# THE BELL SYSTEM TECHNICAL JOURNAL

VOLUME XXXIX | MARCH 1960 | NUMBER 2




# Fundamental Considerations in the Design of a Voice-Switched Speakerphone

By A. BUSALA

(Manuscript received November 23, 1959)

*The speakerphone offers the advantage of hands-free telephony by replacing the familiar handset with a separate microphone and loudspeaker. The convenience of such an arrangement justifies extensive efforts to overcome its acoustic and transmission limitations. Voice switching overcomes some of the limitations but adds problems of its own. This study analyzes these problems and outlines a method that has proved helpful in the design of a voice-switched speakerphone and the evaluation of its performance.*

## I. INTRODUCTION

A speakerphone[1] is a telephone whose familiar handset, which places the transmitter close to the talker's lips and couples the receiver tightly to his ear, is replaced by a separate microphone and loudspeaker that can be set on a table a few feet from the user. This arrangement gives the customer definite advantages. In addition to leaving his hands free during a telephone conversation, it is a great help for physically handicapped persons and offers other convenient features, such as the possibility of having a small group join in the conversation and the avoidance of fatigue during lengthy calls.

However, these advantages are obtained at the price of some limitations.[2] To make up for the loss introduced by moving the instruments away from the head, gain is required in both the transmitting and the receiving paths. This gain is limited by a "singing" problem. A signal from the microphone reaches the loudspeaker via the sidetone path and comes back to the microphone through the acoustic coupling in the room. Too much gain in this loop causes singing. Even before reaching this condition, the loudspeaker-to-microphone acoustic coupling is the source of other undesirable effects. Incoming speech at the speakerphone end is fed back from the loudspeaker through the acoustic path to the microphone, whence it is returned to the distant talker with a certain delay. This is a form of talker echo. Similar reasoning would show that the distant party is also subjected to listener echo. Furthermore, when a speakerphone is used, the acoustic properties of the room and the ambient noise level are important, and often have adverse effects. This is in contrast to the performance of the regular telephone, whose instruments are so close to the user's head that transmission is largely independent of the surrounding conditions.

Voice switching is an answer to some of these limitations. Obviously, a voice-operated device that would apply to the speakerphone circuit the Vodas technique,[3] allowing only one direction of transmission to be fully active at a time, would eliminate both the singing problem and the talker and listener echoes. But the application of voice switching has problems and limitations of its own. There is the inherent limitation of one-way-at-a-time communication, which certainly detracts from the naturalness of the conversation. This is a penalty that one would have to pay, but is probably well worth the advantages gained. Then there are the problems connected with the switching operation. It is the purpose of this study to make a systematic analysis of all these problems and to determine their relation to the circuit design.

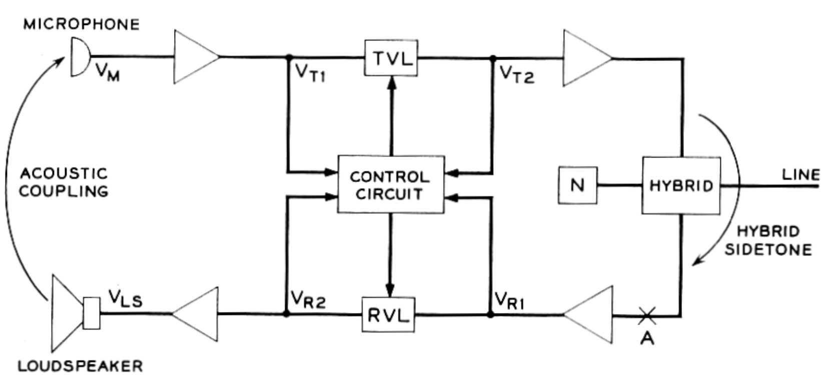


Fig. 1 — The master circuit.

## II. THE MASTER CIRCUIT

The general circuit of the essential elements of a voice-switched speakerphone is shown in Fig. 1, and will be referred to as the "master circuit". It consists of a microphone and a loudspeaker connected to a transmitting and a receiving branch with amplifiers to give the desired transmitting and receiving gains. A hybrid coil connects the transmitting and receiving branches to the line and to the sidetone-balancing network. With the practical limitations in sidetone balance and the inevitable acoustic coupling between loudspeaker and microphone, a substantial amount of loss must be inserted in the transmitting or the receiving branches to avoid singing. Therefore, a transmit variolosser, TVL, and a receive variolosser, RVL, have been included in Fig. 1. It is obvious that the loss will have to be switched* from one branch to the other, depending on whether the circuit is transmitting or receiving, and that one control circuit must operate simultaneously on both variolossers to insure that the minimum loss be always present.

The operation of the control circuit must be determined by some inputs. Four voltages whose ratio is not just a constant are available; these are two voltages, $V_{T1}$ and $V_{T2}$, in the transmitting branch and two voltages, $V_{R1}$ and $V_{R2}$, in the receiving branch. In general, all four voltages could be used simultaneously for control purposes.

## III. THE SUBMASTERS

The master circuit is general, but the use of four inputs to the control circuit makes it difficult to analyze its performance. Once the condition of the circuit (whether it is transmitting or receiving) and the amount of loss switched are known, the voltage $V_{T2}$ is related in a definite manner to $V_{T1}$, as is the voltage $V_{R2}$ to $V_{R1}$. This means that there are only two independent variables, and therefore two inputs — one voltage from the transmitting branch and one from the receiving branch — are sufficient for control purposes. There are four possible combinations of input voltages that can be used, and these correspond to four different circuit configurations, which will be called the "submasters". They are shown on Fig. 2.

If the control circuit is made intelligent enough, all the submasters can be made to give the same performance. Each one, however, has its own characteristics and will be examined individually later.

---

* While the term "switched gain" is often used in connection with a voice-operated device, the following description will be in terms of switched loss.

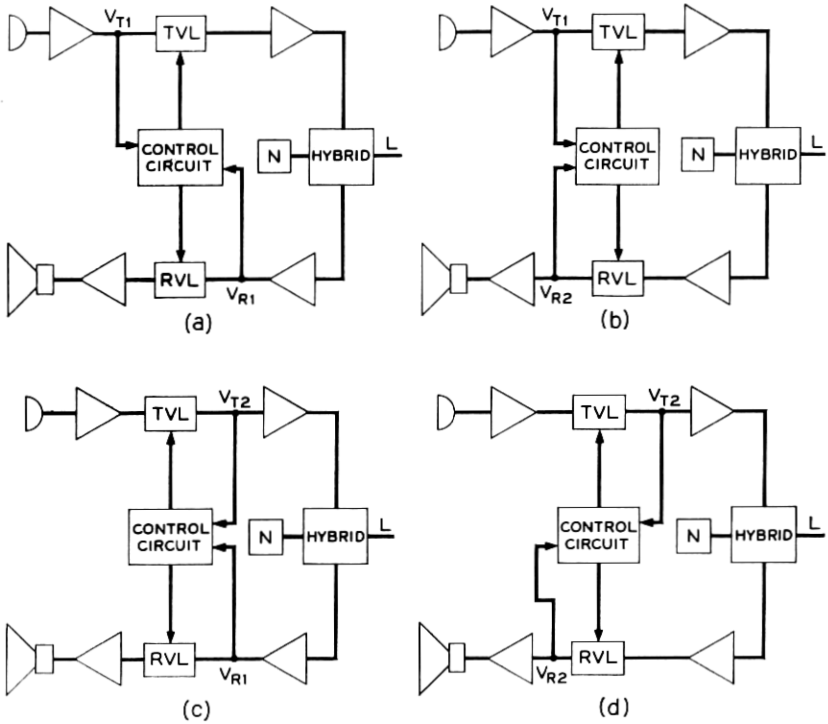


Fig. 2 — The submasters: (a) submaster #1; (b) submaster #2; (c) submaster #3; (d) submaster #4.

## IV. STEADY-STATE AND TRANSIENT PROBLEMS

The basic purpose of the voice-switching operation is to limit transmission (at least at full gain) to one direction at a time in order to avoid singing. It is the function of the control circuit to choose the direction of transmission on the basis of the two inputs and to operate accordingly on the variolossers. The critical parameters on which good performance of the speakerphone depends are the criterion applied by the control circuit to decide in which direction to have full gain and the time required to take the necessary action.

Starting, for instance, from a quiescent condition (which can be assumed to be the receiving one), the control circuit should switch the speakerphone into transmit as soon as there is a microphone input, in order not to lose the initial part of the signal. This implies that it would be desirable to have the control circuit very sensitive to the transmit input and to have a very short operating time. However, there are limitations to these requirements. For instance, if the circuit is too sensitive to microphone signals, room noise could easily impair reception of incoming signals. Also, if the control circuit were designed to switch back rapidly from transmit into receive every time the signal at the microphone fell below a certain threshold, the weaker sounds, especially consonants at the end of a word, would tend to be cut off and the transmitted speech would sound "choppy". Consequently, it seems desirable for the control circuit to possess a certain "hangover" in the transmitting condition after the signal has ceased. However, there are restrictions to the duration of this hangover. The longer the circuit remains transmitting, the more difficult it becomes for the distant party to break in. In other words, the initial part of an incoming signal might be lost. Therefore, the possibility of avoiding choppiness of the transmitted speech without impairing the received signal depends on the existence of satisfactory compromises for the sensitivity and the time constants of the control circuit.

