THE FUNDAMENTALS OF TELEVISION

The moving finger writes, and having writ moves on
Nor all thy piety nor wit
Shall lure it back to cancel half a line
Nor all thy tears wash out a word of it.
Omar Khayyám.

General

Television is the art of seeing from a distance, almost instantaneously, a scene which is so situated, by virtue of its remoteness or position, that it cannot be viewed by optical means.

It is essential at the outset to determine what is implied by the term 'Scene' from a technical standpoint. Fundamentally, a scene may be described as an area from which light is travelling to the eye. It may be regarded as being composed of a very large number of infinitely small sub-areas known as elements. If each element is sending the same amount of light to the eye, the scene is said to be blank and has no pictorial interest, but if there is considerable variation between the amount of light proceeding from the various elements, then the scene possesses detail, and conveys intelligence to the eye of the observer. It is the function of a television system to transmit this intelligence to a distant point, and reconstitute there an image or picture which will send to the eye of an observer substantially the same intelligence as he would receive if he were standing in front of the original scene.

Some connecting link is required between the sending point and the receiving point, and two types of such connecting links are known, both electrical in their operation, these being radio transmission and cable transmission. In each case the intelligence is represented during the course of transmission by an electric current; we can now say, therefore, that we must provide at the sending end transmitting apparatus to transform the intelligence contained in the radiation of light by the numerous elements of the scene into electrical intelligence: the connecting radio or cable link will then reproduce that electrical intelligence at the receiving point, where we must provide further apparatus to transform it back into a luminous picture.

It is in this process of transformation at each end that a fundamental difficulty arises. The electric current in the connecting link can be completely specified by stating its amplitude at any instant of time. In other words, we only need to employ two dimensions, time and amplitude, in order to describe the current completely.

On the other hand, two dimensions are insufficient to describe a picture.
11. They are sufficient to describe a blank scene, because we can specify it completely by saying that it has a certain degree of brightness at a certain time. It follows, therefore, that a normal communication channel can be used with the simplest of transformation apparatus at each end to transmit the intelligence contained in a blank scene whose brightness is varying, because obviously, we can arrange that variations in brightness will be represented in the communication channel by variations in the amplitude of the current.

A picture, however, is more elaborate than a blank scene because it contains detail, and in order to describe a picture completely we need three dimensions: time, amplitude and position, the additional dimension of position being necessary in order that our description may cover the various items of detail located in various parts of the picture. The communication channel is unable to transmit such a picture directly, since owing to the absence of a third dimension or variable there is no property of the communication channel which we can utilise to convey information as regards the detail. The most the channel can do is to average out the brightness of all details in the picture, and reproduce it as a blank scene having a certain value of brightness. This difficulty may be made clear by employing an analogy.

Suppose that the scene to be transmitted is the page of a book, and that the communication channel is a telephone: imagine that a man A at the sending end wishes to represent the page to a man B at the receiving end, so that the latter can reproduce the page on a sheet of paper in exactly the same form as the original page. We know that the man A cannot send the whole of the intelligence contained in the page to B instantaneously, because the natural properties of the communication channel, consisting of the two men and their telephone system, are insufficient; but, if A reads out the words in order line by line, he can eventually transmit to B the whole of the intelligence in the page. In fundamental terms, what A is doing is to reduce the three dimensions which describe the intelligence of the page to two, in order that they may suit the communication channel.

The television system operates in a manner precisely similar to the above analogy. It examines the picture element by element, and in fact "reads it out" to the receiver, which then assembles the received intelligence in the correct order and reconstitutes the picture. The process by which the picture is examined is known as scanning.

Scanning

In practice the scanning of the original scene is very similar to the process of reading a book, since the scene is scanned in horizontal lines beginning at the top left-hand corner and ending with the bottom right-hand corner. The total assembly of lines employed to scan the full area of the scene is known as a field, and in order that transmission may be continuous, the scanning of the scene is regularly repeated, an agreed number of fields per second being successively transmitted. If the number of fields per second is more than a certain minimum figure, the eye of the observer at the receiving end, assisted by its natural property of persistence of vision, receives the impression of continuous transmission, and the reproduction of action and movement in the original scene becomes possible.

Synchronising Signals

Returning to our analogy of the two men on the telephone, if A reads out the page of the book to B, then unless B is told how to arrange the words, he will not necessarily reproduce the page in the same form as A sees it. He may arrange the words in too few lines, or too many lines, or he might even write them out in a single but very long line. To get over this difficulty, A must clearly insert additional information at intervals to enable B to arrange his words properly. At the end of each line he must say " Begin a new line," and if more than one page is to be read, at the end of each page he must say " Begin a new page." He must, in fact, synchronise the writing of B with his own reading. We may, in fact, describe the standard phrase " Begin a new line," which is spoken at the end of each line, as the line synchronising signal, and the phrase " Begin a new page " as the page synchronising signal.

An exactly similar state of affairs exists in a television system. At the end of the scanning of each line, the scanning process is momentarily interrupted and a line synchronising signal is sent to the receiver to tell it to begin a new line, and at the end of each field or picture which corresponds, of course, to the end of a page in our analogy, a picture synchronising signal, differing in character from the line synchronising signal, is sent to the receiver to tell it to begin a new picture.

We have now established that it is necessary in a television system to transmit two sets of signals from the sending to the receiving end : the vision signals, which will carry the intelligence corresponding to the brightness and detail of the picture, and the synchronising signals, which will tell the receiver how to reconstitute the picture so that all the various details take up their correct position. Now since these two sets of signals have separate functions to perform at the receiver, they must be effectively transmitted as separate entities, each signal retaining its own individuality. Fundamentally, this calls for two separate transmission channels, but this is an uneconomic solution to the problem, and it would be preferable to transmit both signals on one channel. Since the vision and synchronising signals have characteristics in common, they cannot be mixed in a straightforward way but must be combined in a manner which will allow their separation in the receiver. We shall see later how this can be achieved.

D.C. Waveform

If we consider for a moment the nature of the waveform of the electrical signals which in a sound-communication channel represent the sound to be transmitted, we find that it is a mixture of alternating currents or voltages having any frequency from about 40 to approximately 10,000 c/s. The important point is that all these signals consist of alternating current, (A.C.), i.e., a current which consists of alternate positive and negative half-cycles operating about a central position known as a datum line, and the area under the curves on each side of the datum line being equal.

In television, however, neither the vision signals nor the synchronising signals are represented naturally by an alternating current but by a direct or unidirectional current (D.C.). This does not here imply a steady supply such as that generated by a battery or a system of D.C. mains, but a current which operates on one side only of an initial datum line, and not on both sides as in the case of A.C. The difference between the two is illustrated in Figs. 1 and 2. Fig. 1 shows one cycle of an alternating current having a positive half-cycle *ABC* which is positive with respect to the datum line *OX*, and a negative half-cycle *CDE* which is negative with respect to the datum line, the latter, of course, representing the zero of voltage or current. Fig. 2 shows the same cycle of amplitude variation in the form of a unidirectional current. In this case, the whole of the waveform *ABCDE* is in the positive sense with respect to the datum line *OX*, which again represents a zero of current or voltage.

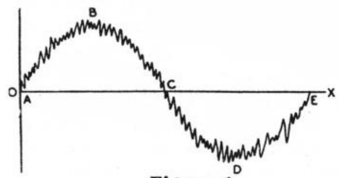

Figure 1

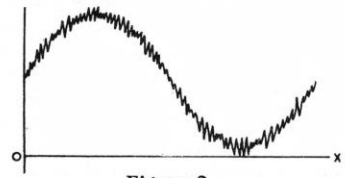

Figure 2

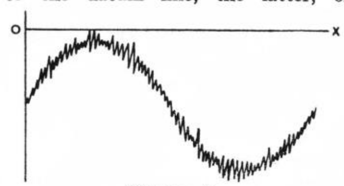

Figure 3

A third variation is possible, represented by Fig. 3, which again represents a unidirectional current, but having a negative sense with respect to the datum line *OX*.

It is evident that the currents which represent sound in the sound-broadcasting system have the form of Fig. 1, since the original sound wave is an oscillation about a central datum line, and the current which represents it must have the same form.

In a vision system, however, the vision signal represents the light and shade of the picture at any moment, and such light and shade is obviously positive at all times with respect to the datum line representing zero brightness or black, since we can conceive of no such thing as negative brightness. The vision signal, therefore, must have the form of Fig. 2. In other words, it is a unidirectional current. Its amplitude may and will vary from moment to moment, but it will always be positive with respect to zero.

The synchronising signals also are represented most naturally by a unidirectional current, and in this case it is immaterial whether the sense of the representative current is positive or negative, as all that is required is a burst of energy at the right moment which can be adapted to say to the receiver "Begin a new line," or "Begin a new picture," as the case may be. By arranging that the sense of the synchronising signals is opposite to that of the vision signals, i.e., negative, then a solution to the problem of maintaining the individuality of the two sets of signals in the transmission presents itself, for if we mix the signals in such a manner that the two datum lines coincide, we have formed one composite signal or waveform in which all excursions on the positive side of the datum line represent vision signals and convey information as to the brightness of picture detail, while those on the negative side of the datum line are concerned with synchronising alone. Thus the picture and synchronising signals may be separated from each other at the receiver by methods of discrimination based on "sense," which is for many reasons more satisfactory than other methods which have been tried.

A typical composite waveform of this sort is illustrated in Fig. 4. *OX* is the datum line, and the portion of the waveform *ABC* is a representation of a vision current during the time of one scanning line. It is confined entirely to the upper or positive side of the datum line *OX*. The other part of the waveform *CD* is the line synchronising signal, taking the form of a simple burst of energy in the negative sense, and this is similar in general characteristics to the simple negative unidirectional sine wave of Fig. 3, but differs in that it is a square pulse. The datum line *OX* thus forms an impenetrable barrier between the two types of signal. It will be evident that the complete waveform of Fig. 4 may be transmitted by one communication channel, but that the two components of it, viz., the vision and synchronising signals, though in effect mixed because thay are passing through the same channel together, are still retaining their individuality, since they differ in the important characteristic of sense, and may be separated at the receiver by some device which responds only to signals in one sense or the other.

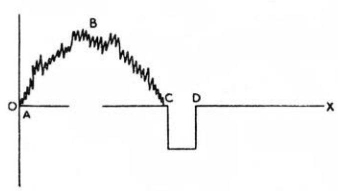

Figure 4

We have seen that the datum line *OX* represents the transmission of zero picture brightness, or black, and now that it has been made to coincide with the datum line from which the synchronising signals operate in the negative direction, it also represents the state when no synchronising signals are being transmitted. It therefore represents the transmission of no intelligence whatever, and is known as the **black level**. It is possibly the most important characteristic of a television system.

Picture/Sync Ratio

Referring to Fig. 5, which shows Fig. 4 redrawn with the addition of one or two explanatory details, if *BB* is the black level, then there will be another level on the positive side of the datum line, and shown as *WW*, which will represent the current corresponding to the whitest part of the line. This is termed the white level. There is yet another level on the negative side of the datum line, shown as *SS*, which will be the amplitude when the synchronising signals are being transmitted. This is known as the sync level. The ratio p/s of Fig. 5, i.e., the ratio of the amplitudes of the picture and synchronising signals, is an important value, and is known as the picture/sync ratio.

