THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOLUME 18, NUMBER 1 JULY, 1946

The Sound Spectrograph*

W. KOENIG, H. K. DUNN, AND L. Y. LACY
Bell Telephone Laboratories, Inc., New York, New York

(Received May 4, 1946)

The sound spectrograph is a wave analyzer which produces a permanent visual record showing the distribution of energy in both frequency and time. This paper describes the operation of this device, and shows the mechanical arrangements and the electrical circuits in a particular model. Some of the problems encountered in this type of analysis are discussed, particularly those arising from the necessity for handling and portraying a wide range of component levels in a complex wave such as speech. Spectrograms are shown for a wide variety of sounds, including voice sounds, animal and bird sounds, music, frequency modulations, and miscellaneous familiar sounds.

I. GENERAL

IN many fields of research it is necessary to analyze complex waves. If these waves are steady in time, the analysis presents no particular difficulties. If, however, the wave is complex in its frequency composition and also varies rapidly in time, the problem is very difficult. Numerous methods have been employed in the past to try to show changing energy-frequency distribution; several examples have appeared in the pages of this journal. Figure 1, for instance, shows a series of harmonic analyses of the successive periods of a vowel sound.[1] The dotted lines mark the regions of resonance which change continuously throughout the production of the sound. By performing this operation on a whole sentence, an effort was made, as shown in Fig. 2 taken from the same paper, to represent the time variations in the energy-frequency distributions. Here the frequencies of the various resonant regions are represented by the solid lines and their relative amplitudes are roughly indicated by the widths of the lines. The generation of this graph represented a formidable amount of time and labor.

Figure 3 shows another representation of a changing wave form.[2] This is an oscillogram of a series of 11 steady tones sent over a radio channel and received through a bank of narrow band filters whose outputs were commutated at the rate of $12\frac{1}{2}$ times per second. A slight gap was left between cycles, as marked at the bottom of Fig. 3. The successive cycles show varying profiles, due to the fact that the frequency response of the radio channel was continually changed by selective fading. Similar pictures would have been obtained if a varying signal such as speech had been impressed on the bank of filters without

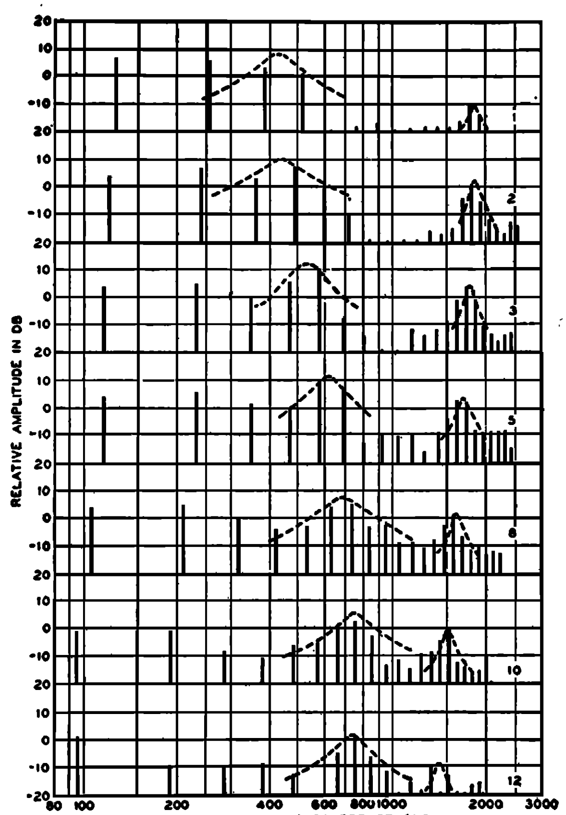

FIG. 1. Illustrating one method which has been used in the past to show how vocal resonances change with time. This is a series of harmonic analyses of successive periods of the vowel in the word "out."

* Paper presented before the meeting of the Acoustical Society of America, May 10, 1946, New York, New York.
[1] J. C. Steinberg, J. Acous. Soc. Am. 6, 16–24 (1934).
[2] R. K. Potter, Proc. I.R.E. 18, 581–648 (1930).

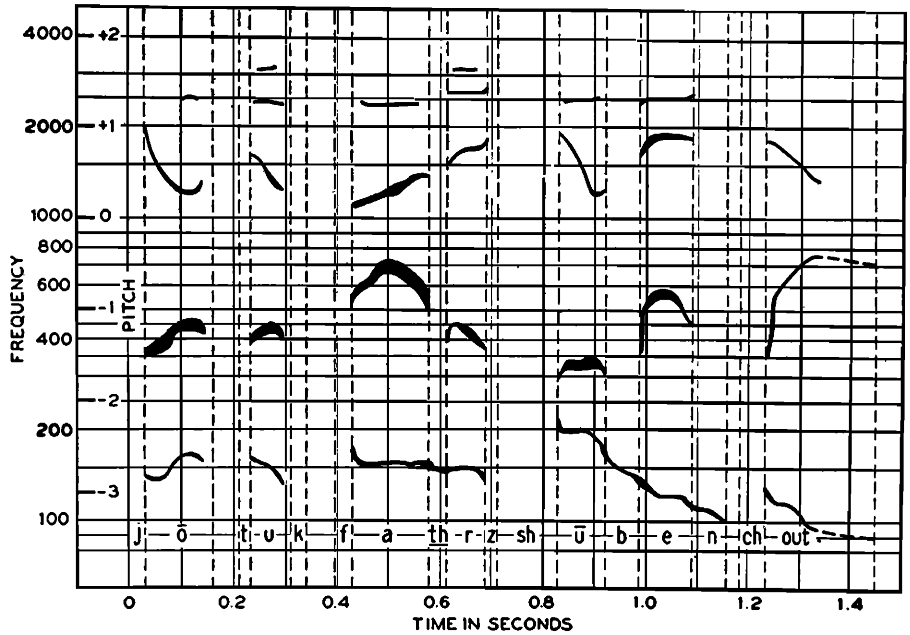

FIG. 2. A plot of the time variations of the vocal resonances in a short sentence, compiled from a series of analyses like those in Fig. 1.

the radio link. Instead of making oscillograms, however, the output was at that time displayed on a cathode-ray tube. The changing profiles of these patterns portrayed the changing energy frequency distribution in speech. An attempt was made to learn to recognize these word patterns, but with little success.

Figure 4 shows another device for portraying a complex wave.[3] Here a kind of three-dimensional model was developed by analyzing the amplitude of each harmonic component of a piano note as a function of time, plotting the results on cards, and cutting out the profiles.

Still another method is illustrated in Fig. 5. Here the frequency range was divided into ten bands[4] by means of band filters. The output of each band was rectified and recorded with a string oscillograph, so that each oscillogram shows the variation of amplitude with time. Despite the rather small number of bands, a kind of speech pattern can be discerned in this array of oscillograms.

Figure 6 shows this process carried further. These are solid models built up of oscillograms of about 200 overlapping frequency bands, cut out in profile and stacked side by side. In the upper model only the high peaks in the speech are prominent. In the lower model the level differences among the various regions have been equalized by electrical compression which will be explained subsequently. Of particular interest in these models is the sharpness of the wave front which appears at the beginning of some of the words. It can be seen from these models that in speech the energy-frequency distribution is very complex and changes form rapidly with time.

The production of solid models, while useful for particular purposes, is hardly a practical method for everyday needs. Furthermore, it is difficult to portray the results usefully in a two-dimensional picture. If, however, we substitute for the third dimension in these models a system of varying shades of gray or black with the highest amplitudes represented by dark areas and the lowest amplitudes by light areas on a flat surface, then we have a method which can be rapid and convenient. This is the method of the sound spectrograph.

[3] O. H. Schuck and R. W. Young, J. Acous. Soc. Am. **15**, 1–11 (1943).
[4] Homer Dudley, J. Acous. Soc. Am. **11**, 169–177 (1939).

II. GENERAL PLAN AND FIRST MODEL OF THE SOUND SPECTROGRAPH

Figure 7 shows in highly schematic fashion the basic method of the sound spectrograph, as originally proposed by Mr. R. K. Potter.[5] It is necessary, first, to have a means of recording the sound in such a form that it can be reproduced over and over. The means shown here is a magnetic tape, mounted on a rotating disk. In recording, some predistortion of the signal may be desirable and is therefore indicated in connection with the recording amplifier. With speech, for example, it has been found advantageous to raise the amplitude of the higher frequencies by about 6 db per octave in order to equalize the representation of the different energy regions.

Second, a means of analyzing must be provided. Most convenient is the heterodyne type of analyzer employing a fixed band pass filter, with a variable oscillator and modulator system by which any portion of the sound spectrum can be brought within the frequency range of the filter.

Finally, the output of the analyzer must be recorded in synchronism with the reproduced sound. The simplest method is by means of a drum, on the same shaft with the magnetic tape, carrying a recording medium which should be capable of showing gradations of density depending on the intensity of the analyzer output. Each time the drum revolves, the stylus which marks the paper is moved laterally a small distance, and the oscillator frequency is changed slightly. Thus a picture is built up which has time as one coordinate and frequency as the other, with intensity shown by the density or darkness of the record. It may be necessary or desirable to distort the amplitudes in the analyzer output, depending on the recording medium used and the use to be made of the spectrograms. This function is indicated in the figure by the compression in the last amplifier.

The first spectrograph set up in the laboratory differed in some particulars from the arrangement shown in Fig. 7. Instead of a recording drum on the same shaft with the magnetic tape, use was made of a machine built for radio fac-

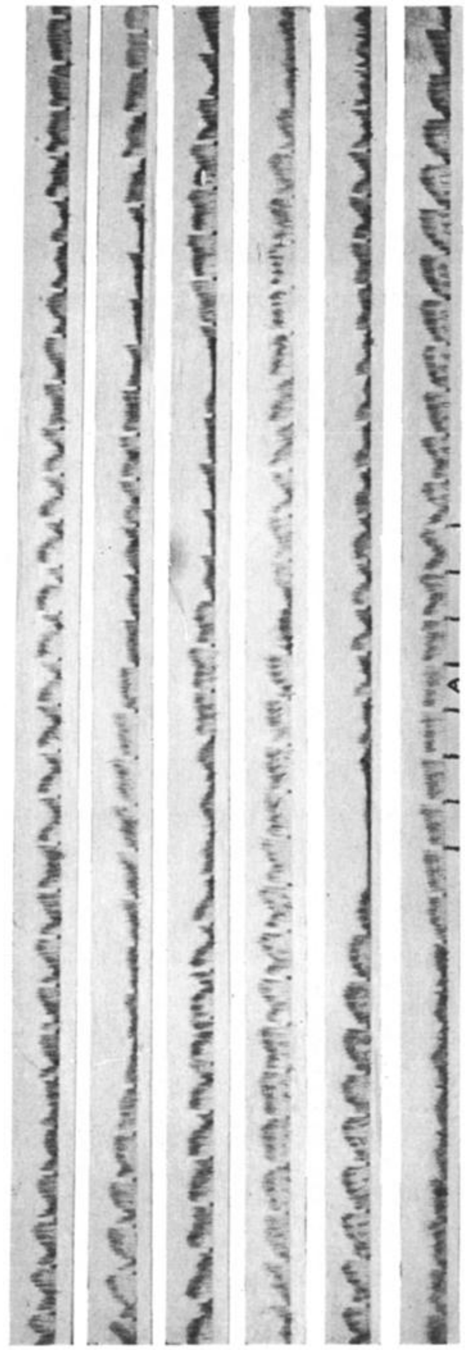

FIG. 3. Oscillogram of a series of 11 tones transmitted over a radio channel and received through narrow band filters whose outputs were repeatedly scanned with a commutator. A single cycle is included in the section marked *A*; here the circuit was momentarily almost flat; the other cycles show changing profiles due to the changing response of the circuit through selective fading.

[5] See introductory paper of this series.

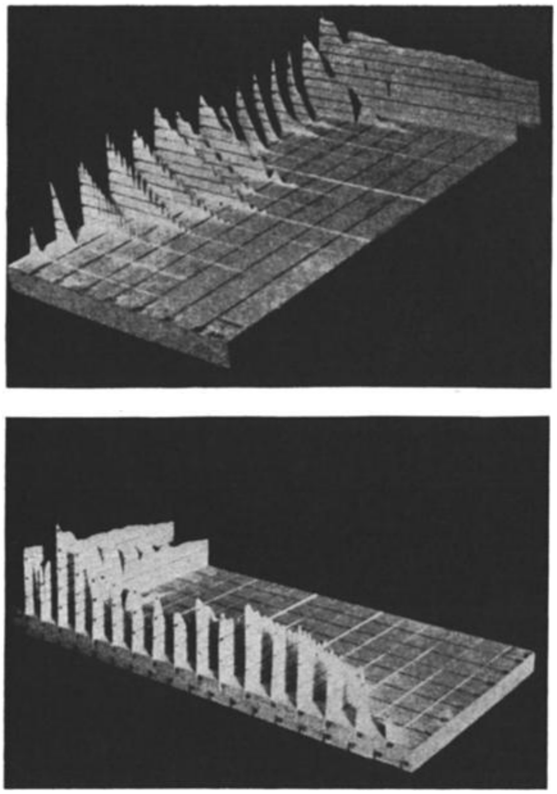

FIG. 4. Three-dimensional models portraying the amplitudes of the several harmonics of a piano note as functions of time.

simile reception, which happened to be available. This device had a cam-driven arm sweeping a stylus across a strip of conducting paper. The paper had a light-colored surface which became progressively darker as the current passing through it increased. Each sweep of the arm was started by a synchronizing signal which in this case came from a contact connected to the disk carrying the magnetic tape. The paper was automatically advanced 0.01 inch between sweeps. The machine had to be modified to the extent of slowing down the motion and providing a new cam to make the motion more uniform.