From this typical example of switching operation, it is apparent that a good performance of the speakerphone is often a compromise in a number of characteristics, such as how fast it switches, how easily it stays in one condition once it has switched, how sensitive it is to room or line noise, or how easy it is to break in during a back-and-forth conversation.

In order to achieve the best compromise, it is helpful to classify the fundamental problems of voice switching into two groups, the steady-state and the transient problems. The ones in the first group concern the performance of the circuit in each condition, transmitting and receiving, with steady or slowly varying input signals. The ones in the second group concern its performance under transient conditions during the short time intervals when the input signals are rapidly varying.

The steady-state problems are:

i. Singing. This occurs, as already mentioned, when not enough loss is present in the transmitting or in the receiving branches.

ii. Transmit Blocking. In this case, the circuit is in the receive condition when it should be in the transmit condition. It can result from one of two conditions: a steady noise coming from the line may prevent the circuit from going into transmit, even with large signals into the microphone (transmit blocking due to noise operation); or a poor sidetone balance in the hybrid may generate a sidetone voltage large enough to hold the circuit in the receive condition (transmit blocking due to false switching).

iii. Receive Blocking. This problem is analogous to transmit blocking, and can also result from one of two conditions: a steady room noise may prevent the circuit from going into receive (receive blocking due to noise operation); or the acoustic coupling between loudspeaker and microphone may generate a pressure at the microphone large enough to hold the circuit in the transmit condition (receive blocking due to false switching).

The transient problems can be classified as follows:

i. Initial Clipping. This is the loss of the first part of a speech signal, and can occur either in transmitting or in receiving. It is, in general, a function of the operating time of the control circuit and switch.

ii. Final Clipping. This is the loss of the last part of a speech signal, and also can happen either in transmitting or in receiving. It is a function of the hangover time of the control circuit and switch.

iii. Echo. This problem is due to the finite decay time of the sound in a reverberant room. The sound put out by the loudspeaker is sustained in the room for a certain time after the incoming signal has ceased. This sound may be picked up and retransmitted by the microphone, giving the effect of a short, transient echo to the far end. The symmetrical problem of an echo effect in transmitting rather than in receiving is much subordinated, since there is no delay in the sidetone path through the hybrid, and any echo coming back from the line is very likely to be below a troublesome level.

## V. THE SWITCHING DIAGRAM

To analyze the performance of a speakerphone, it is helpful to use a plot that will be called the *switching diagram*. This graphical method offers the advantage over an algebraic analysis of showing at a glance the condition of the speakerphone and the margins against singing and blocking for any combination of input voltages. Furthermore, the diagram can be used to represent the effects of nonlinearity in the different parts of the circuit and the performance of the speakerphone during transients. In one convenient form of the diagram, the abscissa represents the voltage at the loudspeaker terminals and the ordinate the voltage across the microphone, with both voltages being on a logarithmic scale.

Rather than studying a single frequency, let us consider speech, assuming that the plotted voltages are weighted averages of some kind representing the speech power in the frequency band of interest. For simplicity, it will also be assumed that all the elements of the circuit have a flat response. Frequency shaping can be used advantageously, especially in eliminating some particular problems, such as receive blocking due to false switching and echo. Its effect could be analyzed in a similar manner by extending the application of the switching diagram.

### 5.1 *The Acoustic Coupling and the Hybrid Sidetone*

In the switching diagram one can plot two lines, which will be called the *acoustic coupling* and the *hybrid sidetone* [Fig. 3(a)]. The first represents the voltage that appears at the microphone terminals for a given voltage at the loudspeaker because of the acoustic coupling between loudspeaker and microphone. If all elements are linear, this is a 45° line whose position is a function of the acoustic environment, the distance between microphone and loudspeaker and the efficiencies of the two instruments. This line will be plotted for the limiting conditions in which we expect the speakerphone to operate, i.e., minimum distance between microphone and speaker and minimum room constant.[4] If one allows for overloading

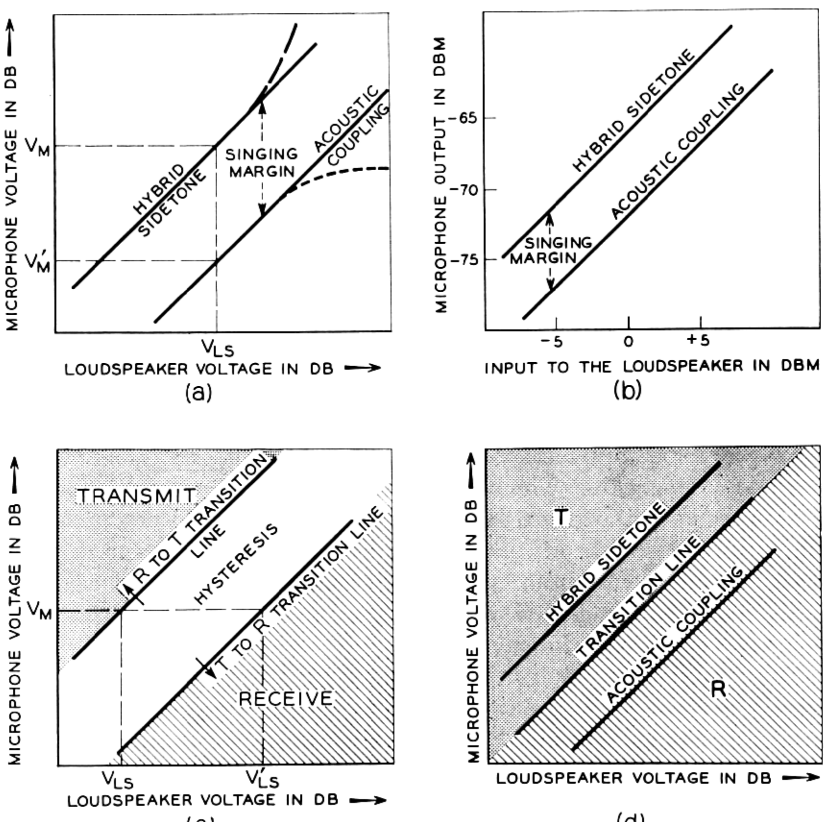


Fig. 3 — The switching diagram, showing sidetone, acoustic coupling and transition lines.

of the instruments, the acoustic coupling line will, in general, bend at high levels, as shown by the dotted line in Fig. 3(a).

The other line, the hybrid sidetone, represents the voltage appearing at the loudspeaker for a given voltage at the microphone. If all elements are linear, this is also a 45° line whose position is determined by the transmitting and receiving gain, the sidetone balance and the amount of loss switched. In general, the amount of loss switched in the transmitting branch may be different from the loss switched in the receiving branch. In this case, we would have two sets of hybrid sidetone lines, one obtained with the speakerphone in the transmit condition and one with it in the receive condition. The hybrid sidetone line for the poorest sidetone balance to be expected is plotted. If overloading of transmit and receive amplifiers is taken into consideration, the hybrid line will bend at high levels, as shown by the dashed line in Fig. 3(a).

An immediate consideration is that, to avoid singing, the hybrid sidetone line must lie above the acoustic coupling line. This is easily shown. A certain microphone voltage, $V_M$ [Fig. 3(a)], will generate a loudspeaker voltage, $V_{LS}$, because of electrical sidetone, and this in turn will develop a microphone voltage, $V_M'$, because of acoustic coupling. To prevent singing, $V_M'$ must be smaller than $V_M$, which is equivalent to saying that the hybrid sidetone line must lie above the acoustic coupling line. As noticed before, the position of the acoustic coupling line is well defined once the conditions (distance between microphone and loudspeaker and acoustic environment) under which the speakerphone is to operate are defined. The hybrid sidetone line, on the other hand, can be shifted in the plane by changing the amount of loss switched. The greater the loss switched, the farther the hybrid sidetone line moves above the acoustic coupling line. The amount of loss switched can, therefore, be used in the design to insure that the circuit is free from singing problems. Since these two lines are plotted for voltages that represent average speech energy, a certain margin must be included to take care of deviations from the average due to standing waves in the room or poor sidetone balance at single frequencies.

This situation may be further illustrated by an example. Fig. 4 shows the signal levels in a typical speakerphone. For a power at the loudspeaker terminals of 1 dbm,* the pressure at the microphone 30 inches away in a room with a constant of 800 is found to be 71 SL.† For a certain transmitter this corresponds to a power available at the microphone

* Dbm is the power in decibels referred to one milliwatt.