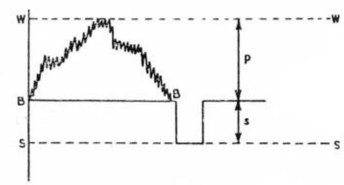

Figure 5

Two different values of picture/sync ratio are in use in the Marconi-EMI system: a ratio of 1 : 1 in most of the vision frequency circuits and a radiated ratio of 7 : 3, to which further reference is made in Item 1.1. The point of this difference is as follows:

It is unnecessary to radiate a synchronising amplitude as great as the picture amplitude corresponding to a picture/sync ratio of 1 : 1, as it has been found by experiment that with such a ratio and at such receiving situations where interference ruins the picture, the synchronising still remains perfectly steady, which is an obviously unnecessarily high standard of excellence of synchronising. It is far better to arrange that the synchronising of the receiver just fails when interference on the picture itself is of such a degree as to spoil its entertainment value. A radiated picture/sync ratio of 7 : 3 is therefore employed. To obtain such a radiated value the radio frequency circuits of the transmitter must be supplied with a picture/sync ratio of 1 : 1, i.e., more synchronising amplitude than is finally needed. This is because the synchronising signals fall upon the bottom bend of the modulation range of the transmitter and the process of modulation automatically reduces their amplitude relative to that of the vision signal. This matter is further developed in Item 1.1.

Vision Frequency Range

We have seen, then, that a television system operates by repeatedly scanning a scene and so generating an electric current, which at any moment

13

represents the brightness of elemental areas in the scene taken in a prescribed order, and, by means of radio transmission, reproducing this current, together with synchronising signals, at the receiver. We have now to determine, more intimately, the nature of this current and, in particular, the frequency band which it occupies. This frequency band will depend upon the number of lines "L" into which each picture is divided, and the number of pictures "P" which are sent per second, and also the shape of the picture. The product of the number of lines per picture and the number of pictures per second, viz. LP, is the number of lines transmitted per second and is known as the line frequency: P is known as the picture frequency.

If we analyse the vision frequency waveform we shall find that it is composed of a pure direct current, which we will designate f_o, and a large number of A.C. frequencies consisting of the picture frequency P and a great many of its harmonics, together with the line frequency L and a great many of its harmonics. In theory, for perfect reproduction we need an infinite band of frequencies, but in practice, an upper limit must be set, the lower limit being, of course, zero frequency, i.e., D.C. The highest frequency which it is necessary to preserve in the practical vision frequency waveform may be obtained from the following formula :—

$f = (L^2RP)/2$ where L and P have the meanings already specified and R is the ratio of width to height of the picture ; R is known as the aspect ratio. For example, for a system of 405 lines per picture, 25 pictures per second, and $R = 5:4$, the value of f, the highest frequency needed, is 2.56 Mc/s.

This formula is derived as follows :—

Imagining a square picture divided into L lines both horizontally and vertically, then we shall have produced a grille, containing L^2 small squares. If each alternate square is black and white respectively, corresponding to the presence of detail in the picture, then a rise in current will represent a white square and a fall in current the adjoining black square, and since one rise and one fall in current constitutes one cycle, whether in pure A.C. or variable D.C., then $L^2/2$ ripples of current will be created by scanning the grille. If the grille be scanned P times per second there will be $(L^2P)/2$ ripples per second, and if the picture is wider than its height in the ratio R there will be L^2R (instead of L^2) elemental areas, and the final number of ripples per second, corresponding to the finest detail, becomes $(L^2RP)/2$. Strictly speaking, the value of R to be used in the above formula should be somewhat greater than the actual aspect ratio of the picture because of a feature of the system known as suppressed scanning. This has the effect of increasing f by some 8 per cent., so that in the above example the value becomes 2.76 Mc/s.

It is an important feature of television that not only must all these frequencies be reproduced at their correct amplitude but they must all take the same time to get to the receiving screen. To achieve this there must be no phase distortion within the system and this usually means that the band width of all circuits must be considerably greater than the value of f given by the above equation.

Interlacing

We must now turn to one aspect in which television scanning differs from the action of a man in reading a book. The band width of all circuits is directly proportional to P, but the value of P also determines the amount of flicker in the system, the flicker decreasing with the increase of P. The flicker also depends upon the picture brightness, increasing with brightness. It follows that an absence of flicker can only be secured by the use of great band width in circuits, which brings in its train technical difficulty and expense.

Fundamentally, P, the number of pictures per second, is also equal to the number of vertical motions carried out by the scanning per second. This must be so if the lines are scanned sequentially, i.e., in their natural order, but we are not obliged to scan the picture in the natural order of the lines, provided the receiver faithfully follows the scanner. We can, for example, scan all the odd lines, missing the even lines, and then let the scanner return and scan the even lines. There will thus be two vertical motions of the scanning for every picture transmitted. The semi-picture containing either odd or even lines and resulting from such a process is called a frame, and the number of frames per second is known as the frame frequency, the process as a whole being termed interlaced scanning. The frame frequency, of course, is always greater than the picture frequency, and it is found that the eye responds to the frame frequency rather than to the picture frequency as regards flicker and is, as it were, deceived into thinking that $2P$ pictures per second are being transmitted. This affects the nature of the transmitted signals in a manner dealt with in Item 1.1.

Summary

We see, therefore, that the fundamentals of television involve the translation of the scene into an electric current by the process of scanning, which may be interlaced to reduce flicker in relation to a given band width, and that the process is constantly repetitive, to ensure the reproduction of a steady picture capable of displaying moving objects in the scene. (The repetitive nature of television signals allows of the extensive use of cathode ray tubes for monitoring the performance of circuits throughout the system.) In addition, synchronising signals linking the operation of transmitter and receiver must be transmitted in a manner ensuring their eventual separation. Since there are obviously a great many ways in which the picture can be scanned, it is necessary for an authoritative body to lay down, for the time being, standards which shall be observed in transmission and which are equally applicable to the transmitter and to the receiver. At present these standards are as follows :

No. of lines per picture .. 405
No. of pictures per second .. 25
Twin interlacing giving a frame frequency of 50

The remaining standards covering the synchronising signals and details of modulation are covered in the next section.

THE SIGNAL WAVEFORM

It is now necesssary to study in more detail the constitution of the signal waveform which has been briefly described in Item 1.0. This waveform, of course, contains all the information which the receiver requires to produce a picture. It will be seen that the waveform, as developed in practice, is somewhat more elaborate than that required for fundamental considerations only.

We have seen that the waveform really consists of two parts, one, that between black and white levels consisting of voltage excursions corresponding to the light and shade of the original scene ; and the other, that between black and sync levels, constituting control information which is sent to the receiver during intervals, allowed for that purpose, between the scanning of successive lines and also between the scanning of successive frames. Of these two parts, the first, the picture intelligence, may be left for the time being, but it is necessary to examine the other, the control information, in more detail.

Fig. 1 shows a diagram of the synchronising signals between lines and between frames. The number of lines per complete picture being fixed at 405, and the number of complete electrical frames, i.e., frames which contain all the scanning lines, being fixed at 25, it follows that there are 10,125 lines per second. To keep the receiver line oscillator in step, it is clear that we must transmit a synchronising signal at the end of each line, or 10,125 line synchronising signals per second, and since the timing apparatus in the receiver will be voltage operated it is necessary that the synchronising signal should develop to its maximum amplitude in the shortest possible time, otherwise the exact moment at which the spot in the receiver will receive its instructions from the synchronising signal will vary, and there will be positional distortion of detail in the picture. To put it in another way, the leading edge of the synchronising signal must have as steep a front as possible.

In the BBC standard waveform the time for the line synchronising signal to develop its maximum amplitude and cut the transmitter from black level to zero output is less than 1/5th of a micro-second, and the line synchronising signals consequently take the form of square-topped pulses operating downwards from black level, as shown in Fig. 1.*

The duration of a line synchronising signal is of importance. It must not be too long, or it would use up time which might be better employed in transmitting picture ; on the other hand, it must not be too short, or the pulse would be too fleeting to give reliable operation at the receiver. In the BBC standard waveform the duration is fixed at 10 per cent. of the time of a line, the time of a line here implying the time occupied by vision and synchronising signals. The time of a line is 98.7 micro-seconds, and for easy calculation this is approximated to 100 micro-seconds. The time of a line synchronising signal is, therefore, 10 micro-seconds and the width of these pulses is adjustable by means of a control on the sync-signal generator.

* *attached to page 3*

It might be thought that the synchronising signals are the only form of intelligence that it is necessary to transmit between lines, and from fundamental considerations this is true, but in practice an additional signal known as the line suppression signal must be inserted. This line suppression signal is in reality a period of black starting from slightly before the synchronising signal and continuing afterwards for a period ; in fact, it may be considered to be a form of level top from which the synchronising signal hangs.

The term **suppression** arose in the following way. In theory the period between the end of one line and the beginning of the next, and between the end of one frame and the beginning of the next should be devoid of any voltage excursions since during these periods no parts of the scene are being scanned. It was soon found, however, that in these intervals there were spurious signals either generated by the emitrons themselves, or due to induction from neighbouring circuits, and it became essential to suppress these unwanted voltages. This could be done most conveniently by mixing in square pulses of the correct duration, known respectively as line and frame suppression pulses. Although these pulses do not appear as such in the ultimate standard waveform, the intervals of black level which their introduction has created, are there (interrupted by the synchronising signals) and for convenience it is customary to show these intervals under the title of the pulses which cause them. Thus, in Fig. 1, there appears a line suppression period of 16.5 micro-seconds and a frame suppression period of approximately 1,400 micro-seconds.

There are, however, two further reasons why the suppression signals are required. The first is that when the line synchronising signal arrives at the receiver it starts the spot flying back to the beginning, and this fly-back takes a certain finite time, which, because of the limitations of practical apparatus, is longer than the duration of the synchronising signal. Clearly, no picture impulses must be allowed to be transmitted before the spot has reached its starting point on the left of the picture, or there will be a shadow on the left of the picture due to the spot being modulated by picture signals while still proceeding at high velocity to the left. This would be a visual defect, and can be eliminated by maintaining the modulation of the spot at black level until such time as the receiver spot has completely returned.

Similar arguments apply in connection with the frame scanning. Generally speaking, commercial-design receivers require that the duration of the line-suppression period should not be less than about 15 per cent. of the time of the line, and the duration of the frame suppression period should not be less than about 10 lines. It will be seen that the practical values shown comfortably cover these minimum figures.

There is a further reason, however, for the provision of the line suppression pulses (but inapplicable to the frame suppression pulses). It will be found later that at various places in the video chain it is necessary to carry out a process known as the restoration of D.C. with respect to black level. This process requires that there should be a period of black regularly occurring at a fixed position during the intervals between lines. Referring to Fig. 1 this period is termed the post line-sync suppression period and is 6 micro-seconds.

A further refinement is necessary. It might be thought that when a clear interval of black level had been produced by the insertion of the line suppression pulses, there would be no objection to the line synchronising signal being inserted straight away at the beginning of this interval. This is, however, inadvisable in practice for the following reason. Let us assume that the beginning of the line sync signal corresponds with the beginning of the line suppression period. Now the voltage at any time during the scanning of a line is, of course, quite arbitrary and at the end of the line just prior to the sync pulses it may be either high, corresponding to white, or low, corresponding to black. In order to execute the sync pulse, the voltage must return to black and beyond, and since it is quite impossible to design any circuit in which a voltage excursion can take place in zero time, should the end of the scanning voltage be high then it will take an appreciable time to get down to execute the leading edge of the line synchronising pulse. On the other hand, should the scanning voltage be low, corresponding to black, then this time will not be required and the voltage can move off at once to execute the leading edge of the synchronising signal, which by comparison with the former circumstance, will be early.