For the analyzing portion of the system, use was made of another piece of available equipment, namely an ERPI heterodyne analyzer. No mechanical connection was provided between the analyzer and the magnetic tape or the recording system, so the frequency had to be shifted by hand after each sweep of the recording arm. In the beginning, the output of the analyzer was rectified and a thyratron threshold arrangement was employed ahead of the recorder, so that it printed only when the analyzer output rose above a predetermined level, and there were no gradations in density. When it was found that intensity could be shown by the density or blackness of the record, the threshold was dispensed with; but because of the small range of currents required for printing the full density range of the paper, it was necessary to compress the sig-

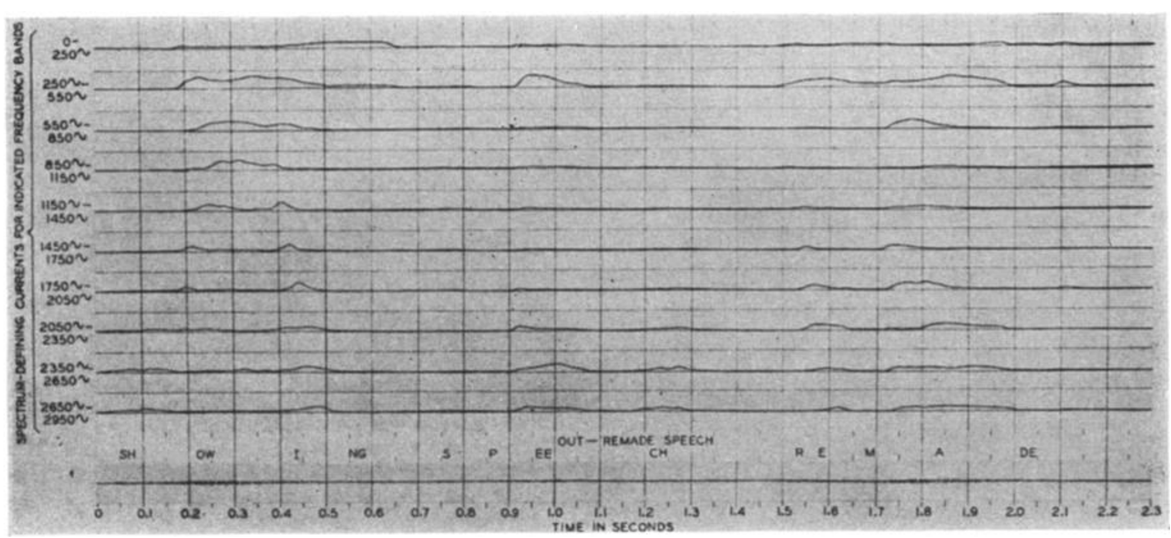

FIG. 5. Another method of illustrating the energy-frequency-time distribution in speech: a series of oscillograms of the rectified outputs of 10 band filters.

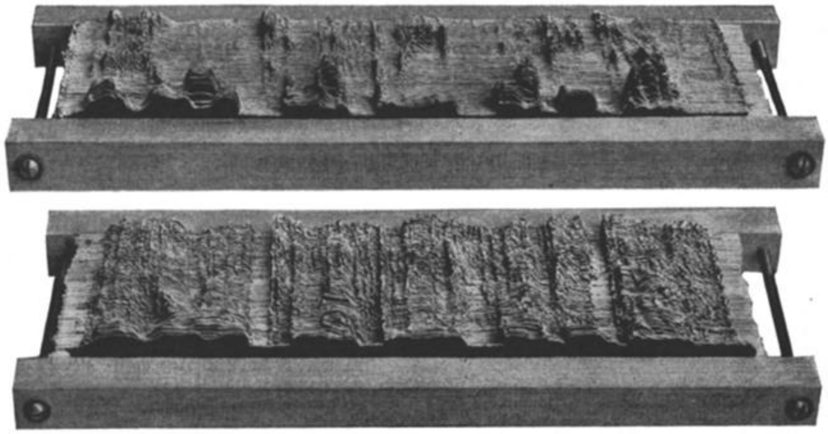

FIG. 6. A solid model built up of oscillograms, similar to those of Fig. 5, for 200 overlapping frequency bands. The words are "visual telephony for the deaf." The two models differ in the amount of amplitude compression.

nal. This was done with a sort of partial automatic volume control arrangement, and adjustments were made until a 35-db signal level range was covered, with what looked to the eye like even density steps on the paper for even db changes of level.

A sample spectrogram of speech made with this arrangement is shown in Fig. 8. The coordinates of time and frequency are indicated. The analyzing band width was 200 cycles, with the band being shifted (by hand) 50 cycles for each new sweep of the recording arm. The sentence shown is a familiar one, containing most of the vowel sounds. The most gratifying feature of these early spectrograms was the clear indication of the almost continual shifting in frequency of the dark bars which represent the vocal resonances. It would take weeks or months of harmonic analysis from oscillograms to obtain the same information. Considerable consonant detail may be seen as well. Predistortion, to the extent of raising the higher frequencies by 6 db per octave, was used in making this spectrogram.

While pitch inflection is not shown in Fig. 8, it may be brought out by two methods. If the analyzing band is made narrow enough, the separate voice harmonics appear and their rise and fall with pitch can be seen. The other method is to permit the beats between harmonics in a wide band to register, producing characteristic

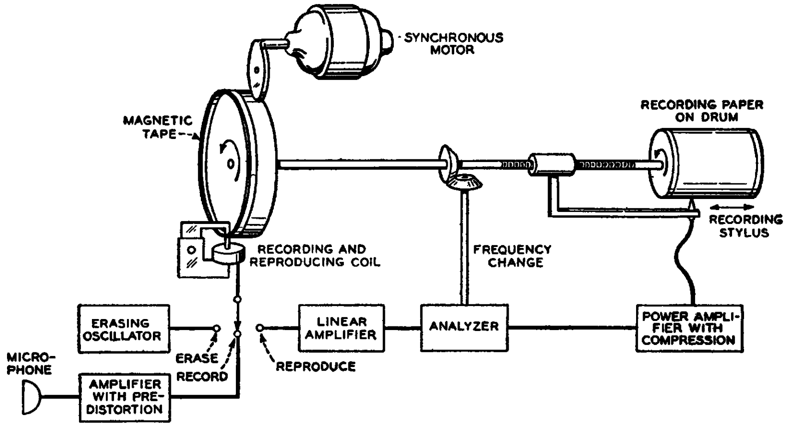

FIG. 7. Schematic representation of the basic method of the sound spectrograph. The sound is recorded on the loop of magnetic tape, and analyzed while repeatedly reproduced. The fluctuating analyzer output builds up a pattern of light and dark areas on the electrically sensitive paper.

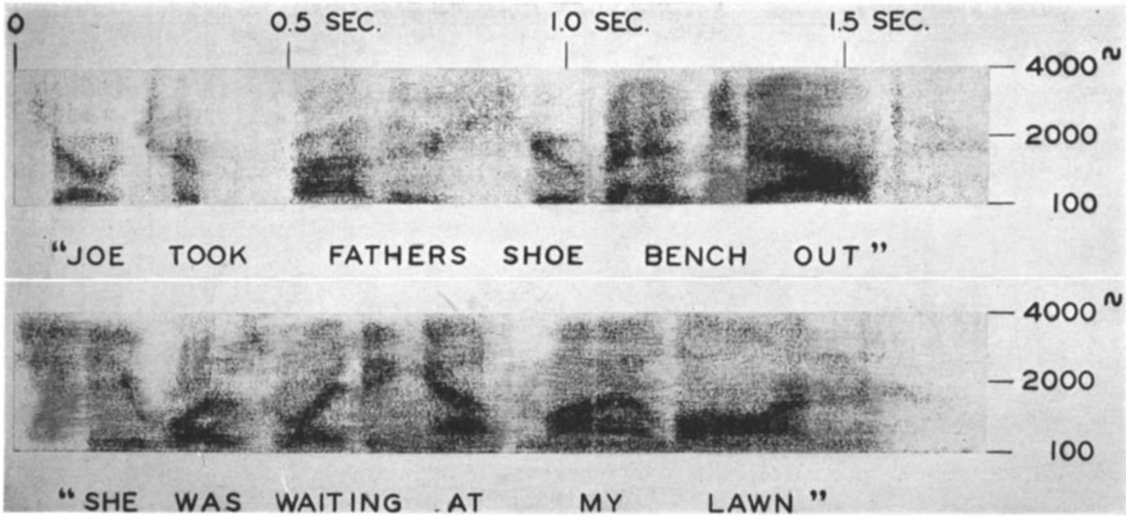

FIG. 8. Spectrograms made with the first laboratory model assembled from available equipment. The vocal resonances are clearly indicated. Further description of the features of spectrograms will be given in connection with subsequent illustrations.

striations vertically across the pattern. Both of these effects will be illustrated subsequently.

The work with this experimental equipment was carried out before our entry into the war. Because of its military interest it was given official rating as a war project and a self-contained model was developed which is described in the next section.

III. PRESENT MODEL

A photograph of the present model portable sound spectrograph is shown in Fig. 9 with the equipment set up as in operation. The recorder unit is at the right, the amplifier-analyzer is at the left with associated control circuits mounted on a panel attached above it, and the power supply is on the lower shelf of the table. The units are interconnected by means of flexible cords and connectors so they can be transported separately. This spectrograph, although basically the same as the early model described above, differs considerably in mechanical and operational details.

The recorder unit, which serves as the magnetic tape recorder as well as the spectrograph pattern recorder, is built around a modified commercial two-speed turntable of the type used for disk recording. The signals to be analyzed are recorded on a length of vicalloy magnetic tape $\frac{1}{4}$ inch wide and between 2 and 3 mils thick, mounted against a shoulder in a step turned on the lower edge of the 13″ turntable platter. About

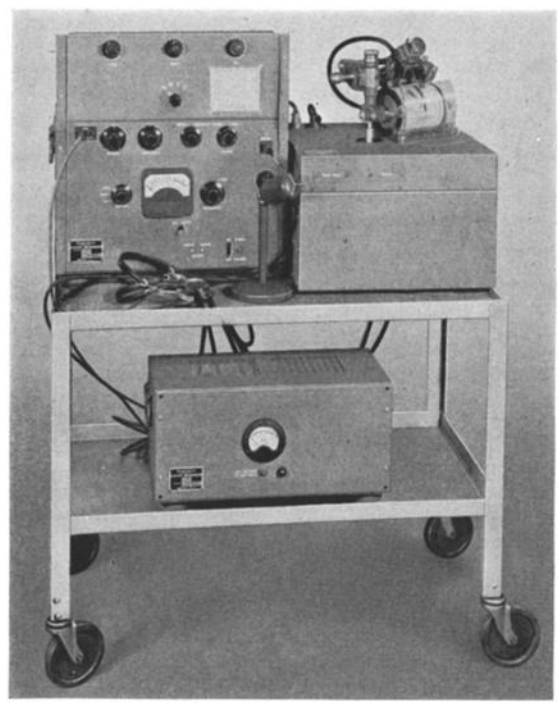

FIG. 9. The present model of the sound spectrograph, built in three parts for portability. The magnetic tape unit is in the right hand box, whose superstructure also carries the paper drum and stylus. The amplifiers, analyzer, etc., are in the left-hand box. The lower unit houses the regulated power supply.

one-third of the tape projects below the platter rim. Precise machining of the step is necessary to prevent eccentricity or wrinkling of the tape. The ends of the tape are sheared diagonally at about 45°, to a length which will provide a 2- or 3-mil butt joint when the tape is mounted. The joint is slanted so that the overhanging sharp pointed edge of the tape trails as the turntable rotates. When the tape is cut and mounted carefully, the joint is hardly noticeable in the patterns. The tape is held to the platter rim by a serving of heavy linen thread wrapped around the upper two-thirds of the tape and cemented in place by several coats of clear lacquer. The turntable is rim-driven by a synchronous motor through friction drive idlers at 25 r.p.m. for recording and 78 r.p.m. for analysis.

Two sets of magnetic pole pieces, one for recording (or reproducing) and another for erasing, are mounted on the bed of the machine so that the overhanging edge of the tape passes between their faces. The recording pole faces are about 40 mils square and are mounted with a pivot and spring so that their inside edges will just pass in a shearing fashion when the tape is removed. In order to insure good contact with the tape at the shearing edges, the pole faces are initially machined at a slight angle (2 or 3 degrees to the tape) to reduce the time for "running in" to a good fit. Very small misalignments of the pole

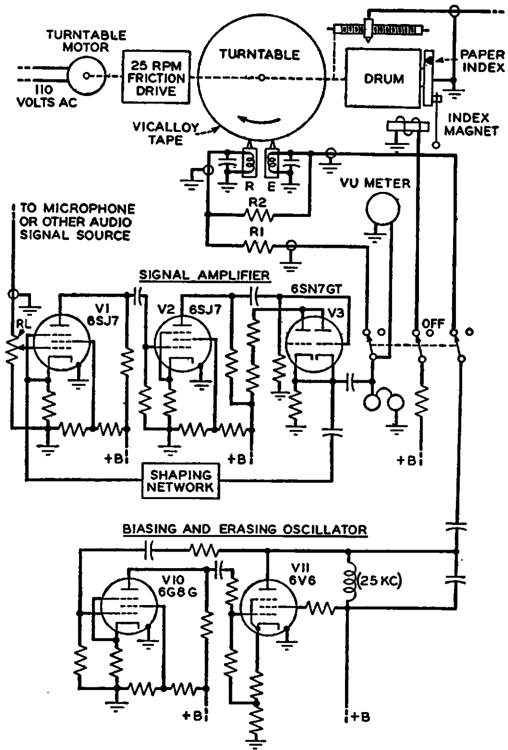

FIG. 11. Schematic diagram showing the spectrograph components arranged for recording a sample on the tape. The erasing and recording processes are continuous until stopped by the switch; the last 2.4 seconds of signal are then on the tape.

faces may cause a decided loss of high frequency response. As shown in Fig. 10, the erasing pole pieces are placed about one inch ahead of the recording pole pieces; their pole faces are slightly wider than the recording pole faces so that erasing will be effective even if there is a small relative misalignment of the two sets of pole pieces. Between the two sets of pole pieces may be seen a spreader with which they may be lifted out of contact with the tape for safety during shipment or handling of the turntable. An oil-saturated wick placed ahead of the pole pieces serves to clean and lubricate the tape as it rotates.