† SL, sound level, is the weighted sound pressure level at a point in a sound field, in accordance with ASA Z 24.1, 1951, Definition 1.380.

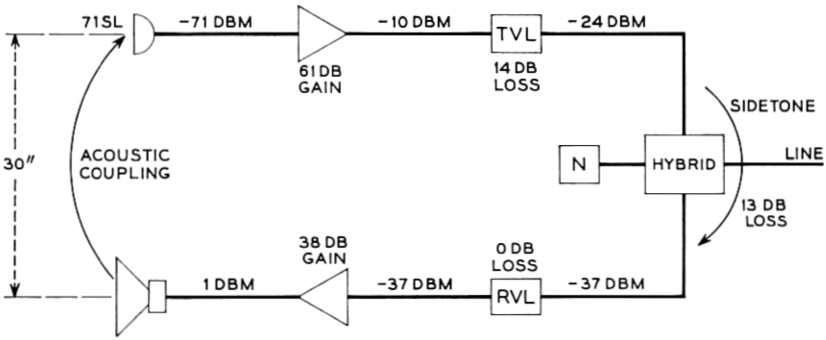


Fig. 4 — Signal levels in a typical speakerphone, illustrating the singing problem

terminals of −71 dbm. These power values define the position of the acoustic coupling line shown in Fig. 3(b). With a gain in the transmitting branch of 61 db, a gain in the receiving branch of 38 db and a hybrid loss in the coil of 13 db, it can be seen that the minimum loss to be switched is 14 db. In Fig. 4 the circuit is shown in the receive condition and the minimum loss of 14 db is inserted in the transmit variolosser. If the speakerphone switched this amount of loss, the hybrid sidetone line would lie exactly on top of the acoustic coupling line. If 20 db were switched, there would be a margin of 6 db, and the two lines would be separated by this amount, as shown in Fig. 3(b).

### 5.2 *The Transition Lines*

In Fig. 3(a) each point represents a combination of a microphone voltage and a loudspeaker voltage. For each such combination, the condition of the speakerphone, depending on the action of the control circuit, can be uniquely defined as being in transmit or receive, or it can be indeterminate, depending on the previous history of the particular combination of microphone and loudspeaker voltages. As shown in Fig. 3(c), the plane can be divided accordingly into three regions, transmit, receive and hysteresis, by two transition lines, the transmit-to-receive and the receive-to-transmit lines.

To make clear the meaning of these three regions and the two transition lines, let us take the master circuit shown in Fig. 1 and open the sidetone and the acoustic coupling paths by opening the circuit at point A in the receiving branch and acoustically isolating the microphone from the loudspeaker. We can then apply a signal at A toward the left that will put the circuit in the receive condition and generate an independent voltage $V_{LS}$ at the loudspeaker. With this voltage kept constant, the microphone voltage is increased until the circuit switches into the transmit condition. If $V_M$ is the microphone voltage at which this happens, then $V_{LS}$ and $V_M$ define one point of the receive-to-transmit (R to T) transition line. Keeping the microphone voltage constant, we can now increase the signal applied at A until the circuit switches back into the receive condition. If $V_{LS'}$ is the loudspeaker voltage at which this happens, then $V_{LS'}$ and $V_M$ define one point of the transmit-to-receive (T to R) transition line. This is illustrated in Fig. 3(c). In similar manner, the remainder of the transition lines could be constructed and the plane divided into the three regions.

For simplicity, at first, it will be assumed that there is no hysteresis region. This means that, for each point of the diagram, the control circuit uniquely defines the condition of the speakerphone. In this case, the two transition lines coincide, and the plane is divided into just two regions, transmit and receive.

The position and shape of the transition line in the diagram is completely determined by the characteristics of the circuit that controls the voice swtiching, and can be chosen by proper design. It is easy to show that, in order to avoid blocking problems, the transition line must lie in the diagram between the hybrid sidetone and the acoustic coupling lines, so that the hybrid sidetone line will be in the transmit region and the acoustic coupling line will be in the receive region, as shown in Fig. 3(d). Let us assume that this condition is not satisfied and that at least a portion of the hybrid sidetone line lies in the receive region of the diagram. This means that, for some values of the microphone voltage, the electrical sidetone through the hybrid develops a voltage at the loudspeaker large enough to put the speakerphone in the receive condition. This is the problem previously called transmit blocking due to false switching. Similarly, if the acoustic coupling line lies in the transmit region, the condition of receive blocking due to false switching develops.

As an example, in Fig. 5(a) the transition line is shown crossing the hybrid sidetone line at point A. For any microphone voltage greater than $V_M$, the loudspeaker voltage due to the sidetone path through the hybrid is large enough to put the circuit in the receive condition. The transmit region is limited at the upper end by the horizontal line through A. If the user of the speakerphone talks at a level greater than the level corresponding to the voltage $V_M$, he is "blocked out"; that is, the circuit stays in the receiving condition and his talking louder, which is the natural reaction, will not place the circuit into transmit.

A similar situation is shown in Fig. 5(b), where the transition line crosses the acoustic coupling line at B and the receive region is limited to

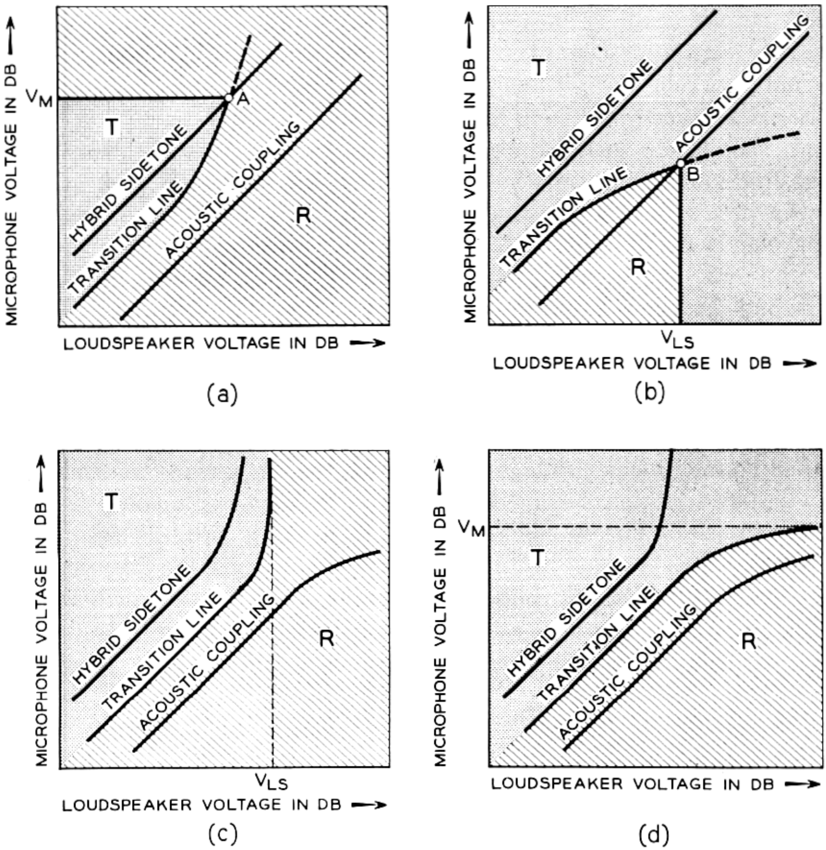


Fig. 5 — Switching diagrams showing blocking problems.

values of loudspeaker voltage smaller than $V_{LS}$. For values of loudspeaker voltage greater than $V_{LS}$, the circuit tends to go into transmit because of the acoustic coupling between loudspeaker and microphone. This condition has been referred to as receive blocking.

The requirement that the transition line lie between the sidetone and the acoustic coupling lines to avoid blocking problems is practically equivalent to the condition that the switching action be of a linear differential type; i.e., the greater the voltage present at the loudspeaker, the greater, by about the same amount, the microphone voltage necessary to switch the circuit into transmit. This is apparent in Fig. 3(d), where the transition line must be close to a 45° line in order to lie, in the range of interest, between the two other 45° lines. The sidetone and the acoustic coupling lines could be spread further apart by switching a greater amount of loss in the transmit and receive variolossers, but this is not desirable, since it tends to make all the switching problems more critical and voice switching less natural.

It should also be noted that, even though, as previously mentioned, the hybrid sidetone and the acoustic coupling lines may bend at high levels because of the effect of nonlinearity in the circuit, the transition line should not be allowed to bend sharply toward a vertical or a horizontal direction, as in Figs. 5(c) and 5(d). For instance, if the transition line had the shape shown in Fig. 5(c), there would be no transmit blocking due to false switching (because the transition line does not cross the hybrid sidetone line), but the problem of transmit blocking due to noise operation could be present. This would occur for line noise levels large enough to develop a loudspeaker voltage equal to or greater than $V_{LS}$. Then no microphone voltage could switch the circuit into transmit. An analogous situation, receive blocking due to room noise, could develop if the transition line were bent horizontally as in Fig. 5(d). In either case, the problem is present only if the bending of the transition line occurs at too low a level. Actually, horizontal bending of the transition line is very unlikely to be at a level to cause receive blocking due to room noise. But this problem may develop for other reasons, as will be shown later.