The practical effect of this would mean that assuming we are scanning the picture from left to right, then whenever the right-hand edge of the picture was white, the synchronising signals immediately following would be late in comparison with those immediately following lines whose right-hand edges were black. This would result in positional distortion in the picture.

This is obviated by having a permanent period of black prior to the line synchronising signal, this period being called the pre-line sync suppression period, which normally lasts $\frac{1}{2}$ micro-second. The matter may be put in another way by saying that the line synchronising signal has $\frac{1}{2}$ micro-second of delay with respect to the line suppression period. Referring to Fig. 1, this period is termed the pre-sync line suppression period.

It will be seen that the sum of the pre-sync line suppression period, the line sync suppression period, and the post sync line suppression period is equal to the line suppression period.

We have so far examined in detail the control information present between the periods of active line scanning, or more briefly, the signals between lines.

The period when the scanning beam is traversing a row of elements is often termed the forward stroke or trace and is equal to the whole of the periodic time of one line minus the line suppression period. The action of the beam in returning to the beginning of a line is called the return stroke or trace, or the line fly-back and occupies a period somewhat less than that of the line suppression period.

It is now necessary to examine the nature of the signals between active frame scanning periods, or, more briefly, the signals between frames. They are of a rather more complicated nature than those between lines and are drawn out fully in Fig. 1, where the upper section shows the signals at the end of the even frames and the lower section shows the signals at the end of the odd frames. There is no difference whatever between these two sets of signals, although a study of Fig. 1 may suggest an apparent difference in that the signals at the end of even frames start at the end of a line scanning period, whereas the signals at the end of odd frames start half way along the scanning of a line. This is, of course, absolutely necessary because an odd number of lines, viz. 405, is scanned in two exactly similar periods called frames. The duration of each frame must, therefore, be $202\frac{1}{2}$ lines. It follows that the commencement of one set of frames, the odd frames, must occur at half way along a line. Since the signals at the end of even frames are identical with those at the end of the odd frames it will suffice to examine one set only, those at the end of the even frames. The first feature is that there is a frame suppression period corresponding in function to the line suppression period but having an appropriately larger duration, chosen to be that of 14 lines, or approximately 1,400 micro-seconds. This allows sufficient time for the frame fly-back to occur in receivers of average specification.

The first 400 micro-seconds of the frame suppression period also contain the frame synchronising signal, the form of which is rather more involved than its counterpart, the line synchronising signal. In theory, all that is necessary, is a long period of zero radiation from the transmitter extending over a number of lines, possibly 5 per cent. of the total. If, however, the signal takes this form the continuous frame synchronising signal will entirely submerge the line synchronising signals, which consist of intermittent periods of zero radiation occurring at the end of each line, and the line synchronising signals will, therefore, cease to occur while the frame synchronising signals are being sent. Though in theory this does not matter, in practice the line oscillator at the receiver will in this period deviate from its proper frequency, and will subsequently require some time at the beginning of the frame to be pulled back into step, with the result that there may be distortion of the first few lines. To avoid this it becomes necessary to ensure continuity of the line synchronising signals and the frame sync signal must be introduced in such a way that it does not interrupt this continuity. This may be achieved by arranging that whatever form the frame sync signal takes it still contains a component at line frequency, i.e., 10,125 c/s.

The next point is that the frame sync signal must start accurately either half way along line 203, when the scanning is at the end of the odd frames, or at the end of line 405 as when the scanning is at the end of the even frames. (In Fig. 1 the lines are numbered successively as they are scanned, that is to say the odd frames contain lines numbered 1 to $202\frac{1}{2}$ and the even frames contain the latter half of line 203 as far as line 405, thus line 1 is the first line after the end of the even frames.) It is shown in Fig. 1 that the line sync signal at the end of line 202 is followed half a line later (or $\frac{1}{20250}$ second later) by the beginning of the frame sync signal. General conditions of design make it clear that such a signal can best be introduced by having square pulses available at twice line frequency, 20,250 c/s, so that these can be injected at the proper moment. If we arrange to send a train of such pulses, the first one commencing exactly where the frame signal is desired to commence, then such a train constitutes a group of pulses repeated every fiftieth of a second.

Such a group pulse possesses properties sufficient to meet all the requirements of the system. Firstly, although composed of a number of component pulses, it is over all, a pulse regularly repeated at frame frequency and is, therefore, a frame synchronising signal. Secondly, as it consists of a number of pulses at a frequency of 20,250, the accuracy of its starting point can be determined, because of circuit considerations, with much greater accuracy than would be the case if we simply generated a signal pulse having a frequency of 50. Thirdly, since its frequency is twice the line frequency, alternate pulses of the group will occur exactly where a line synchronising signal would have occurred and, therefore, continuity of pulses at line frequency is maintained and the line synchronising circuits of the receiver receive continuous tripping and do not get out of step.

It has been found from experience that the group need contain only a few pulses, and eight are sent, having, of course, an overall duration of four lines. The duration of each component pulse, however, is made to be a good deal longer than that of a line synchronising signal, a duration of 40 micro-seconds having been chosen. By this means the frame sync signal as a whole has more energy in it than if the component pulses were 10 micro-seconds wide, as in the case of the line synchronising signal, and a stronger trip is applied to the frame scanning circuits of the receiver. Because the width of these component pulses is so much greater, they are in practice termed the broad pulses. They are shown in Fig. 1 as occupying the first 400 micro-seconds of the frame suppression period and it is important not to confuse the broad pulses of 40 micro-seconds with the intervals between them of 10 micro-seconds which, of course, are not the active pulses.

In view of the fact that each frame bisects the total time of one picture accurately to $\frac{1}{50}$th second, it follows that interlacing must be automatic if an odd number of lines is chosen. The forward trace of line scanning is, of course, not horizontal, but it is slightly sloping as the scanning system is being operated by the line and frame deflection system simultaneously. By the time a spot has moved all the way along a line from left to right, it has moved downwards the space of two lines so that by the time it is half way across it has moved down the space of one line. Referring to the lower diagram in Fig. 1, consider what happens when the scanning is half way along line 203. The frame fly-back then starts. If we may assume for the moment that this is instantaneous and if we may also neglect as irrelevant the operation of the line scanning during the period of frame fly-back, then the latter will cause the spot to move vertically upwards to a point one line width above the first line previously scanned so that the ensuing lines will, if the interlacing is perfect, automatically place themselves symmetrically between the already scanned lines (the lines of the odd frame). Perfection of interlacing is, in practice, somewhat hard to obtain for a number of reasons which cannot be discussed here as they properly belong to the sphere of receiver design.

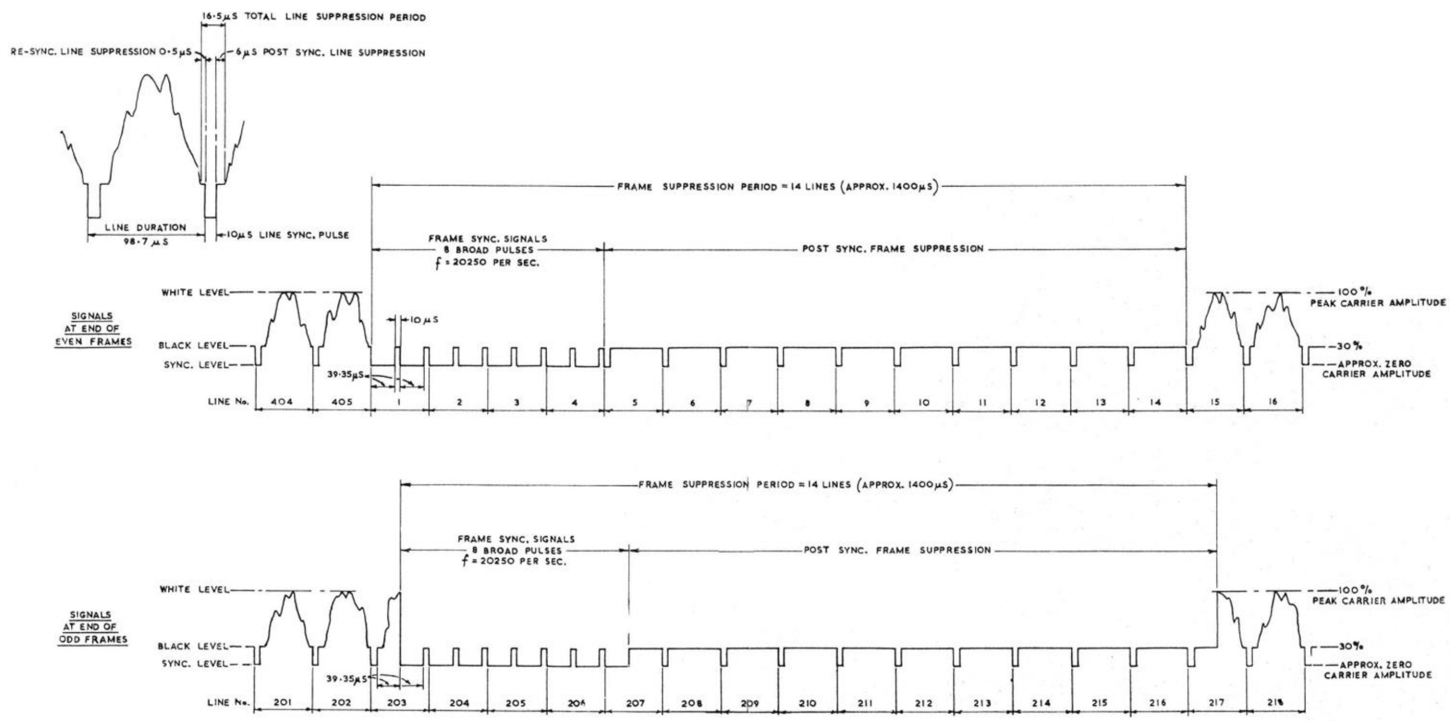

Fig. 1. The Signal Waveform

THE RESTORATION OF D.C.

In a study of modern television circuits, it will be found that a principle known as the restoration of D.C. is here and there applied. It is a process akin to rectification which enables us to achieve more than one important result. In the M.-E.M.I. System in particular, D.C. restoration is effected at various places with various objects in view, and in this note the various applications of this principle are discussed.

D.C. restoration may be said to have originated in the following way. In sound apparatus, the current to be dealt with is alternating, that is to say it consists of an oscillation equally on either side of a central datum line. This is because the sound itself which is dealt with by a broadcasting system is an alternating oscillation, and the current passing through the chain which properly represents it is an alternating current. In order to effect the transmission of television, it is necessary to derive from the scene a current proportional to the light variations produced by the operation of scanning and to transmit this current to a distance. It is at once apparent that the variations in light are all positive, there being no such thing as negative light, therefore the current which truly represents such light is a direct current, and not as in the case of a sound system an alternating one. Consequently, the true television transmitter must be a transmitter of D.C., and not of A.C.

From any scanning apparatus therefore, we primarily derive a direct current whose amplitude is continuously varying. The natural amplification chain, if such a current is to be faithfully reproduced, would consist entirely of D.C. amplifiers. There are, however, grave practical objections to the use of such amplifiers, and it is desirable to use A.C. amplifiers. If the vision current is passed through such amplifiers, however, it must lose its original character, as originally it is a D.C. and after passing through such an amplifier it must emerge as an A.C. The process known as the restoration of D.C. however, enables the alternating current emerging from such a system of A.C. amplifiers to be so treated that it is restored once more to its original form, and appears at the end of the amplification chain as though that chain had consisted of D.C. amplifiers.