FIG. 10. View with cover removed, showing the magnetic tape and the pole pieces. The butt joint in the tape is exaggerated in this picture.

The superstructure mounted above the turntable (Fig. 9) serves as the spectrogram pattern recorder. In addition to the conventional cutting head carriage and lead screw, a 4″ diameter metal drum, mounted on a shaft parallel to and below the lead screw, is driven through 1:1 spiral gears from a vertical extension of the turntable shaft. A flexible stainless steel stylus about

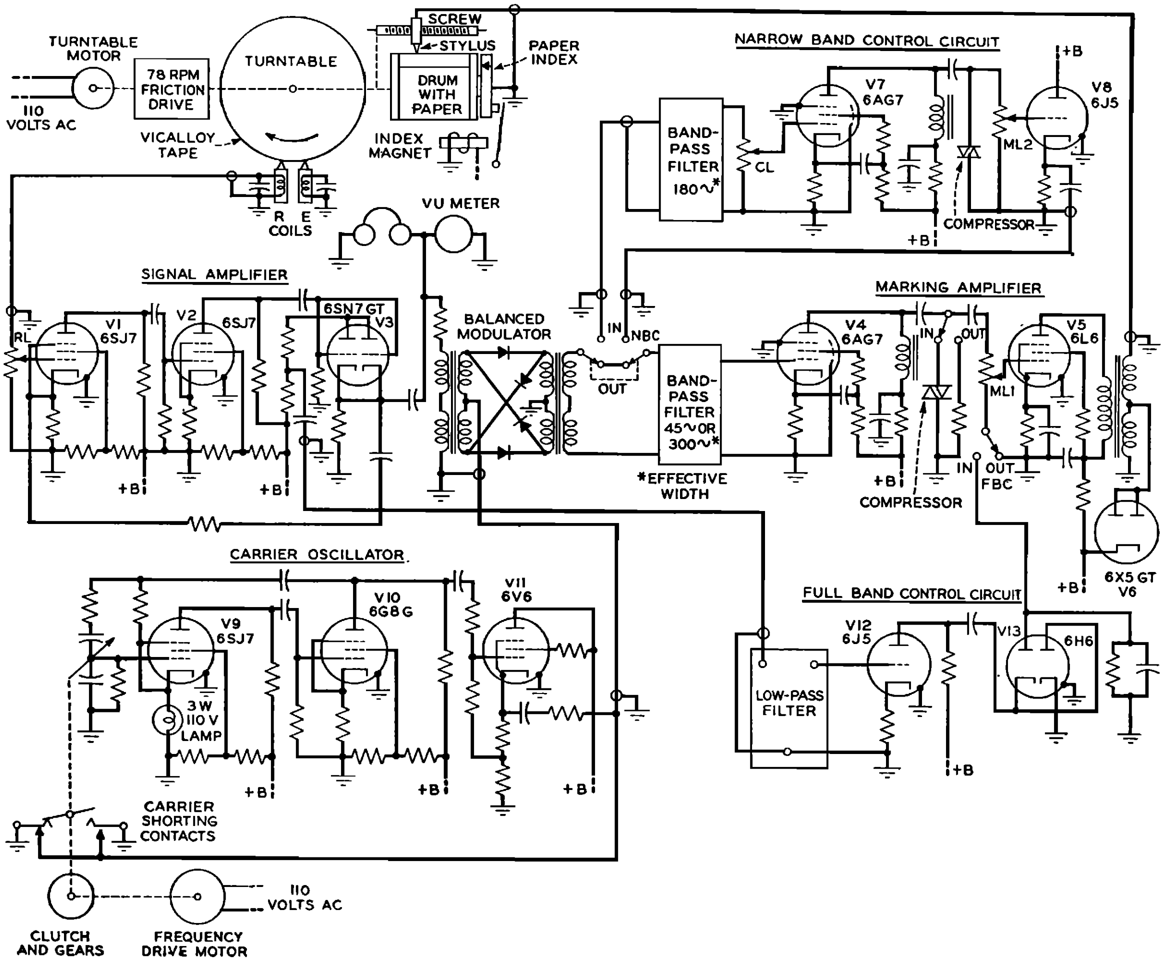

FIG. 12. Schematic diagram showing the spectrograph components arranged for reproducing and analyzing the recorded signal. The carrier oscillator is slowly swept through the frequency range while the signal is repeatedly reproduced and impressed on the modulator. The output of the band pass filter is amplified and marks the paper with a density corresponding to the instantaneous level. The control circuits govern the amplitude compression.

10 mils in diameter is mounted on an insulating block fastened to the cutting head carriage, which can be lowered to or raised from the recording position manually. An automatic paper index mounted on the drum shaft indicates the position for placing the lap in the facsimile paper so that the beginning of the sound sample recorded on the tape will coincide with the leading edge of the paper.

The schematic diagram in Fig. 11 shows the various elements of the spectrograph used in the recording condition. The signal amplifier and the biasing and erasing oscillator are in the amplifier-analyzer unit. The signal amplifier is used also for reproducing and the oscillator forms a part of the carrier oscillator for analysis of the recorded signals. Switch contacts for making the change from the "record" to the "reproduce" condition, however, are left out of the schematics to make them more straightforward.

A microphone or other signal source is connected to the amplifier input, and the recording level, which can be read on the VU meter, is adjusted by means of the RL potentiometer. Shaping networks are provided in the feedback circuit so that the relative levels of high and low frequencies in the signal may be changed if desired. The low impedance output of the amplifier is connected to the recording coil R through a high resistance $R1$ which gives essentially constant recording current for equal voltages over the frequency range of the input signal. Erasing current of 25 kc is supplied to the erase coil E by the biasing and erasing oscillator. Some of this current is also applied as a bias to the recording coil R through the resistor $R2$. Small tuning condensers across the recording and erasing coils raise the effective coil impedance at 25 kc but have little effect in the voice frequency range. These condensers also serve to reduce oscillator switching transients which otherwise tend to magnetize the tape so strongly that complete erasure is difficult.

The turntable and tape are driven at 25 r.p.m. which permits recording a sample of 2.4-second duration at a linear tape speed of approximately 16 inches per second. Continuous recording and erasing are effected as long as the switch is in the recording position. The erasing coil, since it is just ahead of the recording coil, erases signals which were recorded 2.4 seconds earlier. When the "record-reproduce" switch is thrown to the right, erasing and recording are simultaneously stopped, thereby capturing the last 2.4 seconds of signal on the loop of tape, after which it may be reproduced over and over.

The paper index magnet is energized in the recording position, and the paper index rotates frictionally with the drum until the index magnet armature engages a pin projecting on the side of the paper index, causing it to remain in a fixed position. When the index magnet is released at the end of a recorded sample, the paper index is free to rotate with the drum again. The paper index then indicates the position on the drum corresponding to the end of the sample recorded on the tape. Since the drum is directly geared to the turntable on which the tape is mounted, the above relationship will remain fixed and the facsimile paper can be placed on the drum so that the end of the paper coincides exactly with the end of the recorded sample. After a desired sample has been recorded, the turntable is stopped and a pre-cut sheet of electrically sensitive paper is secured to the drum by means of rolling springs.

The turntable is then speeded up to 78 r.p.m. by shifting idler pulleys in the friction drive. It should be pointed out that, although the signal sounds unnatural when speeded up approximately 3 to 1 by this shift, its wave form is unaltered. The effect is merely to divide the reproducing time approximately by 3 and multiply all of the frequency components of the recorded signal by the same factor. This operation converts the original frequency range of approximately 100 to 3500 cycles to about 300 to 10,500 cycles. Since it is very difficult to vary a narrow band-pass filter over this range for analysis, a heterodyning process is used with a fixed band filter.

Figure 12 shows the various circuits of the spectrograph in the reproducing condition. Starting at the upper left of the diagram, the signal is picked up from the magnetic tape by the reproducing coil R. A small equalizing condenser in shunt resonates the coil at about 12,000 cycles to keep the high frequency response from falling off. This provides an over-all frequency response which is essentially flat when the sample has been recorded without pre-equalization.

The reproducing coil works into the RL potentiometer which is used to adjust the level of the reproduced signal to a suitable value. The three-stage feedback signal amplifier has a resistive feedback network to provide a flat frequency response with a voltage gain of about 100 times while reproducing.

The output of the signal amplifier is impressed on the balanced copper oxide modulator. Carrier is supplied to the modulator by the R-C carrier oscillator through the cathode follower amplifier stage $V11$. As the carrier frequency is slowly

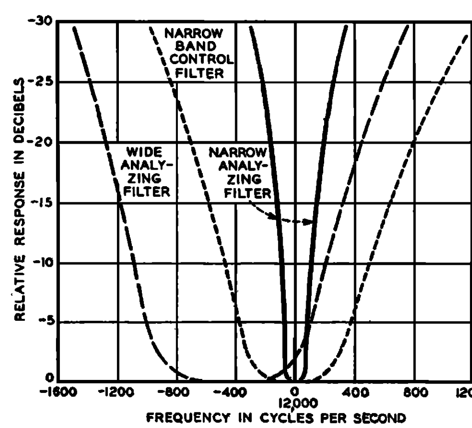

FIG. 13. The characteristics of the several filters used in the spectrograph. The effective band widths are less than those shown by a factor of 3:1, which is the speed-up ratio of reproducing *vs.* recording.

shifted by the frequency drive motor from approximately 22,500 cycles to 12,000 cycles, the lower sideband output of the modulator is swept slowly across the fixed band filter connected to its output. The normal sweep rate is such that the turntable makes approximately two hundred revolutions in the time required for the frequency to change from the highest to the lowest value. Since the stylus advances about 10 mils for each revolution of the turntable, the resulting spectrogram is approximately 2 inches wide for the 3500-cycle frequency range. The carrier shorting contacts, which are operated from a cam on the condenser drive shaft, switch the carrier current on and off to define the high and low frequency boundaries of the spectrograms.

Either of two filter widths may be slected by a switch which selects one of two values of mutual capacity between the anti-resonant sections of the band-pass filter. The characteristics of the two filters are shown in Fig. 13. It will be noted that the mid band frequency of the wide analyzing filter is lower than that of the narrow filter. With this particular filter structure, changing the pass band is accomplished by shifting the lower cut-off; the upper (theoretical) cut-off remains the same. These curves show the actual pass bands of the filters; the effective widths are about one third of the indicated widths, because the frequencies of the speeded-up reproduced signal are spread apart by a factor equal to the speed-up ratio, namely about 3:1. As pointed out previously this does not alter the wave shape of the signal but spreads out frequencies and reduces time, with a net effect of speeding up the analysis by a factor of three.

The output of the analyzing filter is impressed directly on a two-stage amplifier which has enough gain to raise the rather low voltage to a value sufficiently high to mark the facsimile paper. The output of the amplifier is not rectified, but is connected to the stylus through a step-up transformer having a turns ratio of about 1:2. Under some conditions a very high signal may reach the grid of the final marking amplifier stage. If the stylus is not lowered to load the amplifier, very high positive peak voltages tend to appear across the output transformer secondary. The biased diode (V6) is shunted across a portion of the high winding to limit the peak voltage to a value which will not damage the transformer insulation. When the stylus is on the paper, the normal range of marking voltages does not exceed the bias voltage and the diode has no effect. A rather small transformer with ordinary insulation will safely withstand the voltages with the diode protection.

In this connection it may be noted that the process of recording on the paper generates a considerable amount of smoke. A blower is therefore incorporated in the recorder unit to draw the smoke through a charcoal filter. The hinged plastic shield which may be seen in Fig. 9 directs the smoke into a slot in the top of the recorder box.

The marking range of the paper is limited to about 12 db. A 70-volt signal on the stylus will make a barely visible mark and about 300 volts will mark the paper full black at the linear paper speed of approximately 16 inches per second. Since for some applications it is desirable to show components of speech which cover a range of 30 or 40 db, it is necessary to apply a compressing action to the marking amplifier to reduce the signal range to the limited range of the facsimile paper. One kind of compressing action is secured by shunting across a high impedance point a non-linear compressing resistor (thyrite) whose voltage varies as the 3rd or 4th root of the current through it. This form of compression operates directly on the wave shape, and there is no time constant involved. It produces a rather uniform gradation of blackness on the recording paper over a 35 or 40 db range of signal intensity and brings out low level detail in the signal being analyzed. However, the compressing action when secured in this manner tends to degrade the frequency resolution of the filter. Better definition has been secured by the use of control circuits which accomplish compression in a different manner.