### 5.3 *Hysteresis*

Hysteresis, which we have, for simplicity, neglected so far, has important effects on the performance of the speakerphone. What is really of interest is knowing the condition of the circuit: whether it is in the transmitting, receiving or hysteresis region for each combination of the two independent variables, microphone voltage and incoming line voltage.

For this reason, it is useful at this point to make a change of coordinate in the switching diagram and use the incoming line voltage rather than the loudspeaker voltage on the abscissa. It should be pointed out that by incoming line voltage we mean the voltage due to an incoming signal and not merely the voltage appearing across the line terminals, since this could be the voltage due to an outgoing signal from the microphone output or a combination of an incoming signal and an outgoing signal. To make this change in coordinate, we translate the voltage at the loudspeaker into a line voltage by applying to it a correction that takes into consideration, besides the change in impedance, the gain in the receiving branch and the losses in the receive variolosser and the hybrid coil. In this way, a sidetone voltage is plotted in terms of an equivalent incoming line voltage that would give the same loudspeaker voltage.

With these assumptions, the sidetone and the acoustic coupling lines can be drawn on the new diagram. Starting with the acoustic coupling line, it now becomes necessary to define the condition of the circuit — whether it is in transmit or in receive — since the receive variolosser is in the path between the incoming line voltage and the voltage developed at the microphone. As shown in Fig. 6(a), the acoustic coupling line shifts by the amount of loss switched in the receive variolosser upon going from the receive to the transmit condition. Since linearity is assumed, the lines are straight at a 45° angle, and the shift can be considered to be either vertical or horizontal. When nonlinearities are present, as exempli-

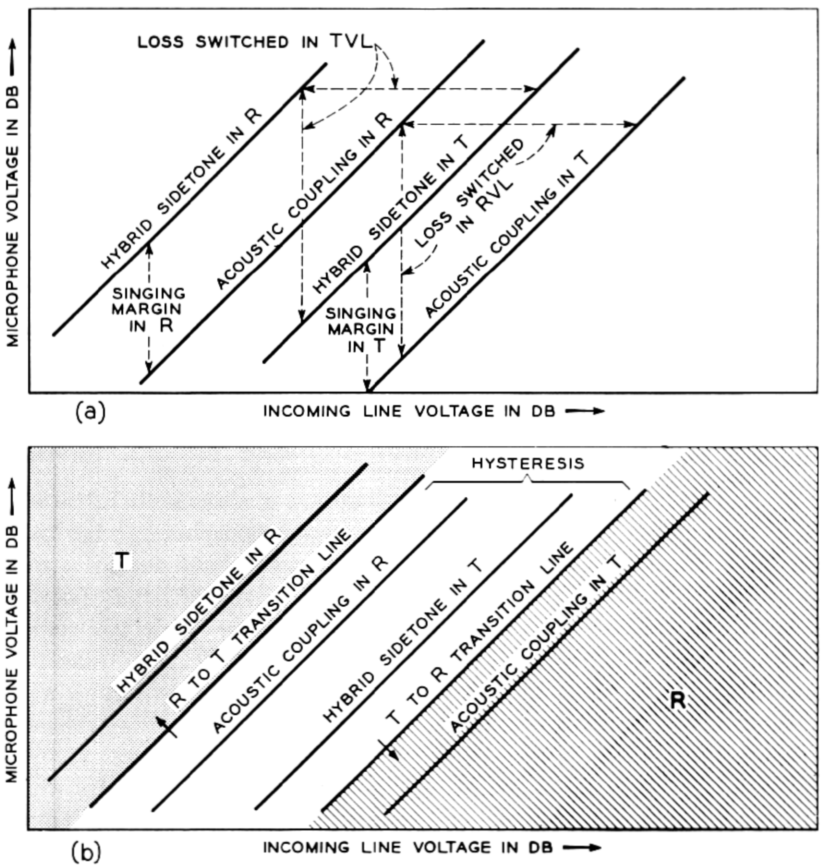


Fig. 6 — Switching diagram with hysteresis.

fied by the dotted line in Fig. 3(a), the shift should be considered vertical if they occur in the circuit before the receive variolosser, and horizontal if they occur after the receive variolosser. Similarly, the hybrid sidetone line has two positions, which are shifted by the amount of loss switched in the transmit variolosser, with similar rules determining whether the shift should be considered horizontal or vertical.

In view of this, the two sets of lines in Fig. 6(a) correspond to the transmit and receive conditions. The line for hybrid sidetone in the transmit condition has been drawn below the line for acoustic coupling in the receive condition. This means that, when no loss is present in either the transmit or the receive variolosser, the circuit is not stable and will sing. If this were not true, there would be no reason to introduce voice switching. It follows that no single transition line can be drawn that will divide the plane into two parts satisfying the requirement that the hybrid sidetone lines lie in the transmit region and the acoustic coupling lines lie in the receive region. It is necessary, therefore, to draw two transition lines, one to define the transmit-to-receive transition, the other to define the receive-to-transmit transition. The first will hold for the circuit in the transmit condition and the second for the circuit in the receive condition. These are shown in Fig. 6(b).

It is evident from the figure that the two transition lines are necessarily distinct and, between them, define a hysteresis region that must have a width approximately equal to the average amount of loss switched. It does not matter how the transition line is shifted in the diagram — from the transmit to the receive condition horizontally, vertically, or both — as long as we deal with the linear part of it.

## VI. SOME GENERAL CONSIDERATIONS

As an application of the switching diagram, it may be useful to consider, before discussing the merits of each submaster circuit, what other general objectives one should strive for in speakerphone design.

We have mentioned that, to reduce switching problems, it is desirable to switch as little loss as possible, consistent with the stability of the circuit. From Fig. 6(a) it is also apparent that, to maintain the same singing margin in transmitting and in receiving, the amount of loss switched in the transmit variolosser should be about the same as that switched in the receive variolosser.

We have also pointed out that, within reasonable limits, the transition line should be a 45° straight line. At high levels, however, the shape of the sidetone, acoustic coupling and transition lines will deviate from this straight line. As long as the levels at which this happens are high compared with those of normal speech, the shape of these curves becomes immaterial. At low levels, on the other hand, it may be advantageous to have the transition lines deviate from the linear characteristic in a predetermined fashion. It has been recognized[5] that, when a voice-operated device selecting one of two directions of transmission on the basis of signal amplitude is located at a point where the signal-to-noise ratio coming from one direction is poorer than that coming from the opposite direction, it is advantageous to use an arrangement whereby the direction having the better signal-to-noise ratio is normally blocked and the direction having the poorer ratio is normally activated. In the case of a speakerphone, the incoming signal-to-noise ratio cannot be controlled and, therefore, can be poorer than the one from the near end. Thus, it is reasonable to design the control circuit so that, at low outgoing levels, the speakerphone is normally in the receive condition. This is equivalent to setting a threshold in the microphone voltage below which the circuit will only be in the receive condition. The transition line will then have the shape shown in Fig. 7(a). It is true that this shape implies transmit blocking for microphone voltages below the threshold, but, as previously mentioned, this may be desirable at very low levels.

When the shape of the transition line is like that of the one shown in Fig. 7(a), the manner in which the line is shifted in going from the transmit to the receive condition and vice versa becomes important. Fig. 7(b) shows a completely horizontal shift, Fig. 7(c) a completely vertical shift and Fig. 7(d) a partly vertical and partly horizontal shift. In the first case, the transmitting threshold remains the same whether the circuit is in transmit or receive. The choice of the threshold level is rather critical in this case, since too low a level will tend to make receiving "choppy", especially in a noisy room, while too high a level will cause clipping in transmitting. With a completely vertical shift [Fig. 7(c)] the threshold drops considerably when the circuit switches into transmit. It is possible, therefore, to choose the threshold level so that it is fairly low in the transmit condition and high in the receive condition, resulting in a very stable circuit, both in transmit and in receive. In this case, however, it might be difficult to switch from transmit into receive, since even a relatively low room noise might keep the circuit in transmit (receive blocking due to noise operation). For these reasons it may be desirable to choose a compromise such as that indicated in Fig. 7(d).