The simple principle of D.C. restoration may be made clearer by reference to the accompanying diagrams. In Fig. 1 a simple square topped waveform is shown. The datum line, i.e. the line about which this curve is described, is the line AB, and it will be noticed that the waveform is entirely situated on the upper side of the datum line. When this is the case, the waveform may be correctly described as a D.C. In Fig. 2, the same waveform is shown but differently situated, so that the datum line CD now passes exactly through the centre of the waveform in such a way that the areas above the datum line are exactly equal to those below it. Such a waveform is described as an A.C. It is at once obvious that the waveform of Fig. 2 is simply that of Fig. 1 minus that of Fig. 3, which shows a steady D.C. component represented by the line GH standing above the datum line EF. Thus, Fig. 2, plus Fig. 3 gives Fig. 1.

A.C. amplifiers, into which category come, of course, resistance and other amplifiers used for sound amplification chains, can naturally pass only

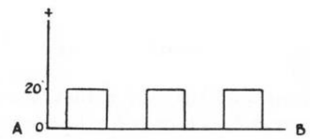

Figure 1

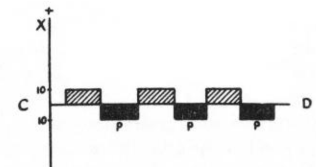

Figure 2

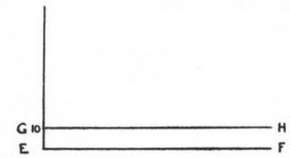

Figure 3

alternating waveforms, such as are exemplified in Fig. 2, and therefore if an attempt is made to pass through them a D.C. waveform such as in Fig. 1, it will emerge as in Fig. 2, having lost the D.C. component of Fig. 3. It is the object of D.C. restoration to restore to the emergent waveform the missing D.C. component (Fig. 3) so that the final result will appear as in Fig. 1. This is obviously important, as if Fig. 1 truly represents a succession of impulses derived from a scanning system, then they cannot be truly represented by Fig. 2, and the immediate result is, of course, that there is some lack of faithfulness in the tonal reproduction of the picture.

In order to see how we may restore the missing D.C. component, let us examine Fig. 2, and see in what way it differs from Fig. 1. It is at once obvious that the lower peaks P of the waveform have slipped below the datum line CD instead of being coincident with it as in Fig. 1. If therefore, we can arrange some circuit which makes it impossible for any potential applied to it to extend beyond a certain fixed value, we should achieve our result. With reference to Fig. 2, for instance, if we could design a circuit which freely allowed variations of potential to take place in the direction of CX but refused to allow variations in the direction CY, then clearly the waveform of Fig. 2 would be stood upon a basis of the datum line CD and would be compelled to execute all its excursions in the direction CX.

19

We can design such a circuit, which is known as a D.C. restoration circuit. In doing so we make use of the special properties of the grid of a valve. Such a circuit is shown in Fig. 4 and it will be seen that it consists merely of a valve in whose grid circuit is a condenser and grid resistance and no grid bias. It will be at once noticed that the circuit is exactly the same as that of a grid rectifier, but for the moment this point should be ignored, and later the relation between D.C. restoration and rectification will be explained.

Let us apply to the circuit of Fig. 4 the waveform of Fig. 2 in such a direction that the peaks P are attempting to make the grid positive, as shown in Fig. 5. It is a property of most valves that for grid voltages on the negative side of zero no grid current is drawn, but when a positive voltage is applied to a grid, electrons are attracted from the filament so that they proceed to the grid, and we say that a grid current flows from grid to filament. Moreover, it is a feature that with increasing positive voltage the increase of grid current is very rapid, that is to say, the impedance of the grid circuit is not constant, but decreases as the grid voltage is increased.

In Fig. 5, it will be seen that we are attempting to apply to the grid a positive excursion *ab* of 10 volts. In practice we do not succeed in making the grid 10 volts positive. Immediately the grid becomes slightly positive, say 1 volt, it attracts electrons from the filament, which accumulate on the grid, thus making it negative until the deliberately applied positive charge is cancelled. When our input pulse rises to 2 volts positive, further grid current flows and the accumulation of electrons is now such as to produce 2 volts negative. Thus, our full pulse *ab* of 10 volts does not really make the grid 10 volts positive but the grid is finally still at zero potential, or possibly very slightly positive, due to the simultaneous presence on the grid of the deliberately applied positive pulse together with a negative charge of almost equal value which is built up simultaneously with the application of the pulse *ab*.

In order to understand the working of this circuit, and incidentally of rectification, it is most important to get a clear idea of this simultaneous existence of positive and negative pulses on the grid. The grid is at approximately zero potential. We have endeavoured to make it positive by 10 volts, but it has defeated us by securing from the filament an equal and opposite negative charge, so that its potential remains at zero. We may say in fact that it is impossible in this circuit to make the grid positive, for any attempt to do so will be met by the heavy deposition on the grid of an equal and opposite negative charge. The insertion of the condenser C_1 ensures that this charge remains there and is not lost to earth. This state of affairs continues during the top part of the pulse *bc*. The pulse now recedes from *c*, via *d*, to *e*. When the pulse reaches *d*, the whole of the original 10 volts has been withdrawn and we are left with the negative charge of 10 volts which the grid attracted to itself, and which is retained by the condenser C_1. The pulse now executes the part of its excursion from *d* to *e*, thus applying a further 10 volts negative to the grid, which is now 20 volts negative. From *e* to *f*, the grid is 20 volts negative, and at *g* the grid will be zero again.

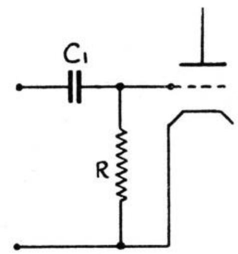

Figure 4

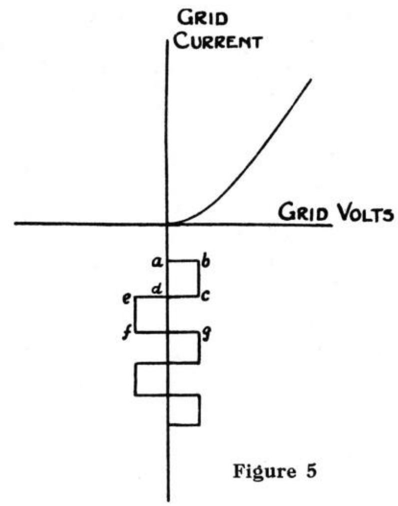

Figure 5

The presence of the grid resistance R is not really essential theoretically in the restoration of D.C. to the particular waveform under discussion, but its performance will be more obvious when we come to consider more complicated waveforms. It will be seen, however, that it provides a continuous path of leakage for the negative charges on the grid and it is probable that during the time *ef* a certain leakage will take place which might tend to reduce the negative grid potential to, say, 19 volts at *f* instead of 20. Accordingly, the grid would try to become 1 volt positive at *g*, but here again this positive voltage of 1, would cause a little grid current to be drawn, and as before the point *g* would take up the same potential as the points *b* and *c*, i.e. zero.

Clearly this arrangement has provided just what we wanted. The grid sets up an impenetrable wall situated at approximately zero potential, on the positive side of which no excursions can be made, and thus the waveform operating on the grid is entirely on the negative side of this wall. This circuit in fact establishes a datum line situated at zero potential. This datum line coincides with the points P in Fig. 2 and the waveform of Fig. 2 thus exists entirely on one side of this datum line, and therefore becomes the waveform of Fig. 1. Thus, we are entitled to say that to Fig. 2 has been added Fig. 3, and D.C. has been restored. Since Fig. 1, if amplified in A.C. amplifiers, would emerge as Fig. 2, the addition of the circuit of Fig. 4 to the end of such a chain of amplifiers will result in the output appearing as in Fig. 1, and thus the whole system of A.C. amplifiers is in fact converted into one of D.C. amplifiers.

We must now consider what is required when the waveform is not quite so simple as that we have been considering and is undergoing changes of amplitude. It is in such cases that we require more particularly the services

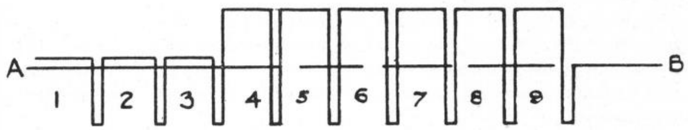

Figure 6

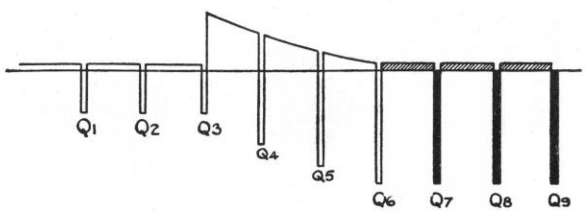

Figure 7

of the grid resistance R. We can, in fact, at once proceed to the consideration of a typical television waveform as illustrated in Fig. 6, which shows the picture output together with synchronising signals formed by the scanning of nine lines.

If AB is the datum line corresponding to black, then the first three lines are black and the last six are white. The waveform of Fig. 6, must be faithfully reproduced at the receiver, and requires throughout a D.C. amplification chain. If we apply it to the A.C. amplification chain it will be executed in the manner shown in Fig. 7. The A.C. amplifiers will endeavour to alter the waveform so that it is an A.C. i.e. the area on one side is equal to that on the other. Lines 1—3 will probably be untouched, but lines 7—9, which exhibit strong D.C. characteristics, will be transformed into the waveform shown in lines 7, 8, and 9 in Fig. 7, in which again the areas above the datum line are equal to the areas below, and there will be an intermediate transitional period, as shown in Fig. 7, lines 4, 5, and 6. Application to a D.C. restoring circuit, however, will restore the waveform of Fig. 7, to that of Fig. 6.
21 Remembering that our restoring circuit sets up an impenetrable wall of potential which cannot effectively be pierced, if we arrange to apply the waveform of Fig. 7 in such a way that the points Q_1 to Q_9 shall attempt to pierce the wall, then points Q_4 to Q_9 will fail to do so and will be unable to descend below the level set by the points Q_1 to Q_3, thus finally they will all

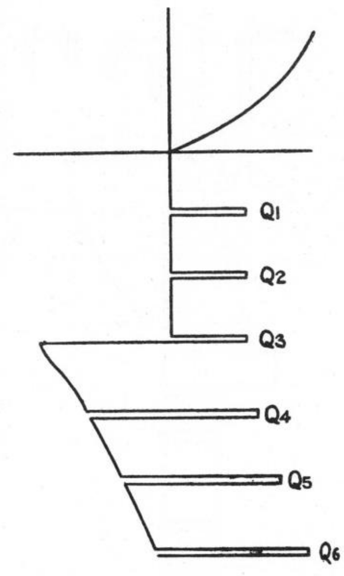

Figure 8

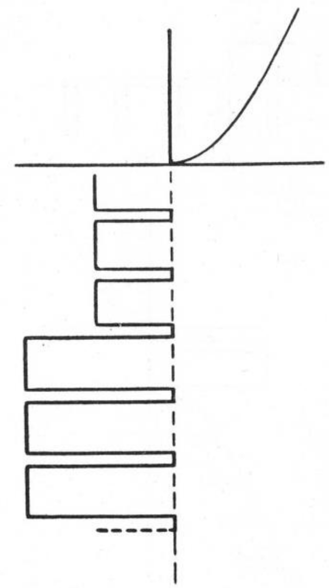

Figure 9

be level, as in Fig. 6. The same circuit as Fig. 4 will, of course, be used, and the waveform will be applied as in Fig. 8. Actually, however, this figure only shows how we shall attempt to apply the waveform, for the points Q will never in fact reach the values of potential at which they are drawn in the diagram. They will coincide on the grid with an equivalent negative potential which will be exactly sufficient to see that they are all at zero potential. That is to say, when we try to apply the initial positive pulse Q_1, a negative pulse equal in value will be applied, driving it back to zero. This state of affairs will remain roughly set until we arrive at the point Q_4, which makes a more determined effort to drive the grid positive, but, like the others, fails, and in turn Q_4, Q_5 and Q_6 are driven back to zero, and the effective excursions of the grid potentials are as in Fig. 9 which as will be seen coincides with Fig. 6.