Two control circuits are provided as shown in Fig. 12. The narrow band control circuit is used with the narrow (45-cycle) analyzing filter and the full band control circuit is used ordinarily with the wide (300-cycle) filter. The narrow band control circuit is a series arrangement which may be switched in ahead of the 45-cycle analyzing filter. It includes a control filter whose characteristic is shown in Fig. 13. Its pass-band is

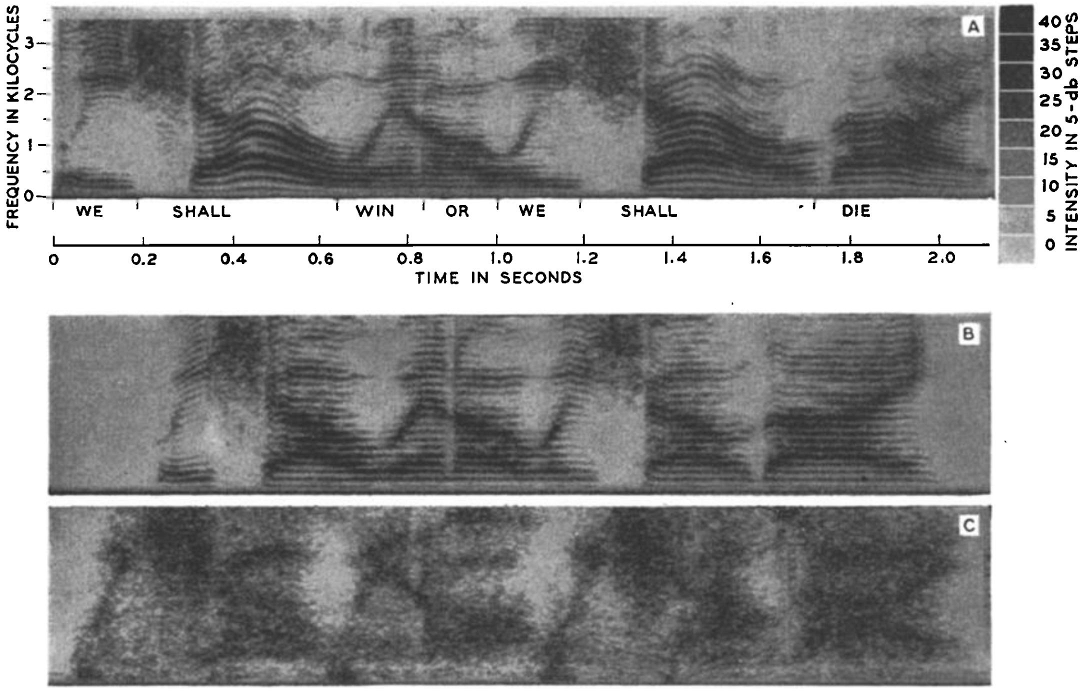

FIG. 14. Spectrograms of speech made on the present model, using a 45-cycle analyzing filter. These and subsequent illustrations have been trimmed and reduced. The originals are 12.5 inches long (for 2.4 seconds) and 2 inches high, recorded on pre-cut sheets 4.5 inches wide. Section (A) normal speech. (B) monotone. (C) whisper.

about 180 cycles (effective) surrounding the pass-band of the narrow analyzing filter. A compressor having the same characteristics as the one described above is inserted in the control circuit amplifier. Since the compressing action takes place ahead of the analyzing filter, the frequency resolution of the analyzing filter is not impaired as it is when the compressing action is placed after it. Compression at this point is permissible because the control filter passes such a narrow band that no important modulation products generated by the compressor fall in the pass-band of the analyzing filter. Various degrees of control can be obtained by changing the working level of the compressor by means of the gain controls provided.

The full band control circuit ordinarily but not necessarily used with the 300-cycle analyzing filter, is a shunt arrangement. The low pass filter passes the entire frequency spectrum of the signal which is then amplified and rectified with a time constant of 2 milliseconds. The filtered or smoothed d.c. output of the rectifier is applied as a bias to the marking amplifier grid thereby controlling its gain. This arrangement serves to control the gain after the analyzing filter by the spectrum ahead of the filter. The direction of control is to reduce gain for higher spectrum energy.

The effect of these control circuits on the spectrograms will be illustrated in the next section.

Returning to the photograph Fig. 9, the location of the various items mentioned above is as follows. The panel attached above the amplifier-analyzer unit houses the control circuits, which were added relatively recently. The two left-hand dials regulate the degree of narrow band control, and the right-hand dial the degree of full band control. The switch on this panel selects the type of control—"full band," "narrow band," "compressor (in the marking amplifier) in" and "compressor out."

The main panel has two attenuators at the upper left, one for controlling the recording level, the other for the reproducing level, both levels

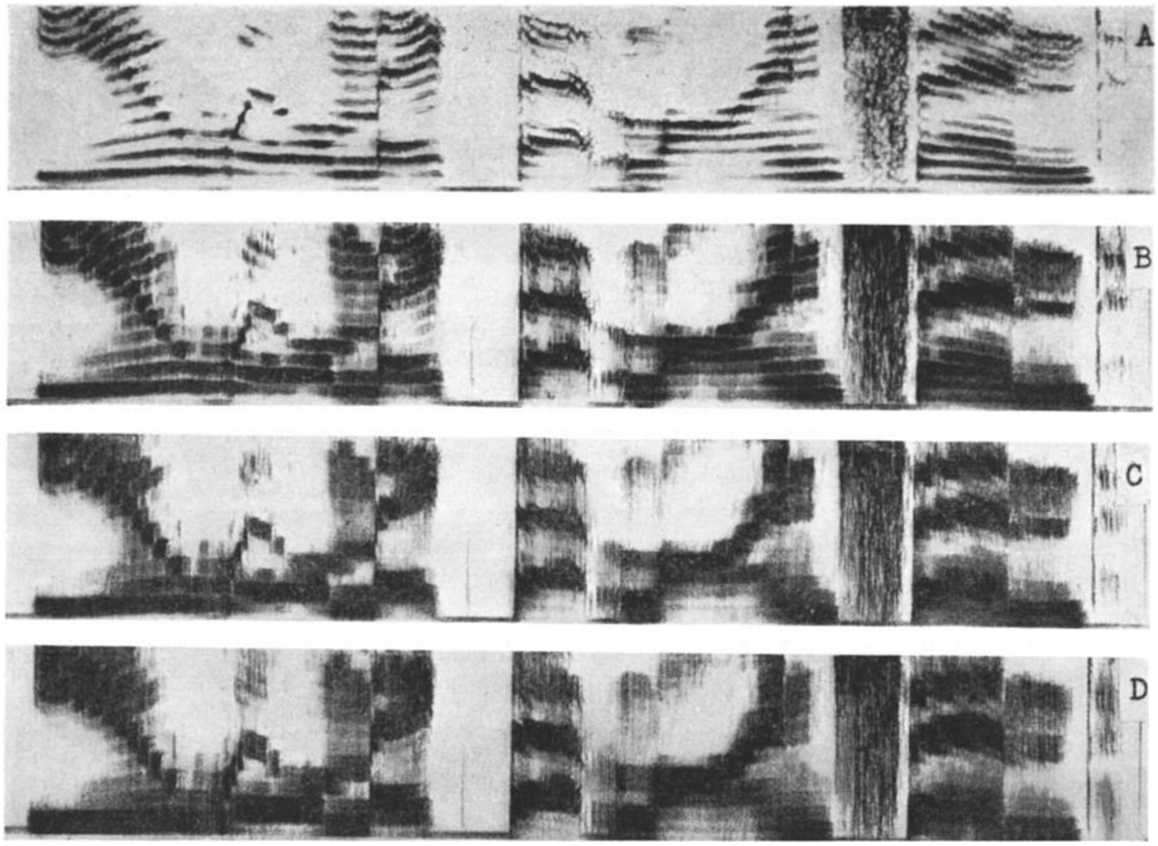

FIG. 15. Showing the effect of widening the analyzing filter, with a high pitched voice. The widths are 90, 180, 300, and 475 cycles in Sections (A) to (D), respectively. The words are "you will make that line send."

being read on the VU meter in the center. The next knob is the recording-reproducing switch, and the next the filter selecting switch. To the left of the VU meter is the predistortion selector, and to the right is the dial for resetting the carrier oscillator at the beginning of each pattern. This dial returns the rotating condenser plates to their starting position; a small motor inside the panel then drives them slowly back through a friction clutch.

Just below the VU meter is a gear shift lever by means of which the rate of frequency sweep can be changed so as to make the patterns 4 inches high instead of 2 inches. The two buttons at the bottom permit reading the carrier and marking voltages, respectively, on the meter.

The power supply unit provides filament and plate voltages to the rest of the equipment. It is highly regulated and operates at 280 volts from 110-volt a.c. power.

IV. SPECTROGRAMS OF SPEECH

Figure 14A shows one kind of spectrogram produced by the later models of the sound spectrograph. The well-known sentence which appears in this spectrogram has been used for testing and illustrative purposes because it contains monosyllabic words with a variety of resonance patterns. The frequency scale of the spectrogram is linear and covers 3500 cycles as shown by the scale at the left. The time scale is also linear and covers 2.4 seconds (in the illustrations the spectrograms have been trimmed at the ends). They have also been photographically reduced; in the originals the vertical height is 2 inches and the length is 12.5 inches, making the time scale slightly over 5 inches per second.

This spectrogram shows a great deal more detail than the ones in Fig. 8; it was made with a much narrower analyzing filter—about 45 cycles wide at the 3-db points. With this filter the

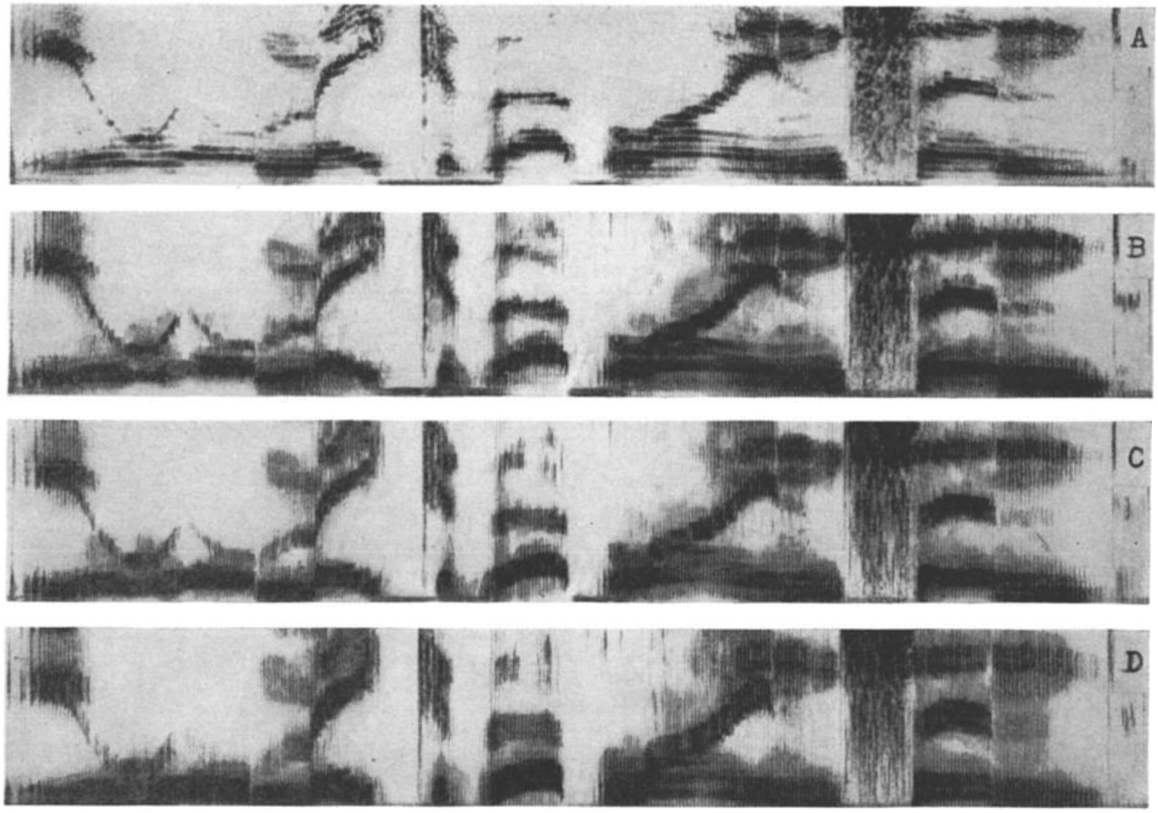

FIG. 16. Same as Fig. 15, but with low pitched voice.

individual harmonics of the voiced sounds can be clearly distinguished. The traces curve up and down as the pitch of the voice is varied in normal speech. The wider the spacing between harmonics, the higher the pitch at any particular instant. In the word "shall" for instance, the pitch first rose and then fell. If the words are spoken in a monotone, the harmonic traces remain level and equi-distant as shown in Section B of Fig. 14. If the words are spoken in a whisper, the spectrogram appears as in Section C. It will be noted that the distribution of dark and light areas is closely similar in all three spectrograms. These dark areas indicate the regions of maximum energy—in other words the vocal resonances. In the whispered words the vowels and consonants all have the same fuzzy texture with no harmonics present. The same texture appears in normal speech in unvoiced sounds such as the "sh" sounds in Section A.

These spectrograms were made with the compressor in the marking amplifier. The gradations of black produced by various signal intensities are shown in the upper right-hand corner. Since the last two steps are nearly alike in blackness, the range is somewhere between 35 and 40 db.