Another feature worth consideration is the volume control. For a speakerphone to work satisfactorily under widely varying room and line conditions, and also to suit different users' needs, it is necessary to have a volume control in the receiving branch. Adjusting the gain of the

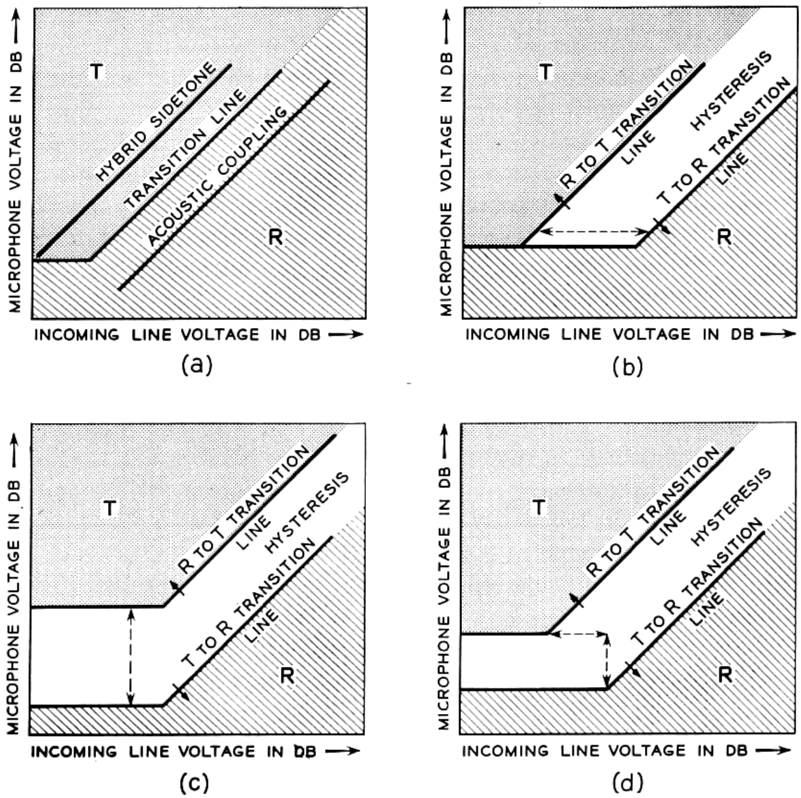


Fig. 7 — Switching diagram, transition lines with threshold.

receiving amplifier or inserting a loss in the receiving branch has the effect of shifting the acoustic coupling lines in the switching diagram without moving the hybrid sidetone lines. Since the circuit must be stable even with maximum gain setting, it follows that, as the volume control is turned down, the acoustic coupling lines move away from the hybrid sidetone lines, as shown in Fig. 8(a), thus unnecessarily increasing the amount of singing margins. Keeping in mind the objective of switching only the amount of loss necessary to make the circuit stable, it is desirable to arrange the volume control in such a way that, when turned down, it will reduce by the same amount the receiving level and the loss switched, in both the transmit and the receive variolossers.[6] This arrangement has the effect [shown in Fig. 8(b)] of shifting the sidetone and acoustic coupling lines for the receive condition toward the corresponding lines for the transmit condition. The lines for the transmit condition

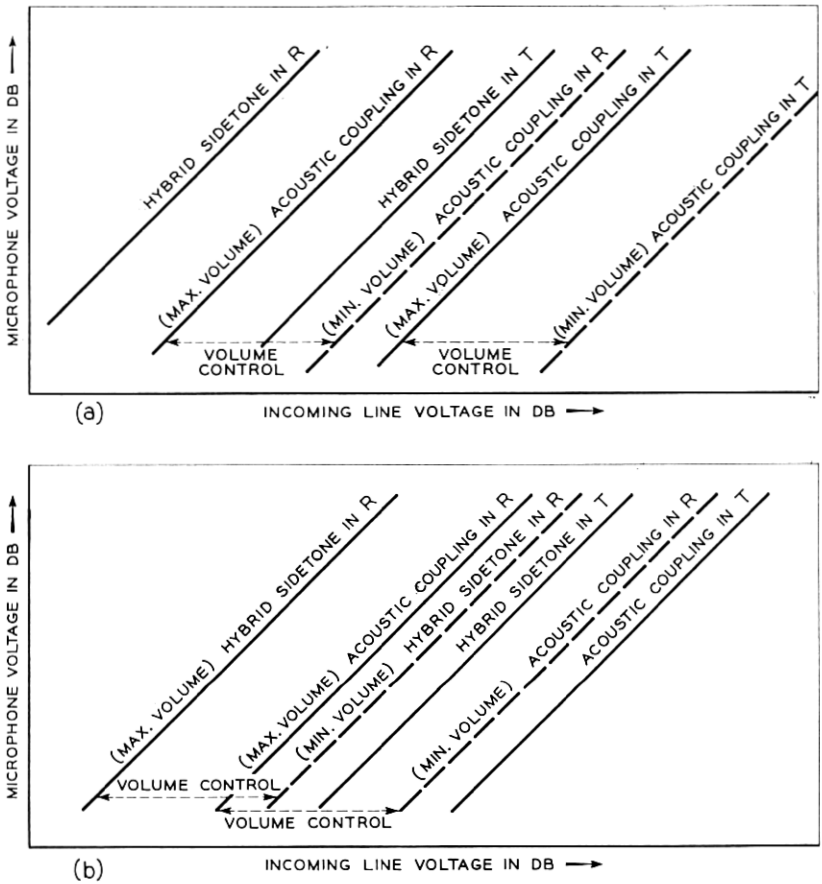


Fig. 8 — Switching diagram, two types of volume control: (a) keeping the amount of switched loss constant; (b) varying the amount of switched loss.

remain stationary and the margins stay constant. The control circuit, of course, must be so designed that, as the volume control is changed, the transition lines will always maintain the proper position with respect to the sidetone and acoustic coupling lines in the diagram.

## VII. FOUR SUBMASTER CIRCUITS

### 7.1 *Submaster #1*

A block diagram for this type of speakerphone is shown in Fig. 2(a), with the inputs to its control circuit being $V_{T1}$ and $V_{R1}$. Let us assume that the control circuit by itself has a simple linear characteristic with a threshold such as the one shown in Fig. 9(a). The circuit is acted upon by the two input voltages and, depending upon their magnitudes, will be in either the transmit or the receive condition. Since the voltages $V_{T1}$ and $V_{R1}$ are directly proportional to the microphone voltage and to the incoming line voltage, there will be a single transition line in the switching diagram identical to the characteristic of the control circuit with no hysteresis.

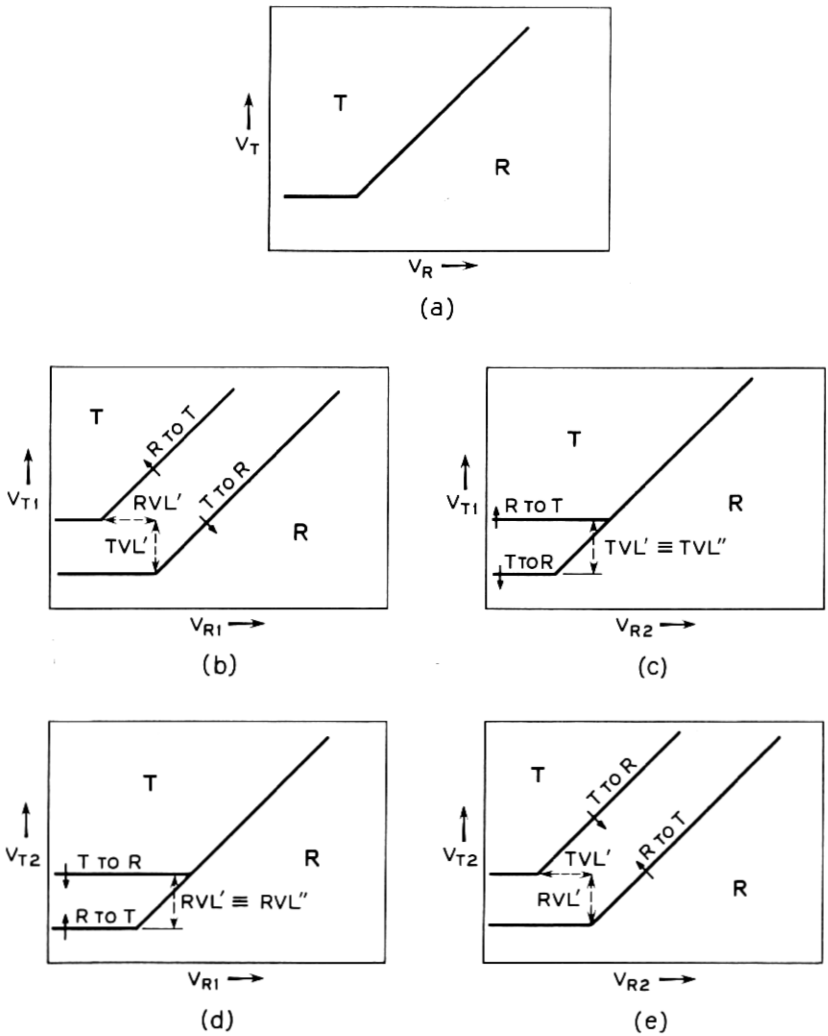


Fig. 9 — Control circuit characteristics.