Let us now consider what happens when the amplitude of the pulses drops, as shown in Fig. 10. After passing through an A.C. amplifier, the waveform of Fig. 10 will appear as in Fig. 11, and it will be seen that the datum line will now tend to reset itself so that the waveform is maintained

Figure 10

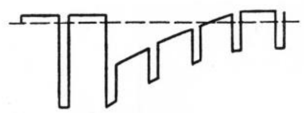

Figure 11

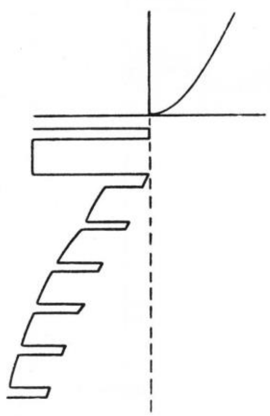

Figure 12

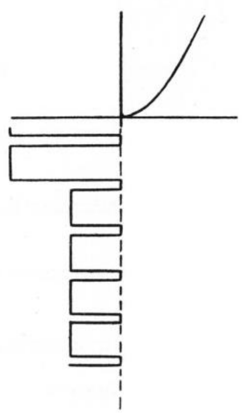

Figure 13

as an A.C. When we apply this waveform to the circuit of Fig. 4, the points Q attempt to recede along the grid characteristic in the negative direction, so that they would no longer coincide with zero grid potential, as the waveform is trying to execute an alternating oscillation about a point on the operating characteristic fixed by the degree of negative grid bias produced by the grid current drawn by the previous pulses of large amplitude. This state of affairs is shown in Fig. 12.

It is here that the grid resistance R plays its part, for it allows the negative charge to leak away. This, of course, pushes the whole waveform to the right, i.e. in the positive direction, a movement which is only arrested when the points Q once more touch the line of zero grid potential, and commence to re-establish a negative grid bias by drawing grid current as before. This state of affairs is shown in Fig. 13. The waveform of Fig. 13 is precisely that of Fig. 10, so that once more the modification of the pulses produced by passing through an A.C. amplifier has been cancelled.

It will be seen that if the leakage through the grid resistance R is too slow, i.e. if the time constant $C_1 R$ is too long, it will be some time before the points Q of Fig. 12 are enabled to touch the line of zero grid potential as in Fig. 13, i.e., the faithfulness of D.C. restoration will be impaired. On the other hand, if the time constant is too short, leakage will occur during each pulse, thus giving it a sloping top which will mar the faithfulness of reproduction of the pulses themselves. The time constant must therefore be chosen with care and a compromise must be effected between these two conflicting requirements.

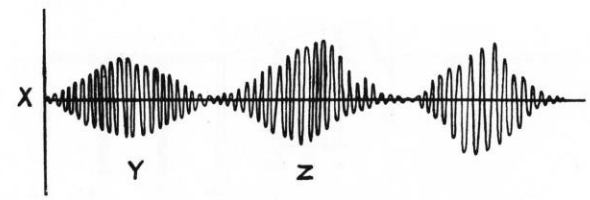

Figure 14

It now becomes possible to understand the relationship between D.C. restoration and ordinary rectification. Clearly there is some relationship, because the circuit used in both cases is exactly the same. In order to make this clear, we shall now consider the rectification of the modulated carrier wave as carried out in a sound receiver in terms of D.C. restoration. Fig. 14 shows the familiar radio frequency waveform which we have to rectify in order to extract the component of low frequency modulation. In a radio receiver it is applied to a circuit in principle identical with Fig. 4.

Let us first consider what would happen if the condenser and resistance of Fig. 4 were both very large. In this case, on application of the waveform of Fig. 14, grid current would be drawn by the successive pulses of increasing amplitude occurring between X and Y, and as usual none of these pulses would be able to drive the grid more positive than zero. Owing to the high time constant little or no leakage would take place between Y and Z, or between Z and any subsequent high amplitude pulse, and D.C. would be re-established along a datum line connecting Y and Z, and the modulated waveform would in fact act on the valve in the position shown in Fig. 15.

Assuming that the total double amplitude of the waveform did not drive the valve operating point as far as the bottom bend of the anode current grid volts characteristic, the valve would give class A amplification and no rectification at all, i.e., there would be a most excellent degree of D.C. restoration but no rectification. In practice, however, the time constant of C_1 R is made very small. This enables such a rapid leakage of charge

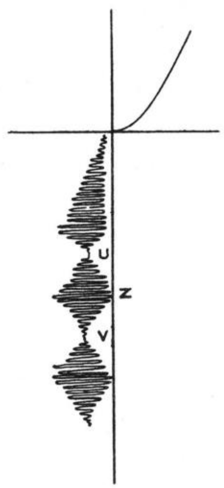

Figure 15

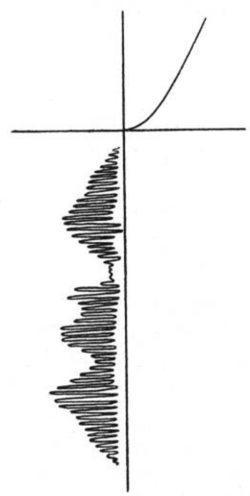

Figure 16

produced by grid current to take place that all the tips of the pulses coincide with the line of zero grid potential, and the gaps indicated by U and V in Fig. 15, where, owing to the high time constant, the tips of the pulses were not touching, are filled up. Thus, if the time constant is small the effect is as in Fig. 16.

Approaching the matter in the usual way from the point of view of rectification, we should say that to obtain a low frequency component from the modulated wave of Fig. 14, we should rectify this waveform, i.e. we should eliminate either that part which exists above or below the datum line. By simple elimination of, say, the lower half, the waveform of Fig. 14, would appear as in Fig. 17.

Thus, D.C. restoration and rectification are really the same process,
23 but two types of results are possible, one labelled *D.C. restoration*, being usually produced by the employment of apparatus in which the time constant of the grid circuit is comparatively long, and the other usually labelled *detection* or *demodulation*, obtained by using a comparatively short time constant. It would be unusual, but not incorrect, to state that in an ordinary radio receiver for sound broadcasting the detector is a D.C. restoring circuit operating upon the radio frequency waveform. The detector is in reality restoring D.C. but the D.C. itself is varying comparatively rapidly, in fact at audio frequency, and if the time constant of the apparatus be made short enough to enable the process of D.C. restoration to follow these rapid audio frequency variations, then the net result is what is ordinarily termed detection or demodulation.

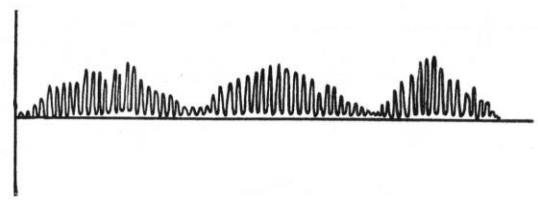

Figure 17

It will be noted that the method of restoring D.C. which has been described above is not perfect in that the time constant is so chosen that the best compromise is effected between the mutually conflicting requirements of avoiding a fall of D.C. during a line and yet of responding with reasonable rapidity to variations in D.C. level. There is, however, another method of D.C. restoration which gives a much more accurate result, and which is used in the modulator of the transmitter for restoring D.C. for the last time in the transmission chain. After D.C. has been restored by this improved process it is not lost again.

It is a feature of this improved method that the datum line does not necessarily have to coincide with zero grid potential of the valve at which D.C. restoration is effected, as it does in the less perfect method, but the datum may be established at any desired grid potential. The fundamental circuit is shown in Fig. 18. In this circuit the signal to which D.C. is to be restored is applied to the grid of the valve V via the condenser C. A source of potential, indicated by the battery B, is connected across the potentiometer P, and a circuit may be made via the switches S_1 S_2 through the resistances R_1 R_2.

Before considering how D.C. restoration is effected in this circuit, let us investigate the value of potential which will be found at the point X

when the switches S_1 and S_2 are held closed. If the slider of the potentiometer P is fixed at the point Y, then the negative end of the battery B is connected to earth and the cathode of the valve V, and the point X is therefore at a positive potential above earth, dependent upon the ratio of the resistances R_1, R_2. If the slider is fixed at the other end, Z, then the positive end of the battery is earthed and the point X now has a negative potential with respect to earth. Thus, when the slider is at Y, the valve receives positive grid bias and when at Z, negative grid bias. Thus, if R_1 and R_2

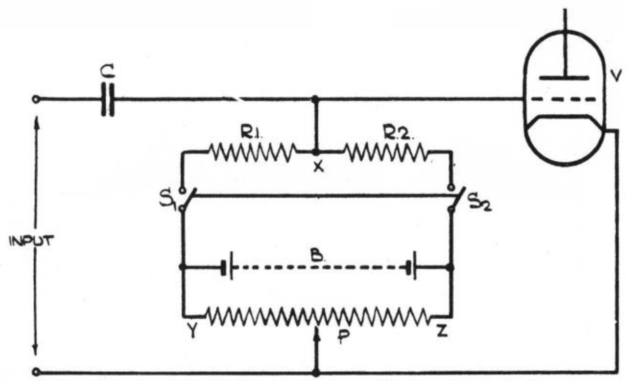

Figure 18

remain fixed, the potential applied to the grid of the valve V will be determined by the position of the slider over the potentiometer P and may be positive, zero or negative.

Let us now assume that S_1 and S_2 are closed and P is fixed at a point which will place a certain negative potential on the grid of V. Then, if S_1 and S_2 are opened this negative potential will remain upon the grid of V, being held there by the condenser C, as there is no grid leak. Let us now suppose that an ordinary vision waveform, together with suppression and synchronising signals, as shown in Fig. 19, is applied to the input terminals of Fig. 18, and let us arrange that the switches S_1 S_2 are maintained open except during the brief intervals t, forming part of the duration of the suppression pulses, as shown in Fig. 19. The suppression pulse is, by definition, a small duration of time during which all valves of the system should be at their proper black level. If, therefore during the suppression pulses the switches S_1 S_2 are closed, then a certain negative potential, determined by the position of the slider of the potentiometer P, is impressed upon the grid of the valve V. The switches S_1 S_2 are then opened before the end of the suppression pulse, and the next line of vision signals together with the synchronising signal is impressed, and the potentials on the grid of the valve V must make excursions which are entirely on the positive side of that potential established by the potentiometer P.