The effect of widening the analyzing filter can be seen in Fig. 15. These are spectrograms of a rather high pitched female voice. The filter used in Section A of Fig. 15 was twice as wide as that in the previous illustration, that is, 90 cycles. In Section B the filter width was 180 cycles, and the harmonics of this high pitched voice are still clearly resolved. In Sections C and D the filter widths were 300 and 475 cycles, respectively; with these wide filters the individual harmonics tend to merge and only the resonant areas can be clearly resolved. The first word in these illustrations shows clearly that the trend of the resonant areas may be opposite from that of the voice pitch. It is evident that the pitch is rising in this first word but the frequency of resonance is rapidly falling so that each harmonic in turn is reinforced momentarily,

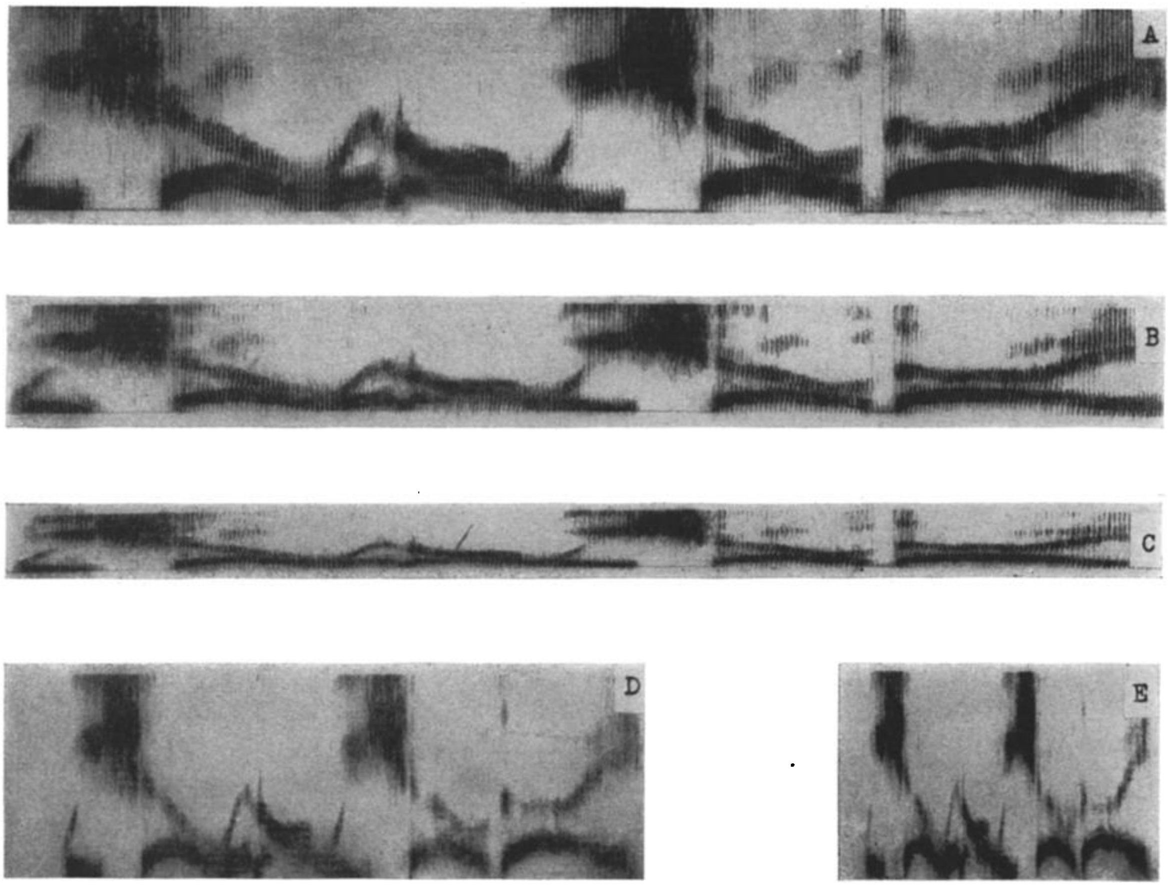

FIG. 17. Spectrograms illustrating various aspect ratios, using the 300-cycle analyzing filter. Section (A) "normal." (B) and (C) frequency dimension reduced by 2 and 4, respectively. (D) and (E) time dimension reduced by 2 and 4, respectively. The words are "We shall win or we shall die," identical sample in all spectrograms.

producing a step effect which is somewhat undesirable for visible speech purposes. With a lower pitched voice such as is illustrated in Fig. 16, the resonance areas tend to form smooth dark bands as soon as the filter becomes wide enough so as not to resolve the individual harmonics. The filter widths in this figure are the same as in Fig. 15. With most male voices a filter about 200 cycles wide would be adequate to smooth the resonance bands. A 300-cycle width has been adopted as a compromise, and is adequate for most voices.

In Fig. 16 it will be noted that there is a distinct pattern of vertical striations in the voiced sounds. This pattern is caused by the fact that more than one harmonic is passed by the analyzing filter. It is well known that two or more frequencies separated by equal intervals will produce beats at the interval frequency. In the case of speech, this frequency of course is the voice pitch. Each vertical striation represents the crest of a beat, and the separation between crests can be seen to vary as the pitch changes. In Fig. 15, where the pitch is very high, the vertical striations are so close together as to be barely distinguishable. Incidentally these vertical striations sometimes persist unbroken across the whole frequency range, which is probably due to the particular kind of phase relations resulting from the mechanism of phonation. Sometimes there are phase reversals in the striations, however; this subject would make an interesting study in itself and probably throw light on the voicing mechanism.

The time and frequency dimensions of spectrograms were originally chosen so as to give ade-

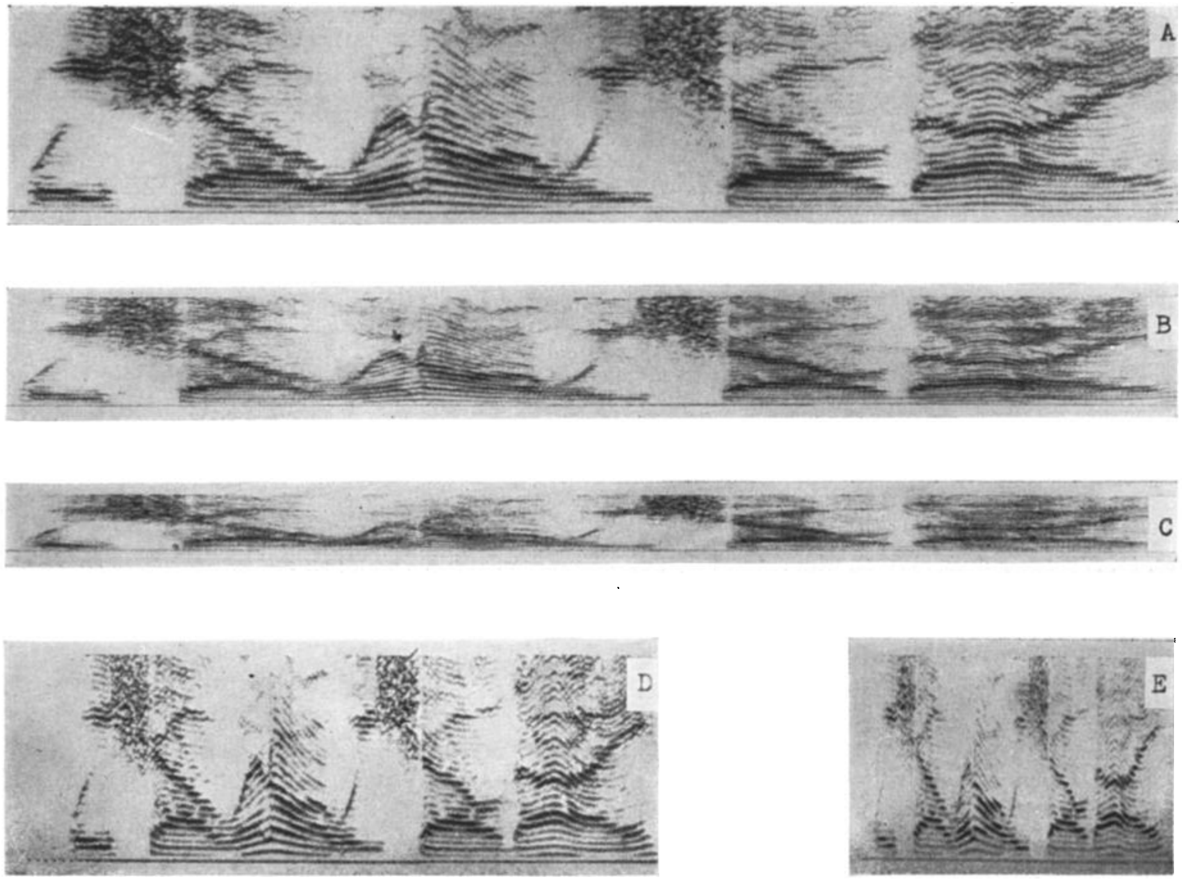

FIG. 18. Same as Fig. 17, but with 45-cycle analyzing filter.

quate resolution for a particular purpose. When the spectrograph was applied to visible speech, the question naturally arose whether some other choice of dimensions might not be more suitable. Figure 17 shows some spectrograms made to explore this question. The upper one (A) has the original aspect ratio, that is, the ratio of the frequency to the time dimensions. In (B) the frequency dimension has been cut in half by increasing the speed of the carrier oscillator sweep, leaving the time dimension the same as before. In (C) it has been reduced by a factor of 4. Changing the aspect ratio in this direction apparently does not contribute to the appearance or legibility of the patterns; rather, it tends to reduce the curvatures of the resonant traces and thereby make the different sounds less easily distinguishable. Spectrogram (D) was made with the opposite kind of change. Here the time dimension was cut in half by decreasing the diameter of the paper drum, while the frequency dimension was left normal. In (E) the time dimension was reduced by a factor of 4. After a study of these and other samples, it was felt that while the aspect ratio originally chosen might not be an optimum, at least there would be no very great advantage in changing it.

Figure 18 shows the same series of spectrograms and the same speech material analyzed with the narrow filter. Here again there appears to be no particular virtue in changing the aspect ratio. In addition there is a distinct loss of detail when the frequency scale is reduced. The "normal" spectrogram is made up of 200 horizontal lines so close together that the frequency is shifted only about $\frac{1}{3}$ the width of the filter. Spectrograms (B) of Figs. 17 and 18 were made with only 100 lines and (C) with only 50. It is apparent from these illustrations that when the narrow analyzing filter is used the number of

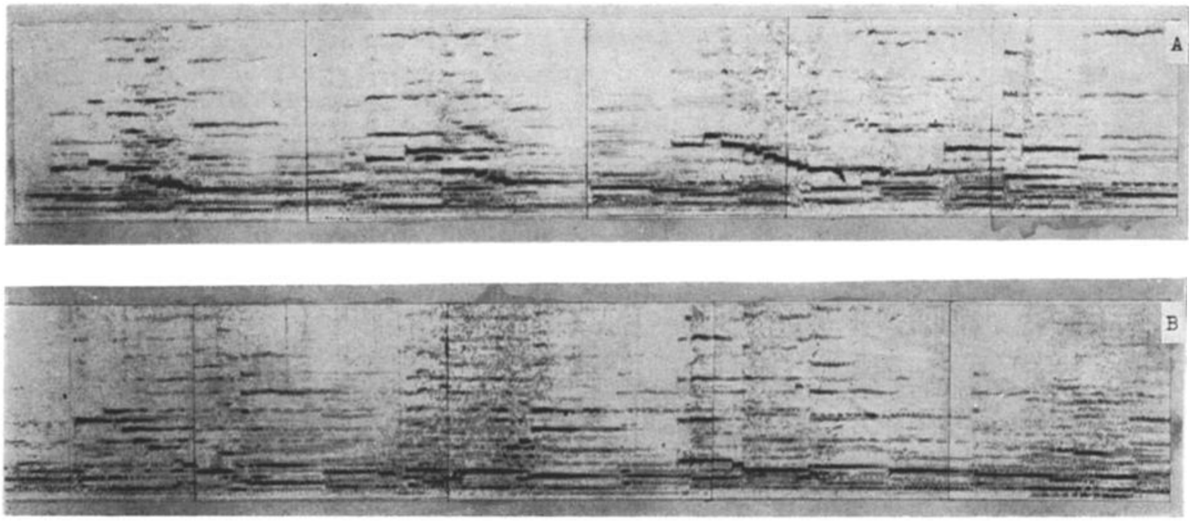

FIG. 19. A passage of piano music with reduced time dimension as in 18(E). This aspect ratio seems quite appropriate for music, though not for speech.

lines cannot be materially reduced, at least with low pitched voices, without loss of frequency detail.

In this connection it may be mentioned that while the normal aspect ratio appears to be nearly optimum for speech, it is not necessarily so for other applications. Figure 19, for instance, portrays a passage of piano music with the time dimension reduced by a factor of 4. With this aspect ratio the musical action is clearly indicated while speech, as shown in the previous illustrations, appears much too crowded in the time dimension. For other applications it has been found desirable to lengthen rather than shorten the time dimension.

Figure 20 illustrates in a somewhat different fashion the effect of changing the amount of frequency overlap in successive lines. Section (A) shows a normal spectrogram with the narrow filter and (B) shows the same sample with the same time and frequency scales; however in making spectrogram (B) three successive lines were made alike, and the frequency was then shifted by three times the normal amount. The effect is the same as though a stylus three times as wide as normal had been used with the lateral shift and the frequency shift also three times normal. The harmonics now appear to rise and fall stepwise rather than smoothly. With the wide filter, however, this process can be carried much further. Section (C) for instance shows the same material analyzed with the wide filter. Here each line has been repeated 4 times and it requires close inspection to distinguish this from a normal spectrogram. In Section (D) each line has been repeated 12 times. Obviously this has carried the process too far, but these illustrations show that if only a wide analyzing filter is of interest, considerable time could be saved by reducing the number of lines in the spectrograms and hence reducing the number of times the sample must be reproduced in the process of analyzing it. In the present spectrograph of course the analyzing time has been set by the requirements of the narrowest analyzing filter rather than the widest.