From the previous discussions, it is clear that the lack of hysteresis gives blocking problems. Therefore, the characteristic of the control circuit must be somewhat more sophisticated and display some hysteresis, such as in Fig. 9(b). One way of obtaining this characteristic and of easily controlling the horizontal and vertical shift of the transition line is shown schematically in Fig. 10(a). Two extra variolossers are employed, TVL′ and RVL′, one for each of the two inputs of the control circuit. The loss of TVL′ controls the amount of vertical shift of the transition line, and it is varied by the control circuit in the same manner as the loss in the transmit variolosser. The loss in RVL′ controls the amount of horizontal shift of the transition line, and it is varied in the same manner as the loss in the receive variolosser. The sum of the losses switched in TVL′ and RVL′ should approximately equal the average loss switched in the transmit and receive variolossers. This type of circuit has been used, for example, in a British speakerphone by Winston Electronics Ltd.[7] It requires extra components such as the two variolossers shown but results in a very flexible design.

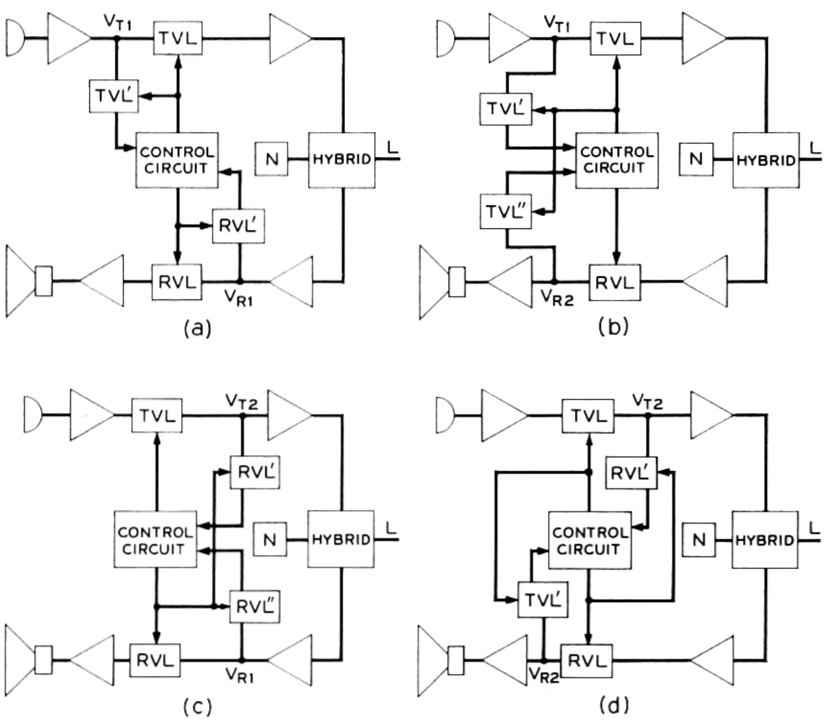


Fig. 10 — Some refinements of the submaster.

### 7.2 *Submaster #2*

A block diagram of this type of speakerphone, with $V_{T1}$ and $V_{R2}$ as inputs to the control circuit, was shown in Fig. 2(b). It has the advantage that even a control circuit with the simple characteristic of Fig. 9(a), without built-in hysteresis, will give satisfactory performance. The reason is that the receive variolosser is located between the incoming line voltage and the control circuit input voltage $V_{R2}$. For a given microphone voltage, the necessary line voltage required to switch the circuit from the transmit to the receive condition is larger than that required to hold it in the receive condition by the amount of loss inserted by the receive variolosser. This is illustrated in Fig. 7(b), where the horizontal shift in decibels can be taken to be exactly equal to the loss of the receive variolosser.

However, to obtain the performance associated with a partially vertical and partially horizontal shift of the transition line like the one in Fig. 7(d), the control circuit must be more complicated and have a characteristic of the type shown in Fig. 9(c). The threshold must be effectively lowered when the circuit is in the transmit condition; a simple arrangement that will boost the gain of the control circuit when transmitting will give the desired characteristic. One way to obtain it is shown in Fig. 10(b), where TVL′ and TVL″ are two extra variolossers, one in each of the two inputs to the control circuit. Both these variolossers should be switched by the control circuit at the same time and in the same direction as the transmit variolosser. The amount of loss switched is about the same in the two variolossers and determines the vertical shift of the transition line.

### 7.3 *Submaster #3*

The inputs to the control circuit for this submaster are $V_{T2}$ and $V_{R1}$, as shown in the block diagram of Fig. 2(c). This type of speakerphone could be considered analogous to the submaster #2 in the sense that a control circuit with a simple characteristic like the one shown in Fig. 9(a) will also result in two transition lines shifted by the right amount in the switching diagram. However, in this case the shift is vertical, as in Fig. 7(c), and equals the amount of loss switched in the transmit variolosser. Receive blocking due to noise operation is likely to occur. As in submaster #2, a better performance could be obtained with a more complicated control circuit characteristic. In this case the threshold should be raised when the circuit is in the transmit condition, as in Fig. 9(d). To accomplish this purpose, Fig. 10(c) shows an arrangement with two extra variolossers, RVL′ and RVL″, which should switch an equal amount of loss. They are operated at the same time and in the same direction as the receive variolosser.

### 7.4 *Submaster #4*

In the interest of completeness, the block diagram of this submaster is shown in Fig. 2(d), with the inputs to the control circuit being $V_{T2}$ and $V_{R2}$. A simple characteristic like the one shown in Fig. 9(a) will give two transition lines in the switching diagram, shifted vertically and horizontally by the amount of loss switched in the transmit and in the receive variolossers. This represents too large an amount of hysteresis. Therefore, as in the submaster #1, the circuit will have blocking problems unless the control circuit characteristic is more complicated. The desired characteristic is shown in Fig. 9(e) and is analogous to that shown in Fig. 9(b) for the submaster #1. Again, a way of obtaining the desired characteristic is to add two extra variolossers, as shown in Fig. 10(d).

## VIII. TRANSIENTS

### 8.1 *Definition of Time Constants*

Since transients concern the time intervals during which the signals are rapidly varying, the time constants associated with the action of the control circuit are very important factors in determining the transient performance. As previously pointed out, two inputs are necessary to determine the action of the control circuit, a voltage from the transmitting branch and a voltage from the receiving branch. It is convenient to define, for each input, two time constants related to the interval necessary for a variation in the input voltage to be detected by the control circuit. Since the control circuit is made sensitive to some weighted average of the input voltages, this process of averaging generally involves a delay. The delay associated with the detection of an increase in the input voltage will be called *build-up time* and will, in general, be different from the delay associated with the detection of a decrease in the input voltage, which will be called *decay time*. There will be, therefore, a build-up and a decay time, $B_T$ and $D_T$, associated with the transmit input and analogous time constants, $B_R$ and $D_R$, associated with the receive input. This is shown graphically in Fig. 11, where, for simplicity, the input voltage is made to vary as a step function and the build-up and decay times are illustrated as the time constants of an exponential curve representing the voltage detected in the control circuit.

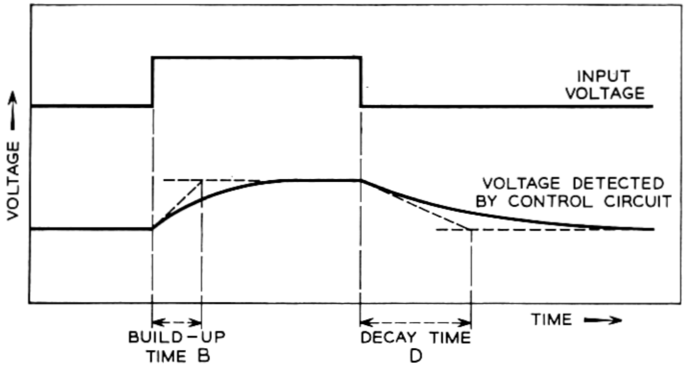


Fig. 11 — Illustration of time constants.

Two more time constants can be defined that correspond to the actual time it takes for the control circuit to switch into transmit or into receive after having detected the two inputs. They will be called $A$ and $R$, the attack and release times of the switch.

It will be assumed in the following that the defined constants are independent, and that the speakerphone is stable in the receive condition when no signals are applied. The latter assumption is in agreement with previous statements about the desirability of having a threshold in the transition line of the switching diagram, as shown in Fig. 7. The delay due to the sound travel time from loudspeaker to microphone and room reverberation will be considered later.

### 8.2 *Relations Among the Time Constants*

Some general observations can be made about the effect of the six time constants on the performance of the speakerphone. When the microphone and the incoming line voltages are varying rapidly, the voltages being detected by the control circuit are affected by the build-up and decay times previously mentioned. Therefore, at each instant one can define effective microphone and line voltages that, under steady-state conditions, would cause the same voltages to be detected by the control circuit as are detected during these transients. The effective voltages are not uniquely defined, since their difference from the actual values depends on how fast the latter ones are varying and on the previous history.