Thus, the restoration of D.C. has been effected at the end of the line of vision signals and the synchronising pulse. The beginning of the next suppression pulse should bring the grid of the valve V to the potential

Figure 19

previously imposed upon it by the closing for the first time of the switches S_1 S_2 during the previous suppression period, but in any case during the next suppression pulse the switches S_1 S_2 are closed again, thereby placing once more on the grid the potential determined by the potentiometer P. This potential is, of course, adjusted by means of P to be the value desired for the black level of the grid of the valve V.

In practice the suppression pulses occur at the rate of 10125 per second and therefore the switches S_1 S_2 are opened and closed at this rate, and during every suppression pulse the correct black level potential is imposed on the grid of the valve V.

It will now be seen why this arrangement is so superior to the alternative method, first described. The disadvantage of the first method is that the continuous presence of the grid resistance R of Fig. 4 in the circuit the whole time would lead, if it were not high enough, to a fall of D.C. during a line, whereas if it were too high variations in D.C. could not be followed. In the circuit of Fig. 18, however, this does not happen because in effect the switches S_1 and S_2 are removing the grid leak from circuit during the whole of the time of the vision signal, and except for the small period, t, the valve has no grid leak whatever. Thus, there cannot be any fall of D.C. during a line, and the restoration of D.C. is perfect.

Furthermore, the question of following fluctuations of D.C. does not now arise because, the black level having been established, the valve is free to make any excursion imposed on it by the vision input, and variations in D.C. are therefore automatically followed.

In practice it is, of course, quite impossible to interpret S_1 and S_2 of Fig. 18 literally as switches mechanically operated 10125 times per second,

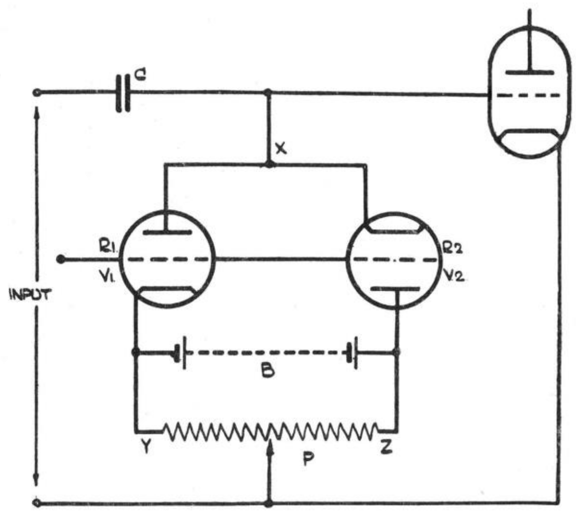

Figure 20

but clearly the switch S_1 and the resistance R_1 can be substituted by a valve, and the switch S_2 and the resistance R_2 by another, and if the valves have negative bias on them the effect will be as if S_1 and S_2 were open, and if the bias is removed then the effect will be as if S_1 and S_2 were closed. The

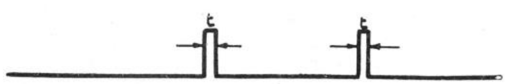

Figure 21

practical circuit therefore will appear as in Fig. 20. In normal operation the valves V_1 V_2 will be held biassed back beyond cut-off at all times except for the short periods t of Fig. 19, when the bias will be removed, thus permitting them to constitute the resistances R_1, R_2. This may clearly be done by applying to the grids of V_1 and V_2, in parallel, a pulse of the type shown in Fig. 21, of frequency 10125, consisting of short positive periods having a duration equal to the time t.

25

Application of D.C. Restoration in M-E.M.I. Vision Circuits

In the foregoing section of this technical note, D.C. restoration has been considered purely as a means whereby proper tonal reproduction of the picture may be secured and yet A.C. amplifiers may be used in preference to a complete chain of D.C. amplification, which would impose severe practical problems. There are, however, various other reasons for the restoration of D.C., and a representative selection of these will now be discussed.

In the Suppression Mixer, where we wish to remove certain spurious signals by combining with the picture signals certain suppression pulses designed for the purpose, we wish to fit the suppression pulses and the vision signals together so that they always bear the same relationship. The particular way in which the two sets of signals are fitted together determines what in the original picture shall come out as dead black in the received picture and as is indicated more fully in my technical note on the Suppression Mixer, D.C. is restored to both picture and suppression signals, in order that a constant datum line may be established carrying the vision signals on one side and the suppression pulses on the other. If this were not done, variations in the total amount of white in the picture would cause variations in the line of cut of the suppression pulses, and the ultimate effect would be that in the received picture the detail in the darker portions would be sometimes good and sometimes bad, the colour of the darker portions would vary and the overall intensity of the picture would fluctuate.

In the Picture and Sync. Mixer, D.C. is restored both to the picture signals and the synchronising signals. Here again it is necessary to fix the common datum line upon one side of which will exist the picture signals and on the other side the synchronising signals, as if this were to fluctuate it would be impossible to maintain steady synchronising.

In the Suppression, Black-out and other generators, D.C. is restored to the generated pulses for the following reason. When a pulse has been generated by a multivibrator or other means, it is rarely a clean square-topped pulse, but the peaks may often have what are best described as 'whiskers' attached to them. If such pulses be applied to a valve at a greater amplitude than the operating characteristic of the valve can take, then the tops of such pulses will be cut off by the bottom ~~band~~ bend of the characteristic, and these 'whiskers' will be removed. In such a case it is clearly desirable to stabilise the position of the other end of the pulse, i.e. the end not being square-topped, which may, of course, be done by D.C. restoration. It is thereby ensured that the pulse, not being able to carry the grid to a more positive value than zero, is obliged to sit at a fixed position on the operating characteristic, and there is no possibility of its avoiding the bottom bend limitation and consequent squaring imposed by the limited valve characteristic.

THE CATHODE FOLLOWER

In modern television technique it will be found there is a wide application of a type of valve circuit known as the cathode follower, and from my various technical notes on the M.-E.M.I. System it will be noted that there is one of these circuits in practically every unit of the System. The circuit is of great importance and provides an excellent solution to more than one difficulty. It is, however, perhaps one of the simplest circuits that has ever been devised. It is of importance that the principles and applications should be thoroughly understood, and they are discussed in this note.

In television technique, just as in that of sound, there arises the problem of providing an amplifier or other unit through which television signals are passing with a suitably low output impedance by means of which the unit may properly feed a line. It may either be necessary to generate an output impedance specifically equal to the characteristic impedance of the line if the latter be long, or to generate an impedance which is low compared with the impedance of the capacity of the line at the highest working frequency if the line be short. In sound apparatus such a low impedance is conveniently supplied by means of an output transformer, and the range of frequencies to be accommodated is such that practically distortionless transformers can be designed for this purpose. In television, however, where the frequency band extends from zero to more than 3 Mc/s it is not possible to create a suitable transformer. It would be possible to find a somewhat unhappy solution to this problem by the use of a circuit involving many parallel connected valves, but it may be solved simply and neatly by employing the cathode follower. Primarily then we may regard the cathode follower as an electronic transformer, or valve circuit, which substitutes for a step-down transformer.

This is not, however, the only application of this circuit. It is frequently found necessary to connect apparatus, having a high internal impedance and carrying the vision frequencies from zero to 3½ Mc/s, to apparatus having an input capacity which is sufficiently great as seriously to attenuate the upper frequencies. The problem necessitates the insertion of a piece of apparatus which will present a negligible input capacity to the sending apparatus, have a high internal impedance, and present a low output impedance to the subsequent apparatus so that the input capacity of this will have no effect. This again is solved by the application of the cathode follower, as not only has this circuit the property of generating a low output impedance and thus simulating an output transformer, but its input capacity can be given a very small value.

There is yet another most useful application which follows at once from the fact that the output impedance of the cathode follower is low over a wide frequency band, including a frequency of zero, i.e. D.C. It is of the highest importance that vision amplifiers and other units through which vision signals are passing, should be supplied with H.T. from a source having as low an impedance as can be obtained, since if this is not so the various signal amplitudes will not be reproduced in relatively correct proportion, and the relative intensities of the tones in the picture corresponding to such signals will be incorrect. We may secure such a low impedance by inserting a cathode follower between the source of H.T. and the amplifier, the impedance then presented to the latter being that of the cathode follower and not of the H.T. supply unit. Since the impedance is low, variations of H.T. voltage will be very small, and the H.T. is now said to be stabilised. A cathode follower used for this purpose is known as a stabiliser.

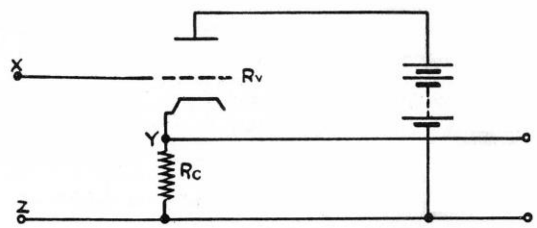

Fig. 1

The cathode follower then has four general applications :—

(1) As a circuit which generates a very low output impedance over a very wide frequency band including zero frequency, or D.C.

(2) As a circuit which presents a very high input impedance over a very wide frequency band.

(3) As a circuit in which, by simultaneously using the properties outlined in (1) and (2) above, a wide range amplifier may be formed, and

(4) as a circuit for reducing the impedance, i.e. improving the regulation, of power supplies, notably those for high tension voltages, and consequently stabilising these voltages.

The fundamental circuit of the cathode follower is shown in Fig. 1, from which it will be seen that it consists simply of a valve and a load resistance, the latter being placed in the cathode circuit. The anode circuit contains only the source of H.T., and the impedance of this must be very low ; in fact the proper functioning of the cathode follower is dependent upon the anode to earth impedance being as near zero as possible. Where a battery cannot be used this condition must be approximated by connecting between anode and earth a large condenser. For simplicity in subsequent diagrams the source of H.T. will be omitted and its impedance will be assumed to be zero.

It might be thought that the circuit of Fig. 1 is simply that of a resistance amplifier in which the load resistance has been placed in the cathode circuit instead of in the anode circuit, but this change is one which fundamentally alters the operation of the circuit, its behaviour being altogether different from that of the straightforward resistance amplifier. This is due to the fact that the presence of the cathode resistance R_c between the cathode and earth considerably modifies the potentials existing between the grid and cathode. It must be remembered that a valve amplifies those potentials which exist between its grid and its cathode, and not necessarily those between grid and earth. If the cathode is earthed by direct connection or by a large condenser, then the potentials between grid and cathode and between grid and earth are the same, but the essential feature of the cathode follower is that there is an impedance between cathode and earth, and that the voltages energising the valve now consist of two components: (1) those deliberately applied at the input, i.e. between the terminals x and z of Fig. 1, and (2) those occurring between the terminals y and z, i.e. those across the cathode resistance R_c. The effect of the presence of this additional voltage component yz. as has been stated earlier, is to give the circuit a very low output impedance and a very high input impedance, the values being respectively much lower and higher than those obtained with the same valve used as a resistance amplifier. Furthermore the amplification has a maximum value of unity.

The output impedance may be designed to be a specific figure, and, as is often required, this figure may be designed to be that of the characteristic impedance of a line, so that accurately matched impedance conditions at the sending end can be obtained. In fact the cathode follower looks to the line like the equivalent circuit of Fig. 2, consisting solely of a generator of internal impedance R_i and of voltage e_i. We shall now proceed to establish a formula for the output or internal impedance of the cathode follower R_i, in order to show firstly that it is unusually low, and secondly how it may be designed to match the impedance of a line exactly.

