser aerial and supplies an input to the vertical pointer movement of the cross-pointer indicator. This input is proportional to the
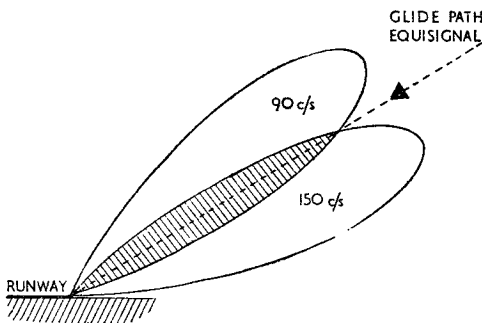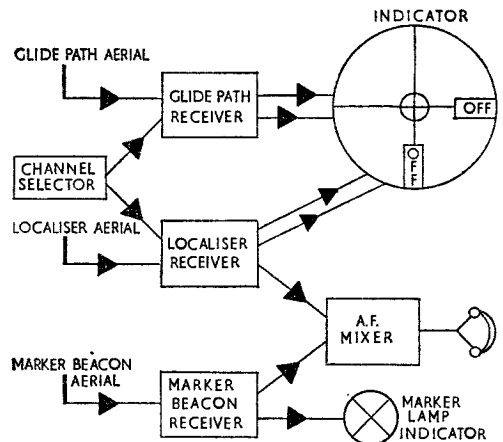

FIG. 5. GLIDE PATH RADIATION PATTERN

FIG. 6. AIRBORNE ILS EQUIPMENT

245

(A.L. 3, September, 1964)

difference in strength of the 90 c/s and 150 c/s modulation and the pointer indicates which way the aircraft must turn in order to gain the equisignal path. Another output is applied to operate the OFF flag positioned over the end of the vertical pointer. This output is proportional to the *sum* of the 90 c/s and 150 c/s modulation intensities and is fed to an additional meter movement which makes the OFF alarm flag visible if the localiser system is not operating correctly. A third output, fed to the audio mixer, gives a coded audio signal in the headphones which identifies the station. This signal can be interrupted in an emergency with a speech transmission from the air traffic control tower.

The glide-path receiver is similar to the localiser receiver but operates on a frequency in the 328 to 336 Mc/s band. The output feeds the horizontal pointer of the indicator to indicate whether the aircraft is above or below the glide path. There is also an output to operate the glide path alarm flag, but there is no audio output.

The marker receiver is tuned to a fixed frequency of 75 Mc/s and supplies an output to operate the marker lamp indicator. This lamp blinks when signals from a marker beacon are received. A coded audio output identifies which of the marker beacons is being received.

Automatic Landing System (ALS)

The accuracy of the localiser and glide-path beams of ILS is not sufficient to enable a pilot to make a landing in zero visibility. However, a system has been developed for use in military aircraft which can be used in conjunction with ILS to enable an aircraft to be *automatically* landed. The ILS localiser and glide-path information is fed to the automatic pilot to guide the aircraft until it is at a height of 300 feet when the localiser beam is switched out and azimuth guidance is thereafter achieved by information obtained from *magnetic leader cables*. A f.m. c.w. radio altimeter is used to feed height guidance information into the automatic pilot. The automatic pilot controls the positions of the rudder, ailerons and elevators to maintain the aircraft in the correct heading and attitude. Pitch and airspeed information is fed to an automatic throttle control so that the airspeed is held constant during approach and the throttles are automatically closed just before the aircraft touches down.
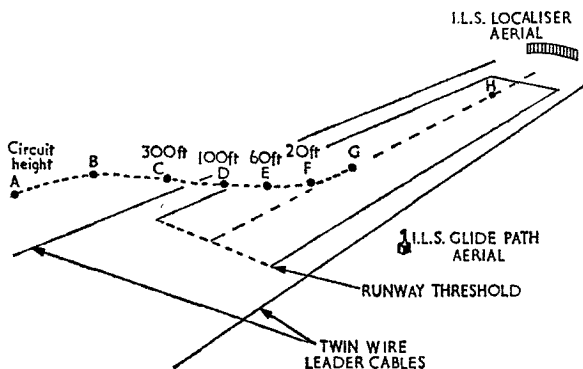

FIG. 7.   THE AUTOMATIC LANDING SYSTEM

The landing starts with an automatic approach in which signals from the ILS localiser and glide-path transmissions are used (points A, B and C in Fig. 7). At a height of about 300 feet (point C) the aircraft comes within range of magnetic fields set up by currents in the magnetic leader cables. The output from the leader cable receiver in the aircraft is substituted for the ILS localiser signal. At a height of 100 feet (point D) the ILS glide-path signal is removed from the autopilot and for a few seconds the aircraft attitude is held constant at a value averaged

during the whole of the approach to this stage. At a height of 60 feet (point E) the *flare* begins and the rate of descent is made proportional to height, i.e. the lower the aircraft the more slowly it sinks. This manoeuvre is accomplished with information obtained from the radio altimeter.