Fig. 12(a) shows qualitatively the shape assumed in some typical cases by the sidetone and acoustic coupling lines when they are plotted in terms of the effective microphone and line voltages. As an example,

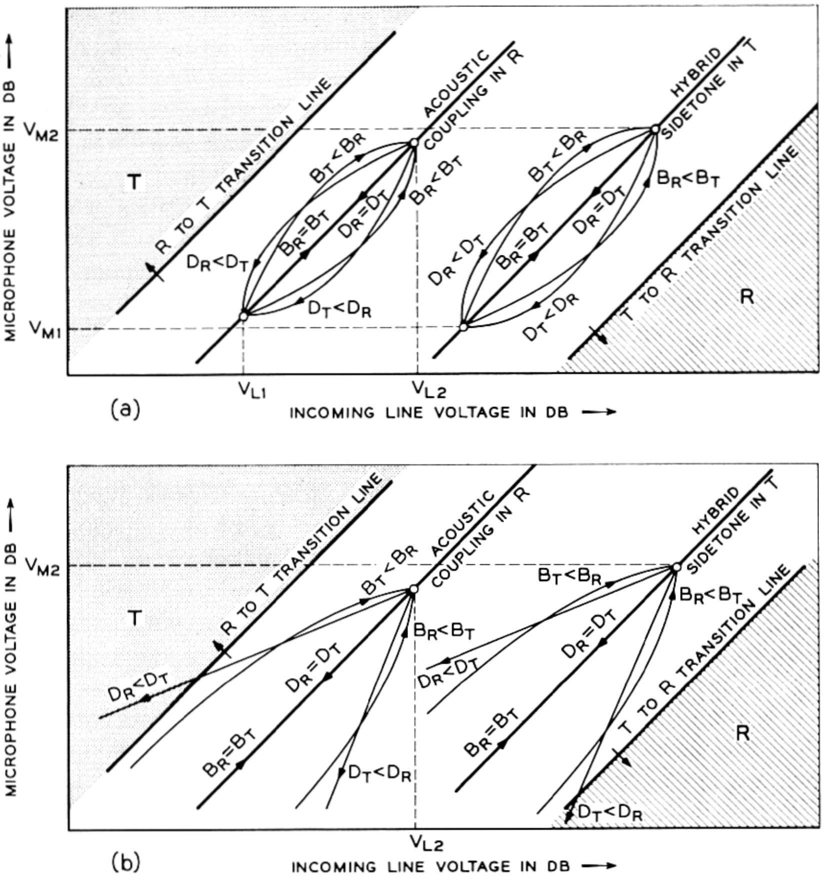


Fig. 12 — Effect of time constants during transients.

let us assume that the circuit is in the receive condition and that an incoming line voltage varies between the values $V_{L1}$ and $V_{L2}$. Under steady-state or quasi-steady-state conditions, the operating point of the speakerphone on the switching diagram would move along the straight acoustic coupling line and the circuit would remain in the receive condition all the time. If, however, the line voltage increased very rapidly from $V_{L1}$ to $V_{L2}$, the effective line voltage would follow the variation, with a certain delay determined by the build-up time of the receive input, and would not reach $V_{L2}$ until some time later. Also, the effective microphone voltage, which represents the acoustic coupling, would follow the increase with a certain delay, determined in this case by the build-up time of the transmit input. As a consequence, the operating point of the circuit on the diagram would trace an acoustic coupling curve whose shape was, in general, a function of the two build-up time constants. In particular, the operating point would follow a curve that lay above the steady-state acoustic coupling line when $B_T$ was smaller than $B_R$ and lay below the line when $B_R$ was smaller than $B_T$, as shown in Fig. 12(a). Only when the two build-up times were identical would the operating point during the transient move in a linear path along the acoustic coupling line, as in the steady-state condition. Similar reasoning would apply when the line voltage decreased rapidly from $V_{L2}$ to $V_{L1}$. The acoustic coupling line would assume new configurations that are a function, in this case, of the decay time constants $D_T$ and $D_R$. These are also shown in Fig. 12(a).

Perfectly analogous is the case of the transmit condition, when the microphone voltage varies rapidly between $V_{M1}$ and $V_{M2}$. The hybrid sidetone line will assume different shapes, which are again functions of the same four time constants, as shown in Fig. 12(a). By making the voltage $V_{L1}$ or $V_{M1}$ approach zero ($-\infty$ on a logarithmic scale) one can see the shape assumed by the acoustic coupling and the hybrid sidetone lines when a line voltage or a microphone voltage is applied or removed suddenly. This is shown in Fig. 12(b) for different combinations of build-up and decay time constants.

From all these curves it is apparent that, unless $B_T = B_R$ and $D_T = D_R$, the margins against false operation of the control circuit will be reduced during the transient either in receiving or in transmitting. For instance, if $B_T$ is smaller than $B_R$, or if $D_R$ is smaller than $D_T$, the acoustic coupling line will shift upward closer to the transition line; and, vice versa, if $B_R$ is smaller than $B_T$, or if $D_T$ is smaller than $D_R$, the hybrid sidetone line will shift downward toward the transition line. If the transition line is crossed during a transient, clipping of the signal is likely to occur. Therefore, it is desirable, in general, to satisfy the relations $B_T = B_R$ and $D_T = D_R$.

However, it should be noted that these relations are not very critical. First of all, even with a ratio of 1:2 in the build-up or the decay times and a sudden variation of 20 db in one of the inputs, the maximum departure of the operating point during the transient from the steady-state line would only be about 6.5 db. A smaller or slower variation would tend to reduce this amount. Secondly, even though the transition line might be crossed by the acoustic coupling or the sidetone lines, if this happens for only a very short interval the control circuit will not have time to switch, due to the finite attack and release times, $A$ and $R$, of the switch. Strictly from this point of view, long attack and release times help prevent clipping.

Other conditions to be imposed on the time constants can be deduced by considering a switching operation. As pointed out at the beginning of Section IV, it is desirable to have the circuit switch into transmit as fast as possible to avoid initial clipping. This means that the build-up time of the transmit input and the attack time of the switch should be very short. Experience shows that, with a total switching time of 10 milliseconds, clipping is hardly noticeable, and there is little to be gained by making this time shorter. How to divide this time between $B_T$ and $A$ does not seem to be critical as long as their sum is short. Since the circuit should remain for some time in the transmitting condition to avoid final clipping of the transmitted speech, it is desirable to make the sum of the decay time of the transmit input and the release time of the switch rather long, but not long enough to impair the break-in ability of the distant party. Experience has shown that a good compromise for the sum of these two time constants $D_T$ and $R$ is around 150 milliseconds. This value and the way it is divided between the two constants are not critical.

### 8.3 *Effects of the Acoustics of the Room*

The time it takes for sound to travel from the loudspeaker to the microphone (a few milliseconds) adds to the values of the build-up and decay time of the transmit input only when we are considering the acoustic coupling during a transient. According to the criteria previously outlined and illustrated in Fig. 12, an increase in $B_T$ actually improves the margin against clipping while the circuit is in the receiving condition. An increase in $D_T$ would tend to decrease the margin, but the few milliseconds added are generally negligible compared with the value of $D_T$. Since the hybrid sidetone line or the other time constants are not affected, there is hardly any problem introduced by this effect.

The reverberation in a room, on the other hand, is the cause of the "echo" problem previously described. It has the effect of making the microphone voltage due to acoustic coupling decay at a slow rate after an incoming line signal has ceased suddenly. The reverberation time constant, $R_V$, therefore combines with $D_T$ in determining the shape of the acoustic coupling line. If, for simplicity, we at first neglect $D_T$ compared with $R_V$, the acoustic coupling line assumes one of the forms shown in Fig. 13. These are similar to the lines shown in Fig. 12(b) for different ratios between $D_T$ and $D_R$. Point A is the operating point for a steady line voltage $V_L$. As this voltage is suddenly removed, the microphone

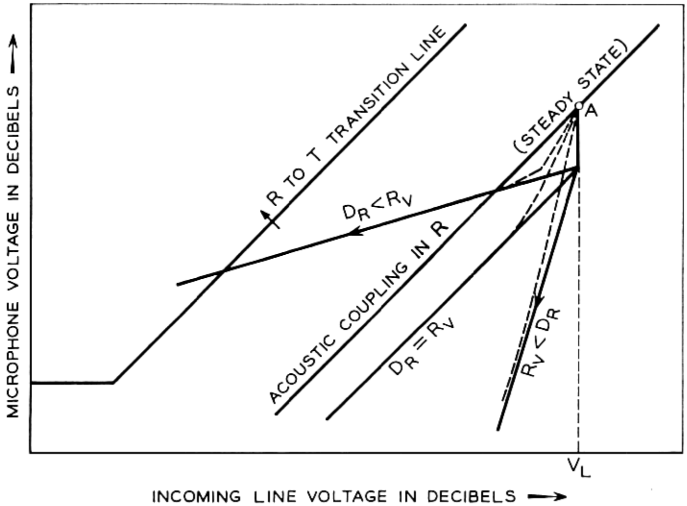


Fig. 13 — Effect of room reverberation.

voltage shows a fast drop, due to the abrupt interruption (within a few milliseconds) of the direct energy coming from the loudspeaker. Then the operating point moves along a straight line whose slope is a function of the ratio between $R_V$ and $D_R$. The larger this ratio, the greater is the chance for the acoustic coupling line to cross the transition line and cause "echo". Therefore, to avoid this problem $D_R$ should be designed to be about equal to $R_V$.