Figure 21 shows the need for some kind of compression in portraying speech sounds and illustrates the action of the full band control circuit mentioned in Section III. Spectrogram A shows the words "one, two, three, four, five, six" made without any compression except a certain amount of limiting action in the last power stage of the marking amplifier. The strongest consonants such as the "t" in "two" and the "s" in "six" can be seen in the spectrogram, but the weaker consonants such as the "f" in "four" and "five" are completely missing. Section B was made with the thyrite compressor in the marking amplifier. This spectrogram shows all the

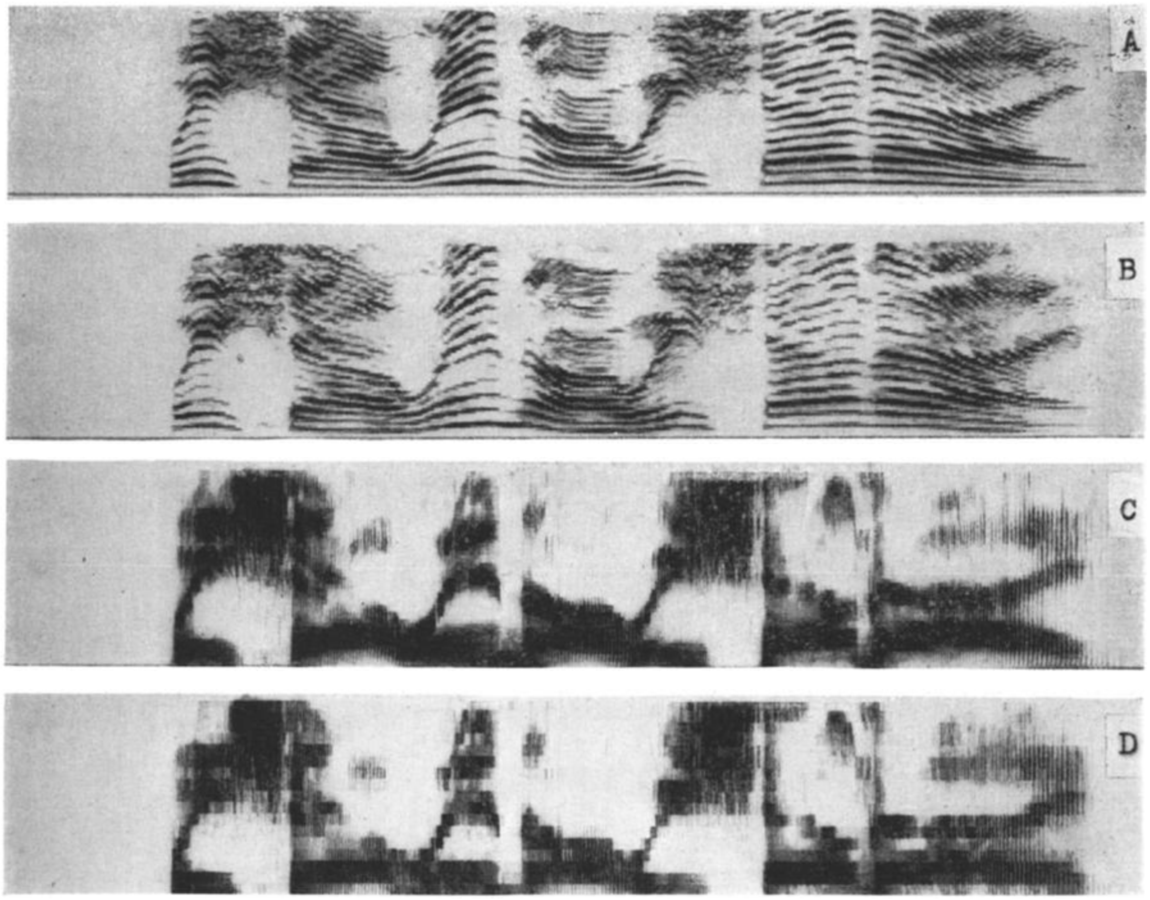

FIG. 20. Showing the effect of degrading the frequency resolution by laying down several identical lines, then shifting the frequency by a correspondingly greater amount to keep the normal frequency dimension. Section (A) normal spectrogram, (B), (C), and (D) 3, 4, and 12 repetitions, respectively.

consonants clearly. However, the low level portions of the vowel sounds are also brought out and the resonance bars are made fuzzy. For visible speech purposes it was desired that the consonants should show as clearly as they do in (B), and at the same time that the vowels should be portrayed as they are in (A). The full band control circuit accomplishes these purposes as illustrated in Section (C). Here the consonants show at least as clearly as they do in (B) and the vowel bars are resolved about as they are in A, with the low level material suppressed. The action of the full band control circuit, as discussed in the previous section, is to adjust the gain of the marking amplifier not according to the momentary output of the analyzing filter, but according to the energy in frequency bands other than the one being scanned at the moment. With this arrangement sounds in which the energy is weak in all portions of the frequency range are amplified, but the low level portions of vowel sounds are suppressed.

Figure 22 illustrates the action of the narrow band control circuit, using the same words as Fig. 21. Again Section (A) was made without any compression except the limiting action of the last marking stage. Section (B) was made with the thyrite compressor. It is clear that a wider range of levels is covered in (B), which is desirable; however, the harmonics have become fuzzier and tend to merge together. Section (C) was made with the narrow band control circuit with a setting such as to show the same range of levels as appears in (B), but here the harmonic traces are much more clear-cut than in (B). The control can be set to bring out an even wider

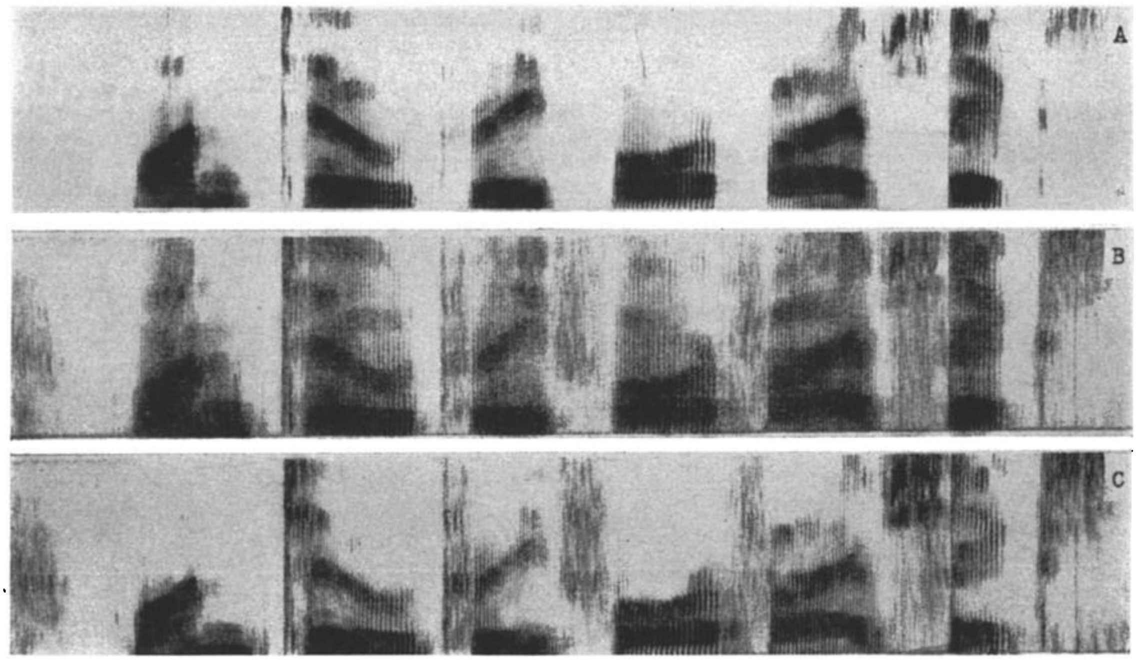

FIG. 21. Illustrating the need for amplitude compression and the action of the full band control circuit. The words are "one, two, three, four, five, six." Section (A) no compression. (B) compressor in the marking amplifier. (C) full band control.

range of levels, if desired, still leaving the harmonics clearly resolved. The action of this circuit is to adjust the over-all gain depending on the energy in the frequency regions immediately adjacent to the one being scanned. The gain is not raised when the analyzing filter happens to pass between two harmonics, but it is raised when the energy level is generally low.

V. OTHER APPLICATIONS

Thus far attention has been concentrated on spectrograms of speech. Other types of signals can be handled as well. The spectrograph is a useful laboratory tool for determining the nature of the time and frequency distributions of energy in sounds other than speech, or in complex waves of any kind whether they exist as sounds or not. Some applications of this type will be illustrated in this section.

Figure 23 shows in Sections (A), (B), and (C) some samples of thermal noise. In Section (A) it was reproduced at a very high level and it shows almost uniform darkness except for occasional light patches randomly spaced. In Section (B) it was reproduced at a lower level and the characteristic pattern of this type of wave can be clearly seen. Since in thermal noise the energy is concentrated in different frequency regions in different instants of time, the spectrogram shows randomly spaced vertical spindles whose length corresponds roughly to the width of the analyzing filter, which in this case was 300 cycles. In Section (C) the 45-cycle filter was used and the spectrogram shows fuzzier patches due to the slower response of the narrow filter. Many familiar sounds contain such random components. Section (D) for instance shows the sound of striking a match. The dark area at the left portrays the sound generated by the abrasive, the longer portion the roar of the flame. Both of these areas contain random noise as previously illustrated. In addition, sharp vertical lines may be discerned across the whole frequency range. Sharp lines of this type represent clicks or crackling sounds which are characterized by a wide frequency distribution of energy and short duration.

Section (E) shows the sound of filing on a metal plate. Three file strokes are included in the picture, and the actual sound of the file

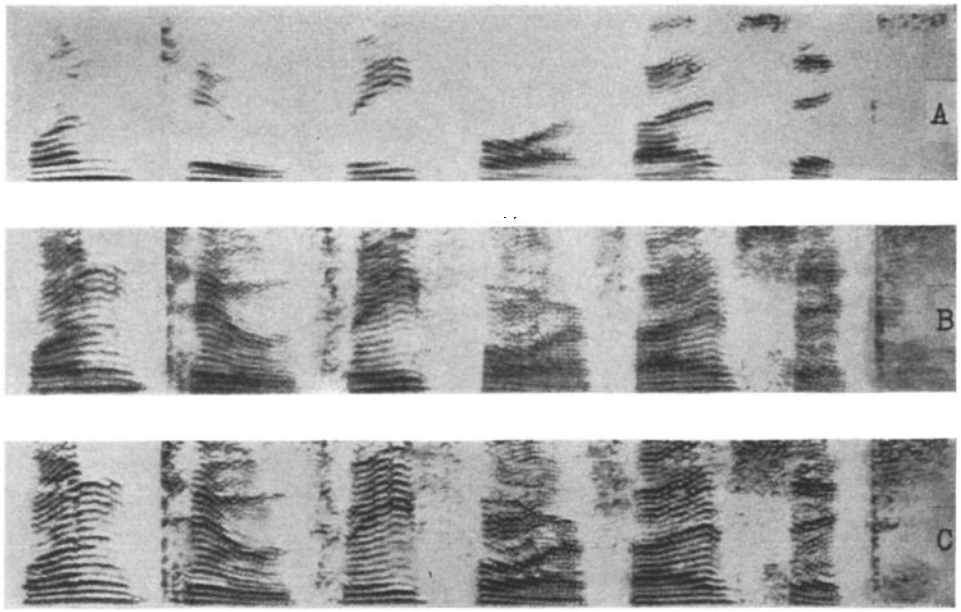

FIG. 22. Illustrating the need for amplitude compression and the action of the narrow band control circuit. Section (A) no compression. (B) compressor in the marking amplifier. (C) narrow band control.

teeth can be seen dimly as rising and falling bands in the lower part of the spectrogram. The dark regions in the upper part of the spectrogram represent resonances in the metal plate which of course remain substantially constant in frequency but vary in amplitude. The definite components can be distinguished from the random components by their more uniform texture. Section (F) shows some machinery noise. There is a low frequency periodicity in the random components, and a dark band across the middle which indicates some discrete components. The small section at the right was made with the narrow filter and shows the four discrete components more plainly.

Whether the wide or narrow analyzing filter should be used depends on the nature of the sound and how far apart in frequency the components are. Figure 24 presents two sounds analyzed with both filters, showing that in some cases the two sets of information from the wide and narrow filters complement each other. Sections (A) and (B) represent the sound of an infant crying. Each cry begins and ends with voiced components of very high pitch. In the middle of each cry the voice breaks down into a very irregular noise. In this example the components are far enough apart so that they are easily resolved with the wide filter. However, the narrow filter permits the exact frequencies of the components to be determined. Sections (C) and (D) analyze the sound of snoring. The wide filter shows that the inspiration phase consists of a series of sharp, regularly spaced clicks, with the frequency distribution modified by resonances in two portions of the frequency range. The expiration phase is practically random noise but also shows reenforcement by resonance in definite frequency regions. The narrow filter reveals that the resonances are unexpectedly sharp considering the nature of the oral cavities.

Figure 25 illustrates some bird songs analyzed with the wide and narrow filters. Sections (A) and (B) show the song of the nightingale. There are two distinct types of song in this illustration. One is a rapidly falling note with a double trill at the end. The other is a rising and falling note with an almost linear trend in both directions. It will be seen that each note is accompanied by some noise components. This is recognized with more assurance with the wide filter than with the narrow filter because the narrow filter tends to produce a fuzzy trailing edge when it scans across a component. Since the wide filter has

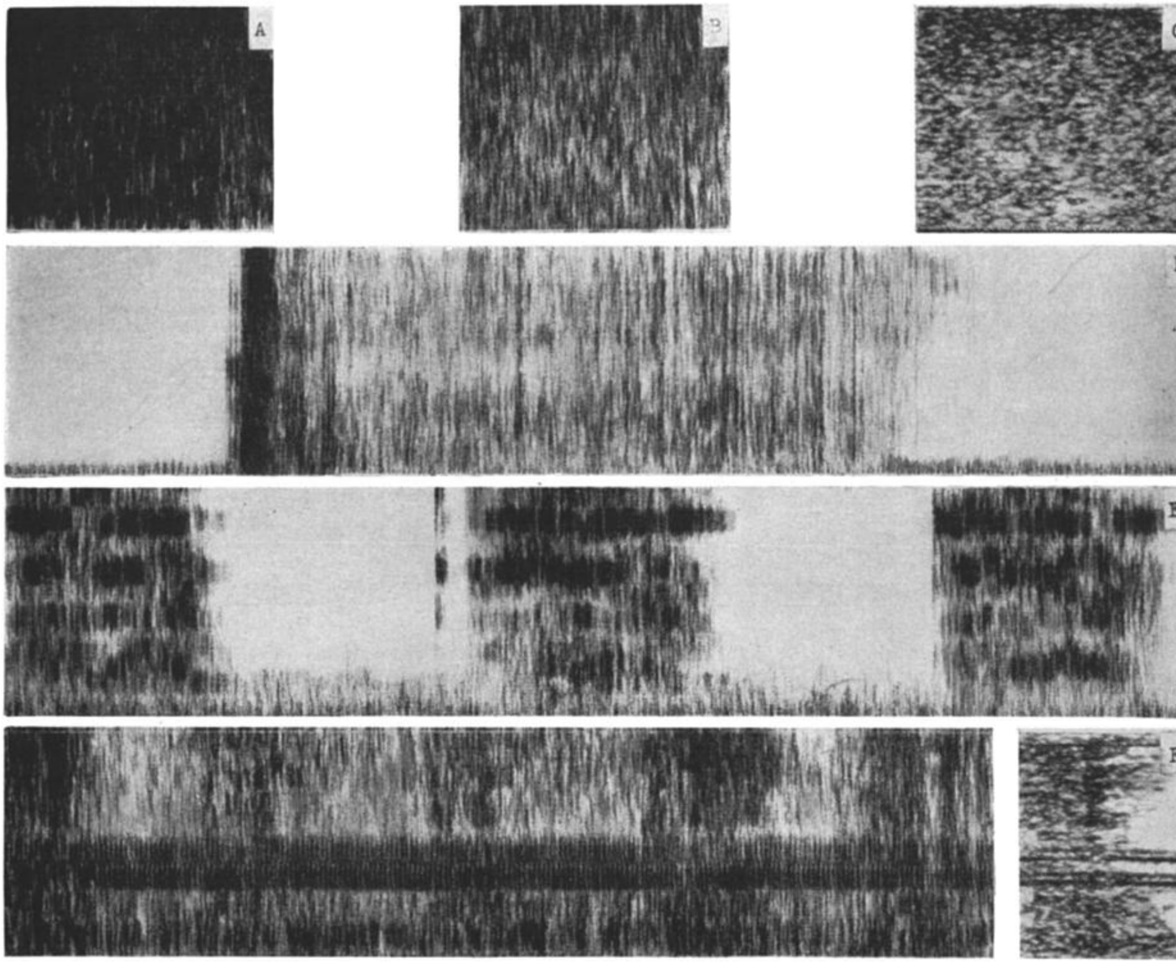

FIG. 23. Some sounds other than speech. Sections (A), (B) random noise at two levels, 300-cycle filter. (C) random noise, 45-cycle filter. (D) sound of striking a match, and the roar of the flame. (E) filing on metal plate. (F) machinery noise.

sharper time resolution, a higher level of signal can be used and thus a wider range of levels can be explored without fuzziness. The noise components in this sample are therefore definitely not a result of the analyzing process.