The final sequence begins at a height of about 20 feet (point F) when corrections are made to remove drift caused by crosswind. The wings are levelled and rudder is applied so that the aircraft touches down not only close to the centreline of the runway but also lined up with the runway (point G). After touchdown the pilot disconnects the autopilot and brings the aircraft to a stop (point H).

Although the system is entirely automatic the pilot can monitor each phase of the landing by watching monitoring signals displayed on a *sequence indicator* and can take over complete control of the aircraft at any instant if he considers it necessary.

Magnetic leader cable. Two cables are laid in the ground symmetrically about the centreline of the runway and supplied with constant currents of 4A, current in one cable being at 1,100 c/s and that in the other cable at 1,800 c/s. The cables extend along the length of runway and for 5,000 feet beyond the threshold. The currents produce magnetic fields as shown in Fig. 8a and the centre of the runway is defined when the 1,100 c/s and 1,800 c/s magnetic fields are of equal strength. The airborne receiver measures the difference between the field strengths and gives an output as shown in Fig. 8b.

The 1,100 c/s and 1,800 c/s signals are generated by an alternator consisting of a 3 phase motor driving two generators which produce currents of 4A at the two frequencies. A simplified block diagram of the ground equipment is shown in Fig. 9. The outputs from the alternator are fed to the two leader cables each of which is terminated in its correct terminal impedance. The equipment can be remotely controlled via landlines. A monitor unit ensures that the currents in the leader cables are constant and balanced within very close limits. If a fault occurs in the equipment or leader cables the monitor unit sets off an alarm and disconnects the mains supply from the alternator.
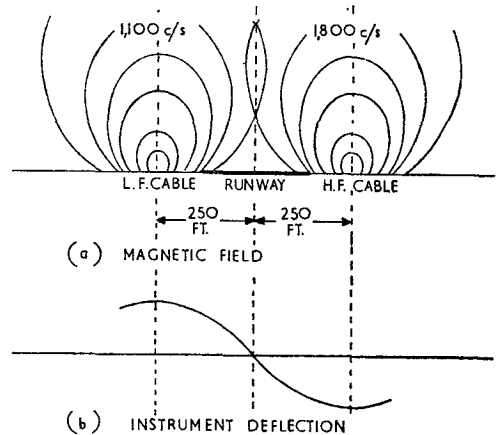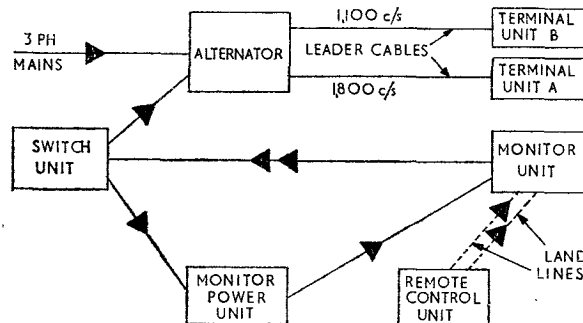

FIG. 8. LEADER CABLE FIELDS

FIG. 9. SIMPLIFIED DIAGRAM OF ALS GROUND INSTALLATION

247

Aircraft radio equipment. A simplified block diagram of the airborne automatic landing installation is shown in Fig. 10. The loop assembly consists of two loops mounted at right angles, each tuned to one of the leader cable frequencies, and a miniature 3-phase motor which rotates the loops at a constant speed. The signals induced in the loop are thus modulated (at 30 c/s) and by comparing the peak values of this modulation at each of the two frequencies a measurement of the two magnetic field strengths is obtained. By this means the effects of field direction and attitude of the aircraft are eliminated.
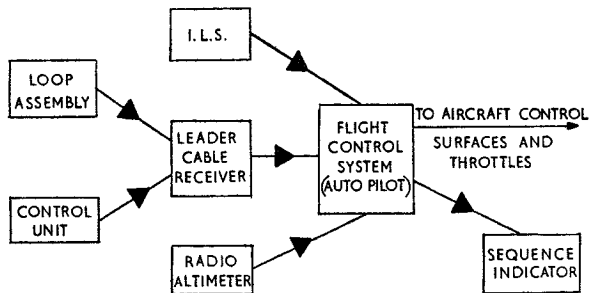

FIG. 10. SIMPLIFIED DIAGRAM OF AIRBORNE ALS INSTALLATION

The loop signals are then amplified in the receiver, separated by filters and applied to the autopilot and to the pilot's instrument display. Outputs from the ILS equipment and the radio altimeter are also fed to the autopilot and provide signals which control the aircraft at various stages in the landing.

A sequence indicator in the pilot's instrument display enables the pilot to assess the function of the equipment throughout each of the landing stages. The equipment is also provided with built-in test arrangements whereby the sensitivity and correct performance of the system can be checked before the aircraft is committed to an automatic landing.