If $D_T$ is not neglected, the curves on Fig. 13 will be slightly different and will be determined by, among other factors, the way the circuit combines the two time constants $D_T$ and $R_V$. In general, the initial drop will not be so abrupt, and the shape will be more like that of the dashed lines. The same conclusions will still be valid.

The reverberation time may vary over quite a range, depending on the size and acoustic treatment of the room, and it may be necessary to adjust the value of $D_R$ at each installation. Experience has shown, however, that, by choosing a fairly long $D_R$, one can take care of the echo problem in most of the rooms encountered in practice without further adjustment. Of course, $D_T$ must also be made relatively long to avoid final clipping of the transmitted signal in a room with short reverberation time. This follows from previous considerations, pointing to the desirability of making the two decay times $D_R$ and $D_T$ about equal.

The general relations derived for the time constants in these sections can be summarized as follows: the build-up and decay times of the transmit input must be approximately equal respectively to the build-up and decay times of the receive input. The sum of the build-up time and the attack time of the switch must be relatively short (about 10 milliseconds), while the sum of the decay time and the release time must be relatively long (about 150 milliseconds). The decay time must be long enough in each case to take care of room reverberation.

### 8.4 *Other Considerations Affecting the Transients*

Important characteristics that affect the performance during the transients are the loss–time curves of the two variolossers. These are the curves that give the variation of the loss with time during the switching interval. To insure proper margin against singing at any instant, the sum of the two losses introduced by the transmit and receive variolossers must be at least equal to the total loss in the steady-state condition. This means that, if the two loss–time curves are plotted on a diagram like the one of Fig. 14(a), the curve for the transmit variolosser TVL must lie at least slightly above the one for the receive variolosser RVL, as indicated by the solid and dashed lines.

That this condition be satisfied is particularly important when the volume control is of the type described in Section VI, which changes the receiving level and the loss switched in TVL and RVL at the same time and by the same amount. A convenient way of arranging this is to have the volume control operate directly on the variolossers, shifting their steady-state receiving condition as shown in Fig. 14(b). As the volume is turned down, a fixed (nonswitchable) loss is introduced in RVL in the receive condition, thus reducing the receiving level by this amount. At the same time, an approximately equal amount of fixed loss is taken out of TVL. The requirements of this volume control arrangement are met if the two loss–time curves satisfy the condition of minimum singing margin at all times.

Besides the relative position of these two curves, their shape is also of interest, because it has a bearing on the unnatural effects of voice switching. To minimize these effects it is desirable to make the rate at which the loss is switched vary in inverse fashion with the amount of energy being changed. For example, in the transmitting variolosser the rate of loss removal in going from the receive into the transmit condition can be rapid at the start of the switching interval, but it should be slower in the final stages when the signals put out on the line are greater. Similarly, in going from the transmit into the receive condition, the restoration of loss should be at a slow rate in the initial stages so that back-

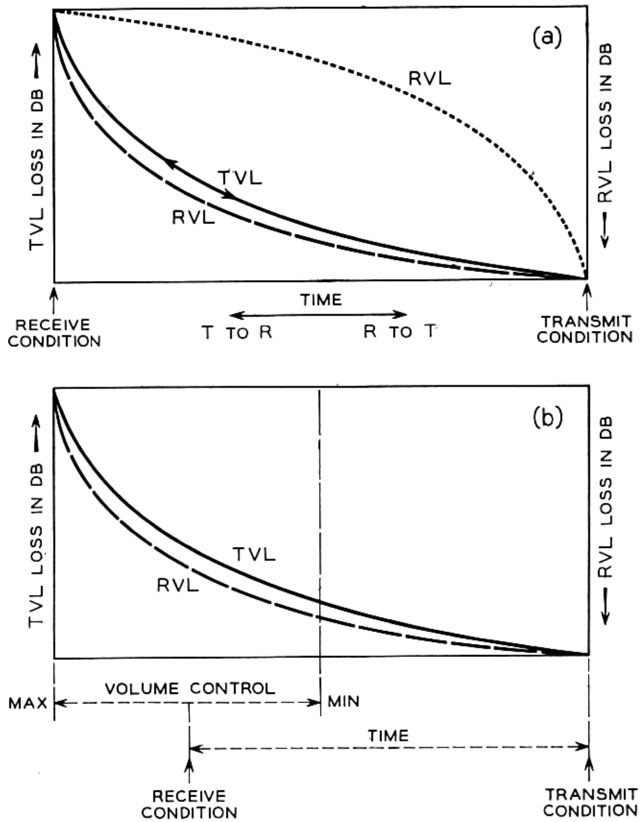


Fig. 14 — Loss–time characteristics of variolossers.

ground noise is not chopped off abruptly after each talk spurt, but a faster rate could be used in the final stages. The resulting shape is the one shown in Fig. 14(a) for the transmit variolosser.

If the same criteria were applied to the receive variolosser, the RVL loss-time curve should have the shape of the dotted line in Fig. 14(a). It is obvious that this RVL curve is not compatible with the chosen TVL curve if the singing margin must be maintained at all times. However, it is more important to minimize the unnatural effects of voice switching in transmitting than in receiving, since the circuit tends to remain stable in receiving when no signals are applied. On this basis, the loss–time curves for the transmit and receive variolossers should assume the shapes shown in Fig. 14(b).

## IX. ANALYSIS OF SOME SPECIAL FEATURES

A problem commonly encountered in the use of speakerphones is sensitivity to local noises, whether they are sharp loud noises like the slamming of a door or the steady type like that of a fan. In the case of sharp noises, if $B_T$ and $A$ are short and $D_T$ and $R$ are long, as was shown to be desirable, the noise will switch the circuit into transmit and keep it there for some interval, interfering with the reception of the incoming signal. One way to reduce this interference is to design the control circuit so that $D_T$ and $R$ are short for sounds of very short duration and have the normal long value only when a sound persists for a period equal to the shortest duration of speech sounds. This "deferred hangover" principle has been successfully used in TASI.[8]

With steady noise, the problem is to keep the circuit from switching into transmit and causing receive blocking. A simple solution consists in raising the threshold of the transition line. However, in order not to impair the performance of the speakerphone in a less noisy environment, the threshold should be adjusted in each case to be just above the noise level. There are ingenious circuits that make this adjustment automatically. In order to operate properly, they must discriminate between speech and steady noise, so that the threshold will not be affected by the former but will be raised by the latter. The fluctuating characteristic of speech energy compared with noise is used for this discrimination.

## X. SUMMARY

In the design of a speakerphone circuit and in the evaluation of its performance from both the steady-state and the transient viewpoints, the switching diagram described in this paper has proved a useful tool. The fundamental problems of voice switching have been classified, and criteria for avoidance of them have been given in terms of the diagram characteristics and the circuit time constants.

Four submasters representing general configurations of a speakerphone with only the essential features, have been examined. Each of them can be made to give comparable performance, and the choice in general will depend on circuit design considerations, mainly on which type results in the simplest and most economic design for a given application.

At the cost of some increase in complexity, features can be added to the submasters to give improved performance in a particular characteristic. Some of the ingenious features that have been used or previously suggested have been examined. In general, these can be analyzed and evaluated by means of the switching diagram.

REFERENCES

1. Clemency, W. F., Romanow, F. F. and Rose, A. F., The Bell System Speakerphone, Comm. & Electronics, No. 30, May 1957, p. 148.
2. Emling, J. W., General Aspects of Hands-Free Telephony, Comm. and Electronics, No. 30, May 1957, p. 201.
3. Wright, S. B. and Mitchell, D., Two-Way Radio Telephone Circuits, B.S.T.J., **11,** July 1932, p. 368.
4. Hopkins, H. F. and Stryker, N. R., A Loudness Rating for Loudspeakers and Power Requirements for Rooms, Proc. I.R.E., **36,** March 1948, p. 315.
5. Wright, S. B., The Vodas, B.S.T.J., **16,** October 1937, p. 456.
6. Tillman, R. J., U.S. Patent No. 2,269,565.
7. Laurence, R. F., A Voice-Switched Loudspeaking Telephone for Public Network Lines, British Comm. and Electronics, February 1957, **4,** p. 91.
8. Bullington, K. and Fraser, J. M., Engineering Aspects of TASI, B.S.T.J., **38,** May 1959, p. 353.