Sections (C) and (D) represent the song of the wood thrush. Here again a higher level can be used with the wide filter. This song is remarkable for two reasons. The middle portion contains a rapid succession of two distinct notes repeated with great precision. The fainter trace at the top of this section is a second harmonic which was not necessarily generated by the bird but may have been produced in the recording and reproducing processes. The last section of the song however, shows two distinct notes produced at once, namely an extremely rapid high pitched trill accompanied by a steady note of lower pitch. Since these two notes cannot be multiples or submultiples of each other, they are not harmonically related and must therefore have been produced by the bird with two distinct emitting mechanisms. The rapid trill is almost completely blurred by the narrow filter. Incidentally, these bird songs (and others to be shown later) were taken from phonograph records. They included frequencies above the normal 3500-cycle range of the spectrograph. The phonograph records were therefore played at less than their normal speed so that all frequencies were reduced sufficiently to fall within the range. When this method is used, the actual frequencies can be

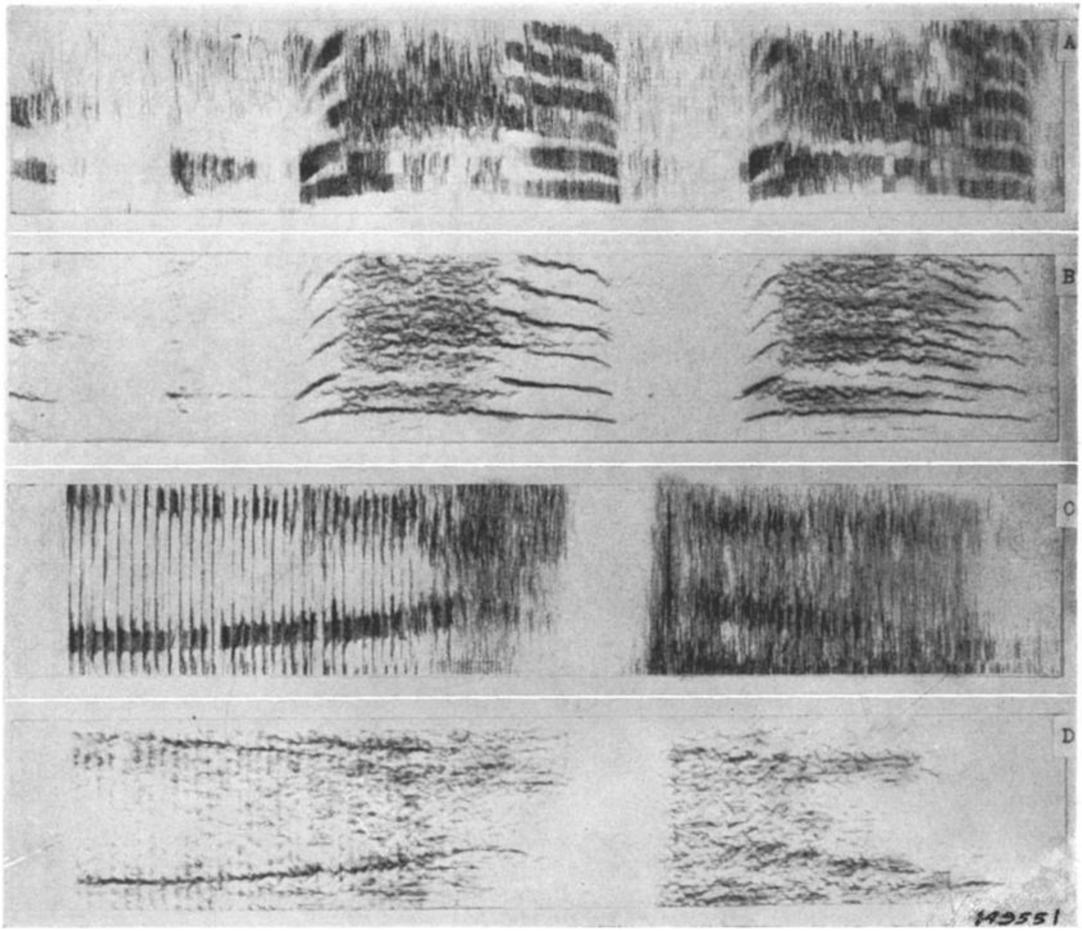

FIG. 24. Comparing the two analyzing filters. Sections (A) and (B) infant crying. (C) and (D) sound of snoring.

computed by dividing the apparent frequencies, as read from the spectrogram, by the speed reduction ratio. The nightingale sample covers about 4500 cycles, and the wood thrush about 6000 cycles.

Figure 26 gives further examples of the behavior of the two filters with signals that vary rapidly in frequency. These are spectrograms of warble tones produced by varying the frequency of a single tone sinusoidally. Sections (A) (narrow filter) and (B) (wide filter) show four different rates of warble from 40 per second to 80 per second. It can be seen that as the frequency of warble is increased, the picture made by the narrow filter tends to exhibit a horizontal structure. In other words, the narrow filter sees this signal as a frequency modulated wave which consists mathematically of side bands around the average frequency. Sections (C) and (D) show higher frequencies of warble and also somewhat greater excursions. The breaking up into discrete components is complete at 100 cycles with the narrow filter and at higher frequencies even the wide filter shows the same effect. Mathematically, both of these representations are correct; the signal may be regarded either as a set of distinct frequency components or as a rapidly varying single component. The spectrograph therefore does not give a false picture. It is simply a matter of choice which interpretation of the signal is more convenient.

Incidentally, all of the illustrations used in this section were made without any form of compression. In exploring an unknown signal it is generally advantageous to make this kind of analysis first because it gives a better picture of

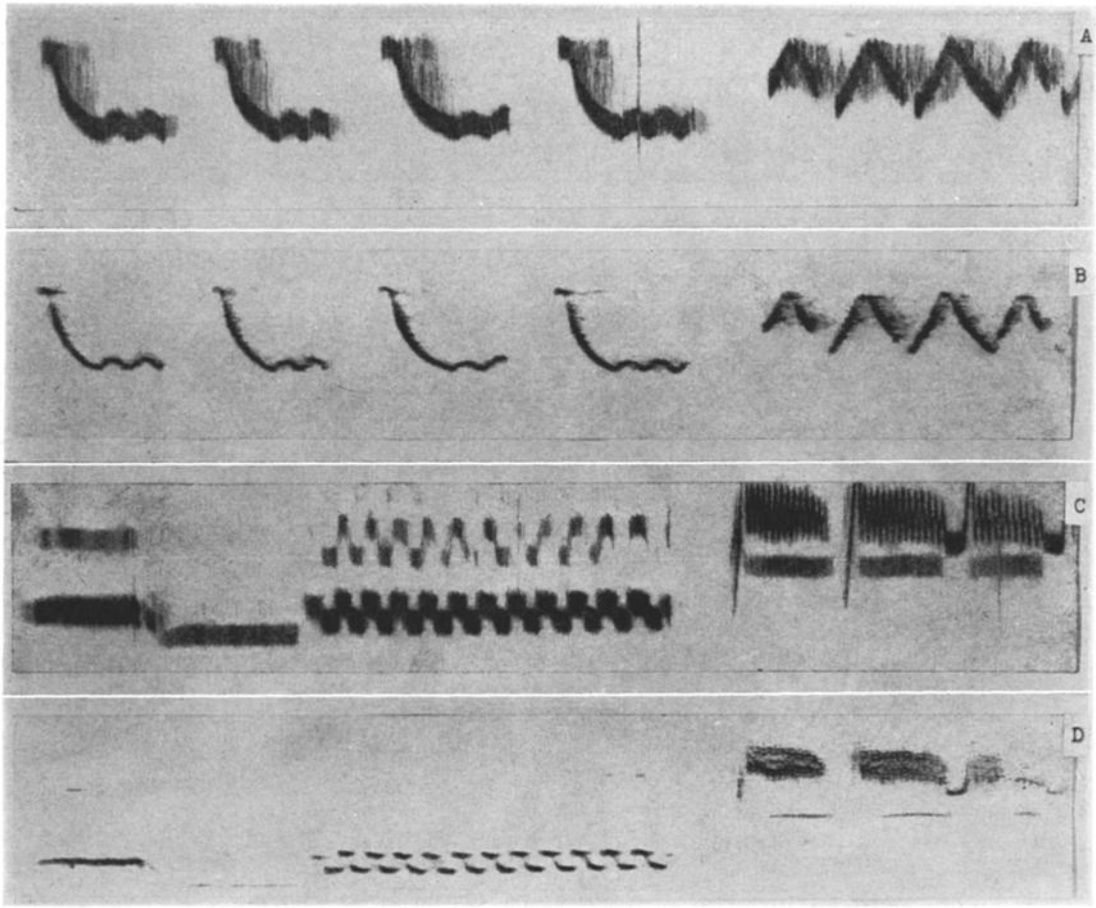

FIG. 25. Two bird songs analyzed with both filters. Sections (A) and (B) nightingale. (C) and (D) wood thrush.

amplitude relations. If the existence of weaker components is suspected, these can be explored with the help of the control circuits if desired.

In Fig. 27 are some spectrograms in which frequency analysis is not the purpose but only the means to an end. Section (A) is the output of an oscillator whose frequency was varied by means of a motor driven condenser. It was desired to determine how the frequency varied with time, and this is immediately apparent from the spectrogram. This type of information could have been obtained by making an oscillogram of the output and determining the frequency *versus* time by counting cycles and making proper interpolations, which would have been an extremely laborious procedure particularly if the determination of slight irregularities were important.

Section (B) portrays the output of a sweep frequency generator showing that the sweep is substantially linear and the return substantially instantaneous. Section *C* illustrates an experiment in which it was desired to determine the manner of acceleration of a phonograph record when released on a moving turntable. The record contained a single frequency tone and a spectrogram of this tone during release proved that the frequency varied linearly from the time the record was released to the time it reached full speed. The duration of the acceleration period can be measured directly on the time scale.

Incidentally, the spectrograms in Fig. 27 cover an 11,000-cycle range instead of the normal 3500 cycles. This was accomplished by recording the signal at the high speed normally used for reproducing. With this arrangement the time covered by the spectrogram is reduced by a factor of three and the frequency scale is multiplied

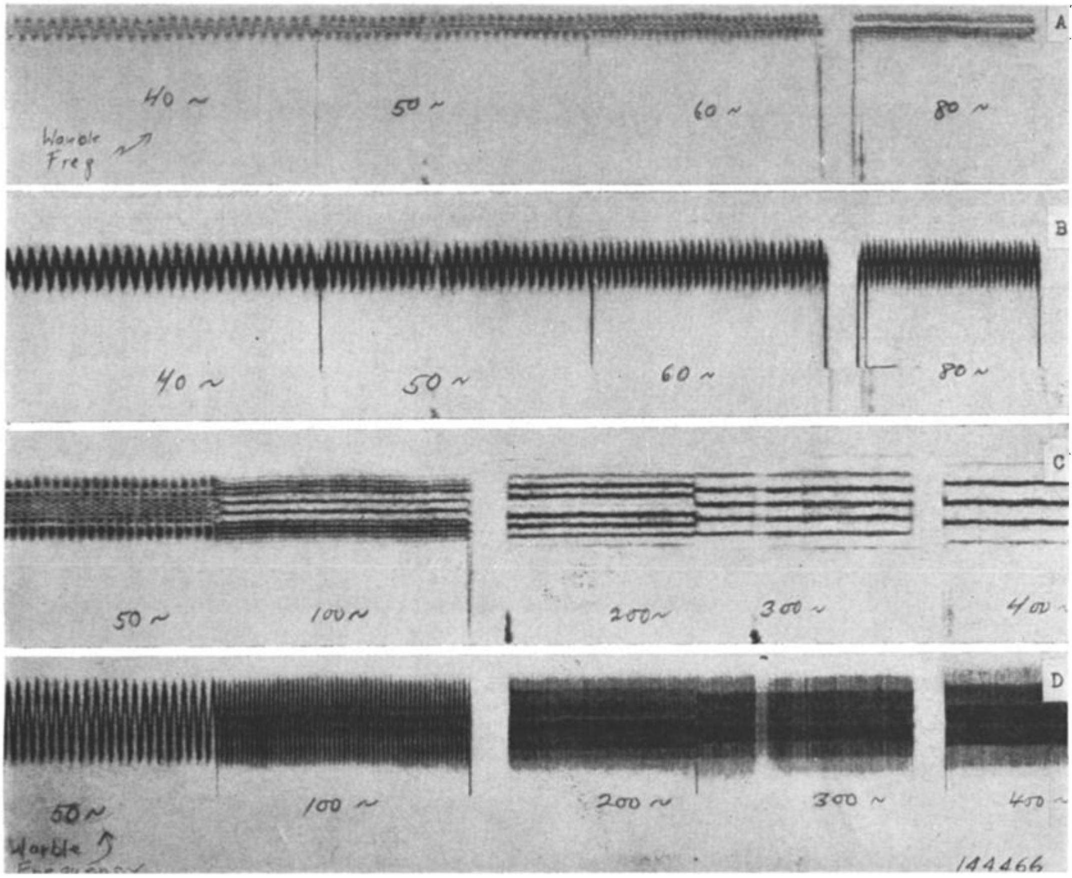

FIG. 26. Tones warbled at various rates. (A) and (C) narrow filter. (B) and (D) wide filter. When the warble rate is comparable to the filter width the warble breaks up into discrete sideband frequencies.

by the same factor. The effective filter width, however, is also multiplied by the same factor, but the apparent width, that is, the width of the trace in the spectrogram, remains the same.