The Output Impedance

We shall first of all find the formula giving the internal impedance R_i of Fig. 2 in terms of the resistance of any load which may be placed across it $-R_l$ as in Fig. 3, the voltage across the load $-e_l$, and the voltage of the equivalent generator $-e_i$, and then apply this fundamental formula to the cathode follower circuit of Fig. 1.

We have $\quad e_l = e_i \cdot \dfrac{R_l}{R_i + R_l}$

therefore $\quad R_i = R_l \left(\dfrac{e_i}{e_l} - 1\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (1)$

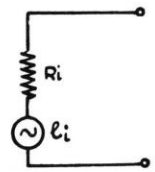

Fig. 2

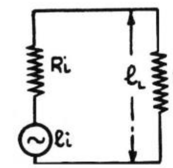

Fig. 3

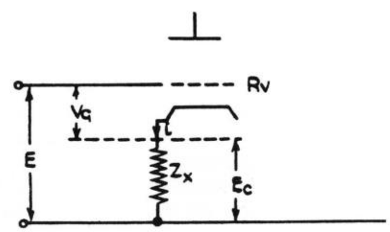

Fig. 4

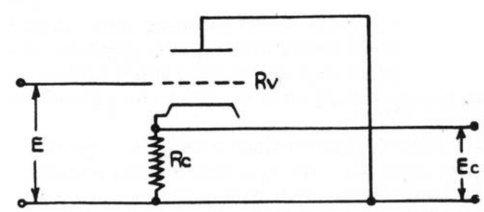

Fig. 5

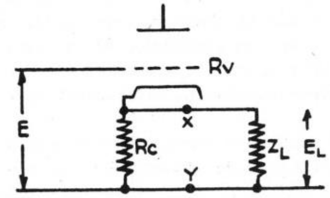

Fig. 6

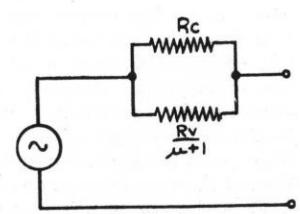

Fig. 7

This is our fundamental formula, and we must now apply it to the cathode follower by determining what will be the values of e_i and e_l in the latter circuit. Let these new values be E_i and E_l.

Considering now the cathode follower redrawn in Fig. 4, let us find the value of E_c for any total load Z_x included between cathode and earth.

We have $E_c = i\,Z_x$

and $i = \dfrac{\mu\,V_g}{R_v + Z_x}$

therefore $E_c = \dfrac{\mu\,V_g\,Z_x}{R_v + Z_x}$

Now $V_g = E - i\,Z_x = E - E_c$

therefore $$E_c = \frac{\mu\,Z_x\,(E - E_c)}{R_v + Z_x} = \frac{\mu\,Z_x E}{R_v + Z_x\,(\mu + 1)} \qquad (2)$$

We can now find E_i by considering the circuit of Fig. 4 to have a cathode resistance R_c instead of Z_x and regard it as a source of voltage E_i, as shown in Fig. 5; then substituting R_c for Z_x, and E_i for E_c in (2), we have

$$E_i = \frac{\mu\,R_c E}{R_v + R_c\,(\mu + 1)} \qquad \ldots\ldots\ldots\ldots\ldots \quad (3)$$

We have now to find E_l. Let the circuit of Fig. 5 feed a load Z_l as in Fig. 6, and let the voltage across Z_l be E_l. E_l will be given by equation (2) if we write

$$Z_x = \frac{R_c Z_l}{R_c + Z_l}$$

and consequently we have

$$E_l = \frac{\mu\,E\,R_c Z_l}{R_v\,(R_c + Z_l) + (\mu + 1)\,R_c Z_l} \qquad \ldots\ldots \quad (4)$$

We can now establish R_i, the output impedance of the cathode follower, by substituting E_i of (3) and E_l of (4) for e_i and e_l of (1). Also $R_l = Z_l$. Therefore we have by substitution

$$R_i = Z_l\left(\frac{\mu\,R_c E}{R_v + R_c(\mu + 1)} \times \frac{R_v(R_c + Z_l) + (\mu + 1)R_c Z_l}{\mu\,E\,R_c Z_l} - 1\right)$$

from which $$R_i = \frac{R_v R_c}{R_v + R_c(\mu + 1)} \qquad \ldots\ldots\ldots\ldots\ldots \quad (5)$$

Equation (5) gives the output impedance of the cathode follower accurately, and is of interest in that its right-hand side expresses an impedance which would be formed by the connection in parallel of two impedances R_c and $\dfrac{R_v}{\mu + 1}$ of which the latter at a glance contains the reason why the output impedance of this circuit is so low, for approximately it consists of

29

the impedance of the valve divided by its magnification factor. The equivalent circuit of the cathode follower may therefore be accurately drawn as in Fig. 7.

The output impedance may for practical purposes be expressed to a close approximation. We simply neglect 1 in comparison with μ, and we have

$$R_i = \frac{R_v R_c}{R_v + \mu R_c}$$

Now the mutual conductance of a valve $S_m = \dfrac{\mu}{R_v}$, therefore we may write

$$R_i = \frac{R_v R_c}{R_v + R_c R_v\,S_m} = \frac{R_c}{1 + R_c\,S_m} \qquad \ldots\ldots \quad (6)$$

Neglecting 1 in comparison with $R_c\,S_m$, then

$$R_i = \frac{1}{S_m} \qquad \ldots\ldots\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

This expresses the extremely simple result that the output impedance of the cathode follower is equal to the reciprocal of the mutual conductance of the valve if the two approximations made above can be upheld, which is in general easily possible.

As an example of the order of low impedance which may be obtained, we may take a valve type H.30 having a mutual conductance of 6 mA/V, in which case $R_i = \dfrac{1}{S_m} = \dfrac{1000}{6} = 166$ ohms.

It should be noted that if the resistance R_c is omitted, the output impedance is then accurately expressed by $\dfrac{R_v}{\mu + 1}$

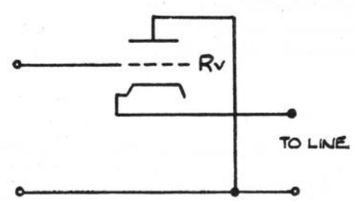

Fig. 8

Design of Cathode Follower to Match the Characteristic Impedance Z_o of a line

Given a line of known characteristic impedance Z_o, we may design a cathode follower to match this exactly. In most cases, however, we may neglect the errors produced by the two assumptions made above as they are extremely slight, and the assumptions will accordingly be made. The problem presents three cases:

(1) Where Z_o happens to equal $\frac{1}{S_m}$. In this case $R_c = \infty$, and no resistance R_c need be provided and the circuit will be as in Fig. 8.

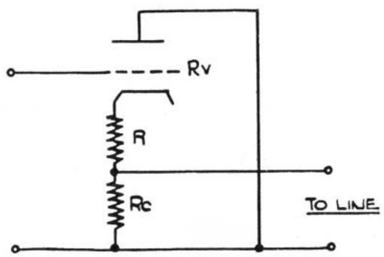

Fig. 9

(2) Where Z_o is less than $\frac{1}{S_m}$. In this case we must choose a value of R_c such that in parallel with $\frac{1}{S_m}$ will give an output impedance equal to Z_o. Clearly from equation (6) the value of R_c will be given by $R_c = \frac{Z_o}{1 - S_m Z_o}$

(3) Where Z_o is greater than $\frac{1}{S_m}$. In this case we cannot obtain an exact match by any circuit so far shown, but the output impedance may be raised to the correct value by the insertion of a small padding resistance R, as shown in Fig. 9. A very simple formula for R results if we make the following three assumptions:

(a) That 1 is neglected in comparison with μ.
(b) That R is neglected in comparison with R_c.
(c) That 1 is neglected in comparison with $S_m R_c$.

We then have that

$$R_i = \frac{1}{S_m} + R \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (8)$$

and that to match a value of characteristic impedance Z_o, which is greater than $\frac{1}{S_m}$, we must choose a value of R given by $R = Z_o - \frac{1}{S_m}$.

A detailed proof of the validity of (8) is given in Appendix 1.

The Input Impedance

It has been stated that the input impedance of the cathode follower is very high due to the fact that the input capacity is much less than that of a triode amplifier. In the latter the well known Miller effect is present, by which is meant that there is effectively thrown across the input a capacity approximately μ times the anode grid capacity in addition to the normal grid cathode capacity already existing. The total input capacity of the triode amplifier therefore is in practice sufficiently great as to render this form of amplifier useless in vision circuits as it stands.

In the cathode follower, however, the potentials on the cathode are in the same phase as those on the grid, and in general are of nearly the same amplitude. The name cathode follower is derived from the fact that the potentials on the cathode follow those on the grid. Accordingly, to a rough approximation, there can be no capacity between grid and cathode. There can also be no Miller effect because there is no anode impedance, and the only capacity present is really that of the grid to the anode. Consequently the total input capacity is very low indeed, and it can be shown by

$$C = C_{ga} + C_{gc}\left(\frac{R_v + Z_x}{R_v + \mu Z_x}\right)$$

where C_{ga} = the anode to grid capacity,
C_{gc} = the grid to cathode capacity,
and Z_x = the total cathode impedance.

Amplification of Cathode Follower

The gain of the cathode follower must always be less than unity. The gain on open circuit, i.e. with the line or other load not applied, will, of course, be $\frac{E_i}{E}$ and the gain with the load connected will be $\frac{E_l}{E}$. These may be easily derived from the preceding equations, and it is only necessary to quote here the results.

For case (2), where Z_o is less than $\frac{1}{S_m}$ and a cathode resistance R_c is used, the open circuit gain may be written down at once from equation (3) as

$$A = \frac{\mu R_c}{R_v + R_c(\mu + 1)}$$

The gain with any load Z_l connected is, from equation (4),

$$A = \frac{\mu R_c Z_l}{R_v(R_c + Z_l) + (\mu + 1) R_c Z_l}$$

For case (3), where Z_o is greater than $\frac{1}{S_m}$ and a padding resistance R is used, the open circuit gain will be, from Appendix 1,

$$A = \frac{\mu R_c}{R_v + (R_c + R)(\mu + 1)}$$

and the gain with any load Z_l connected will be given by

$$A = \frac{\mu R_c Z_l}{(R_c + Z_l) R_v + (\mu + 1)(RR_c + RZ_l + R_c Z_l)}$$

Use of Cathode Follower to form an Amplifier

If an amplifier be formed as in Fig. 10 in which alternate valves are respectively cathode followers and triode amplifiers, then the amplifier so constituted forms a basis of a design which will accommodate a wide band of frequencies as is required for television. Considering Fig. 10, the valve V_1

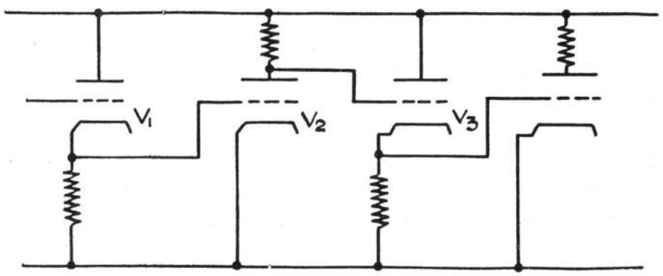

Figure 10

presents a low output impedance to V_2 so that the comparatively high input capacity of V_2 will have a much smaller effect than if V_1 were an amplifier. The valve V_2, having a comparatively high output impedance, needs to feed a circuit having an input impedance as devoid of capacity as possible, and this is provided by the cathode follower V_3, and so on throughout the chain.

The Cathode Follower as a Stabiliser

It has already been pointed out that it is of the highest importance in much television apparatus that the internal impedance of the H.T. supply should be low, and as it is possible to design a cathode follower whose output impedance is much lower than the usual figure for a mains H.T. supply unit, it is of advantage to feed the H.T. to such apparatus through a cathode follower.

The fundamental circuit is given in Fig. 11, where A is an amplifier whose H.T. supply is to be stabilised, and V is a valve connected as a cathode follower, the amplifier A constituting the cathode load for the valve V. It is necessary to apply to the grid of V a positive bias, as otherwise the voltage drop across the amplifier A, i.e. its high tension voltage, constitutes a fixed negative bias for V, and only a small current would flow. R represents the regulation resistance of the H.T. supply.

It is considered that this circuit may be best understood by approaching it in the above manner. Strictly speaking, however, the input to a stabiliser is that voltage applied between grid and earth, and it is this voltage which appears slightly reduced at the output terminals, i.e. between cathode and earth. We should therefore properly regard the amplifier H.T. which is
31 taken from the cathode circuit of the cathode follower as being derived from the positive voltage applied to the grid. Clearly the voltage of the amplifier H.T. will be determined by that applied to the grid of the stabiliser, and any adjustment in value would be made by adjusting the grid positive voltage. Naturally the power drawn by the amplifier will be initially drawn from the H.T. mains unit, and will pass to the amplifier via the anode circuit of the cathode follower.

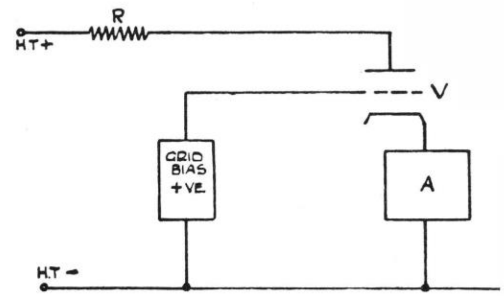

Fig. 11

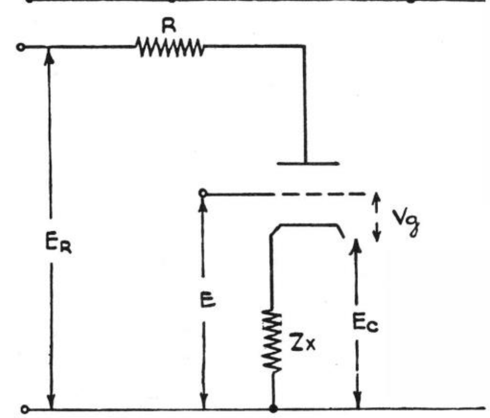

Fig. 12

We must now determine what will be the internal impedance of the H.T. supply as formed by the supply unit in cascade with the stabiliser; clearly it will be approximately equal to the output impedance of the stabiliser valve V, but it is necessary to take into account the value of the regulation resistance R, of the H.T. supply unit, a factor which has been ignored in all previous calculations. The equivalent circuit is shown in Fig. 12. We first proceed as before to determine the value of the output voltage E_c with a total load Z_x between cathode and earth.

Let E_r be the value of applied H.T.

We have $E_c = iZ_x$

and $i = \dfrac{E_r + \mu V_g}{R_v + Z_x + R}$

therefore $E_c = \dfrac{(E_r + \mu V_g) Z_x}{R_v + Z_x + R}$

but $V_g = E - E_c$

therefore $E_c = \dfrac{\mu Z_x(E - E_c) + E_r Z_x}{R_v + Z_x + R} = \dfrac{\mu Z_x E + Z_x E_r}{R_v + R + (\mu + 1) Z_x}$

$$= \frac{E_r + \mu E}{\dfrac{R_v}{Z_x} + \dfrac{R}{Z_x} + (\mu + 1)}.$$

As before we now determine E_i. This will be the value of E_c when $Z_x = \infty$

and therefore $E_i = \dfrac{\mu E + E_r}{\mu + 1}$

We can now determine E_l, the load voltage, which will be that voltage existing across Z_x when Z_x is equal to the equivalent resistance of the amplifier. Writing the latter as Z_l, then let $Z_x = Z_l$, in which case

$$E_l = \frac{\mu Z_l E + Z_l E_r}{R_v + R + (\mu + 1) Z_l}$$

As before, by substituting E_i for e_i, E_l for e_l, and Z_l for R_l in (1), we have

$$R_i = Z_l \left(\frac{\mu E + E_r}{\mu + 1} \times \frac{R_v + R + (\mu + 1) Z_l}{Z_l (E_r + \mu E)} - 1 \right) = \frac{R_v + R}{\mu + 1}$$

The above expression shows that R behaves purely as an addition to R_v, the internal resistance of the valve.

Neglecting as usual 1 compared with μ, we have finally

$$R_i = \frac{R_v + R}{\mu} = \frac{Rv}{\mu} + \frac{R}{\mu} = \frac{1}{S_m} + \frac{R}{\mu}$$

This expression shows that the impedance produced by the stabiliser is constituted by the low value of $\dfrac{1}{S_m}$ augmented only by the regulation resistance of the H.T. supply unit divided by the magnification factor of the stabiliser valve, and in most practical cases the value of R_i is much lower than the value of R, so that the insertion of the stabiliser considerably reduces the regulation resistance of the H.T. supply unit. The value of the amplifier H.T. is, of course, given by E_l.

The action of the stabiliser can be physically described as follows. If the amplifier A attempts to take more current, the drop of voltage across it, i.e. its H.T., will decrease, but since this voltage is the negative bias for the valve V, the difference between the value of this voltage and that of the positive grid bias supply will become more positive, so that V will tend to pass more current and thus restore the H.T. to a value very nearly equal to the initial value.

The Stabiliser as a Smoothing Circuit

It has so far been assumed that the initial H.T. supply derived from the mains supply unit is maintained at a constant value, and it is of interest to consider what will happen in the stabiliser circuit if there is a fluctuation in this voltage due possibly to some disturbance in the mains. Let the initial voltage be E_r and let this be increased by some such disturbance to a value of $E_r + \delta E_r$. This increase of voltage δE_r will produce an increase of anode current δi, and this in turn will modify the voltage across Z_x and effect a change in the effective grid voltage by an amount δVS_m. We may accordingly write

$$\delta E_r = \delta i\ (R + R_v + Z_x) + \mu \delta VS_m$$

Remembering that μ is a negative factor, we may write

$$\delta VS_m = \delta E_c = \delta i Z_x$$

from which $\delta E_r = \delta i\ (R + R_v + (\mu + 1)\ Z_x)$

therefore $\delta i = \dfrac{\delta E_r}{R + R_v + (\mu + 1)\ Z_x}$

The resultant change in output volts δE_l will be $Z_l \delta i$.

Therefore $\delta E_l = Z_l \delta i = \dfrac{Z_l \delta E_r}{R + R_v + (\mu + 1) Z_l}$

$$= \frac{\delta E_r}{\dfrac{R + R_v}{Z_l} + (\mu + 1)}$$

Making a first assumption that $\dfrac{R + R_v}{Z_l}$ will be negligible compared with μ, as will be the case in practice, we may write

$$\delta E_l = \frac{\delta E_r}{\mu + 1},$$

Making a second assumption, as usual, that 1 may be neglected in comparison with μ, we may finally write

$$\delta E_l = \frac{\delta E_r}{\mu}$$

This equation expresses the fact that not only does the stabiliser tend to maintain a constant H.T. voltage during variations of amplifier load, but it also effects a reduction of fluctuations in the initial H.T. supply, reducing them by a factor approximately equal to the magnification factor of the stabiliser valve. This double application of the stabiliser is, of course, very valuable.

The action of the stabiliser in reducing the effect of mains fluctuations may also be physically described as follows. Should a fluctuation occur, tending to increase the H.T. supply, the stabiliser will pass more current which increases the self bias produced by the cathode load resistance, and this tends to reduce the current passed by the stabiliser, thus restoring its output voltage to a value very near to the original.

Appendix 1

To establish the validity of equation (8), we must proceed to find new values of e_i, e_l, and consequently R_i for the new conditions. It will first be necessary to establish a new value for the E_c of equation (2). The cathode resistance is now equal to $R_c + R$. We may therefore refer to equation (2) and write

$$Z_x = R_c + R$$

therefore $$E_c = \frac{\mu E(R_c + R)}{R_v + (\mu + 1)(R_c + R)}$$

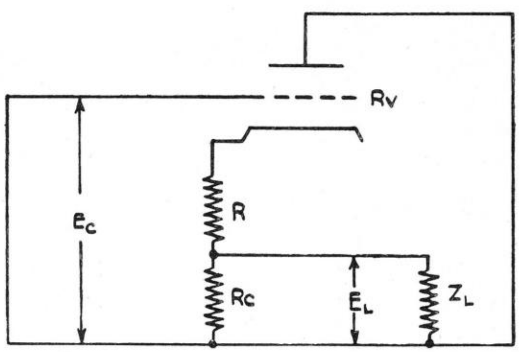

Fig. 13

The open circuit voltage E_i will now be that across R_c

therefore $$E_i = \frac{R_c}{R_c + R} E_c = \frac{\frac{R}{(R_c + R)}\,\mu E\,(R_c + R)}{R_v + (\mu + 1)(R_c + R)} = \frac{\mu E R_c}{R_v + (\mu + 1)(R_c + R)} \quad \ldots\ldots\ldots\ldots \text{(9)}$$

To establish E_l we must connect the load Z_l as in Fig. 13, and determine firstly the value of E_c, the voltage now existing between cathode and earth.

We refer again to equation (2), this time putting

$$Z_x = R + \frac{R_c Z_l}{R_c + Z_l}$$

therefore $$E_c = \frac{\mu E(RR_c + RZ_l + R_c Z_l)}{(R_c + Z_l)R_v + (\mu + 1)(RR_c + RZ_l + R_c Z_l)}$$

E_l will be that proportion of E_c existing across Z_l.

Therefore $$E_l = \frac{\dfrac{R_c Z_l}{R_c + Z_l}}{R + \dfrac{R_c Z_l}{R_c + Z_l}} E_c$$

$$= \frac{\mu E R_c Z_l}{(R_c + Z_l)R_v + (\mu + 1)(RR_c + RZ_l + R_c Z_l)} \quad \ldots \text{(10)}$$

Substituting (9) and (10) in (2), we have

$$R_i = \frac{R_v R_c + (\mu + 1) RR_c}{R_v + (\mu + 1)(R_c + R)}$$

Making a first assumption that 1 is very much less than μ, this reduces to

$$R_i = \frac{R_v R_c + \mu RR_c}{R_v + \mu(R_c + R)}$$

Making a second assumption that R is very much less than R_c, this reduces to

$$R_i = \frac{R_v R_c + \mu RR_c}{R_v + \mu R_c}$$

Eliminating μ by substituting $\mu = S_m R_v$, we have

$$R_i = \frac{R_c + S_m RR_c}{1 + S_m R_c}$$

Making a third assumption that 1 is very much less than $S_m R_c$, we have

$$R_i = \frac{R_c + S_m RR_c}{S_m R_c} = \frac{1 + S_m R}{S_m}$$

$$= \frac{1}{S_m} + R$$

33