Figure 28 illustrates another application in which the determination of frequency variation rather than analysis was the objective. These spectrograms portray the output of an oscillator whose frequency was governed by a fluctuating d.c. voltage on its grid. The instantaneous frequency can easily be determined at any point with a suitable frequency scale. The fuzziness which is apparent in certain sections results from very rapid amplitude or frequency modulation superposed on the lower frequency modulation. Here the spectrograph takes the place of a string oscillograph, and it avoids the necessity for a d.c. amplifier with sufficiently high output current to drive the low impedance string.

Figure 29 presents another frequency modulation series. In this case a complex wave having both odd and even harmonics was varied in frequency at several rates. The high and low points are marked on the spectrograms. In the uppermost sample the frequency became so low that even the narrow filter resolved the beats. In all these spectrograms "harmonics" will be noted which appear to slope in the opposite direction from the real frequency variation. The effect is due to the fact that the filter, with a rather slow response, is scanning across components at such a rate as to generate a kind of interference pattern. Without attempting to go into the matter thoroughly here, the effect might be compared to the phenomenon of the spokes of a wheel apparently turning backwards in motion pictures. Spurious components like those in Fig. 29 are sometimes visible in analyses

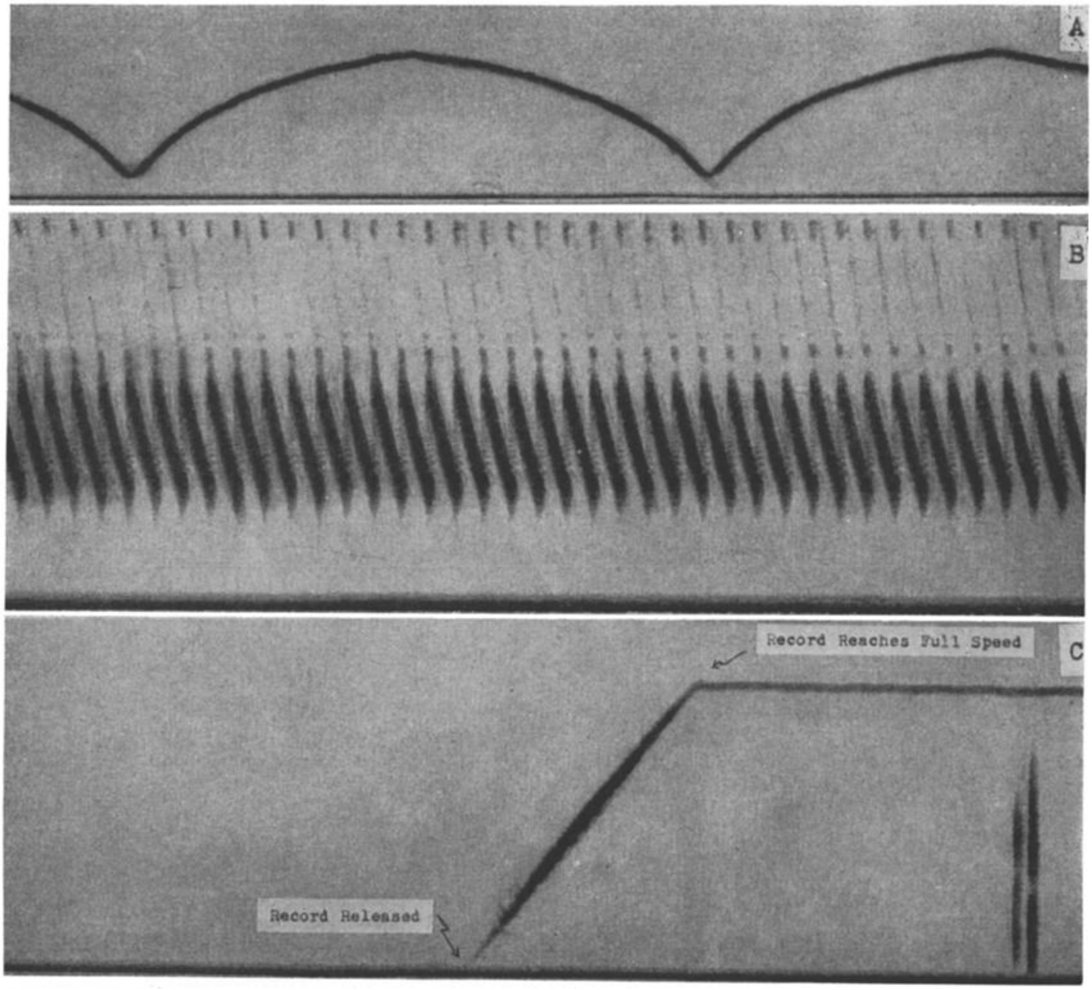

FIG. 27. In these illustrations spectrograms were used to determine frequency *vs.* time rather than energy distribution. Section (A) the output of a motor-driven warbler oscillator. (B) a time base wave. (C) determination of the mode and speed of acceleration of a phonograph disk when released on a moving turntable.

of speech. A good example of this effect can be seen in Fig. 22.

Figure 30 illustrates an interesting application of the spectrograph which should prove very useful. Section (A) shows speech after it has passed through a long loaded line. The line had a cut-off at about 1800 cycles as may be seen by the absence of speech components above that frequency. In addition, however, there is a curious curvature to each syllable. This same curvature can be seen in the small section at the right which represents a click sent over the same line. Ordinarily a click, as mentioned previously, appears as a straight line across the frequency range. The curvature here produced is due to the fact that different frequencies were delayed by different amounts in transmission over the line. Section (B) shows the same material after a delay correction network had been added to the line. Here the speech and the click are nearly normal. In Section (C) the speech was transmitted through the network alone, resulting in the opposite kind of curvature. This illustrates that the spectrograph may be used to investigate the delay characteristics of filters or other networks which are difficult and sometimes practically impossible to handle mathematically when they include dissipation. These characteristics can be determined by simply measuring the de-

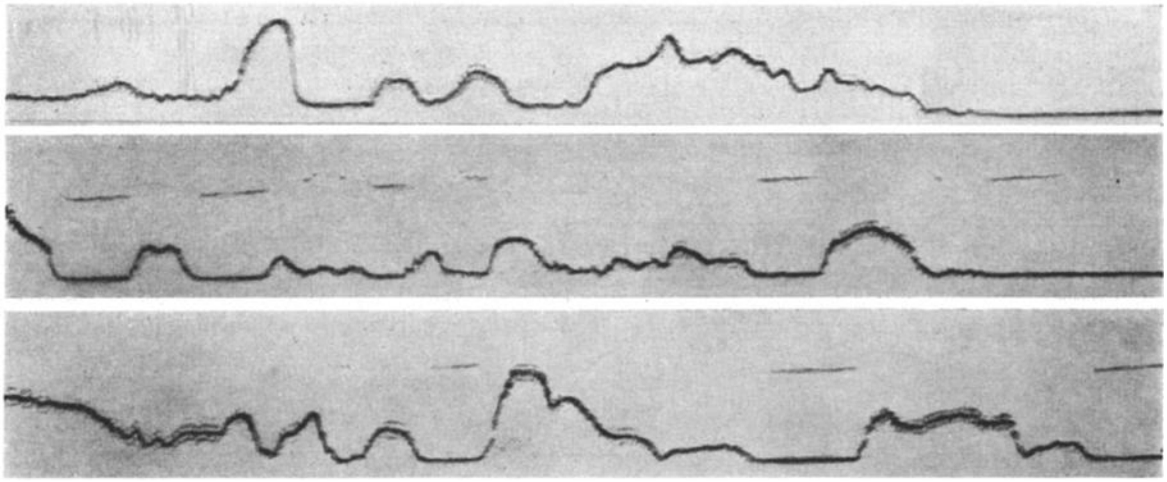

FIG. 28. Another application where spectrograms were used to record frequency *vs.* time. This shows the output of an oscillator whose frequency was governed by a d.c. voltage. The spectrograph here performed the function of a d.c oscillograph.

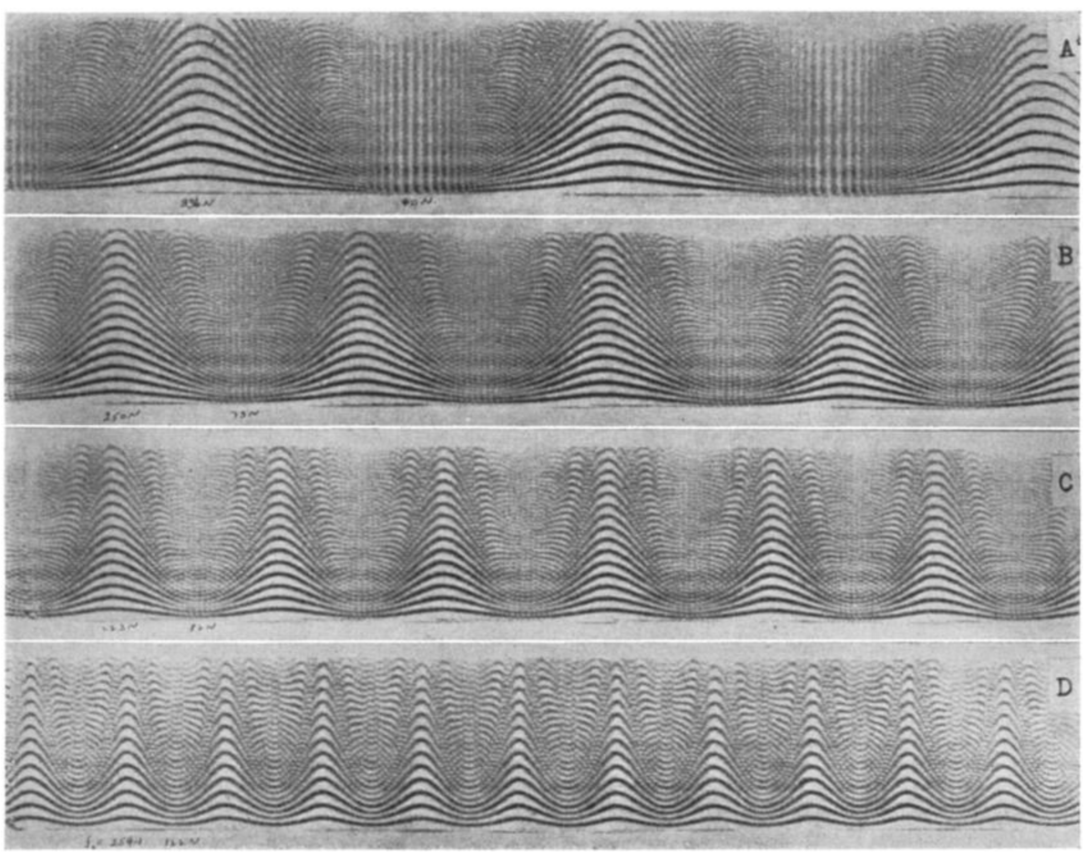

FIG. 29. Here a complex wave rich in harmonics is varied in frequency at several rates. The maximum and minimum frequencies are marked on the margins. In Section (A) the frequency becomes so low that the narrow filter shows vertical striations due to beats. In all the samples, spurious components may be seen which slope in the opposite direction from the true frequency change.

lay *versus* frequency for clicks sent through the networks.

This section will be concluded by presenting a variety of sounds with a wide variety of patterns. First in Fig. 31 are some additional bird songs. Ornithologists have long been trying to analyze and record the songs of various birds accurately. Even after the songs have been slowed down by recording and reproducing at different speeds, it is still difficult for the ear to follow the rapid changes in frequency. Even if the ear were adequate there still remains the problem of a suitable notation for indicating the complex time and frequency relations. With the spectrograph these analyses can be made objectively, and the results recorded unequivocally.

In Fig. 32 are a few additional voice sounds. They require no particular comment except to note that the song in Section (A) is that of a very high pitched voice which is resolved with the wide filter. The amplitude and pitch variations in the vibrato are clearly shown. Presumably the spectrograph will have application in voice training.

Figure 33 contains a variety of animal sounds and Fig. 34 some miscellaneous familiar sounds. The various kinds of components occurring in these sounds have been pointed out in the previous discussion.

VI. SOME VARIABLES

As already indicated above, the general plan of the spectrograph is capable of very wide variations depending upon the application to which it is to be put. For instance, the time scale can be made extremely short for investigating transient phenomena or the like, or it can be made very long for making long spectrograms or for plotting the course of very slowly varying phenomena. The frequency range is also very flexible. Extremely low frequency sounds could be handled by recording very slowly and using a high speed-up ratio. Extremely high frequencies could be handled by recording at high speeds or by the use of modulating processes to bring the high frequencies down to where they could be handled in the normal way.

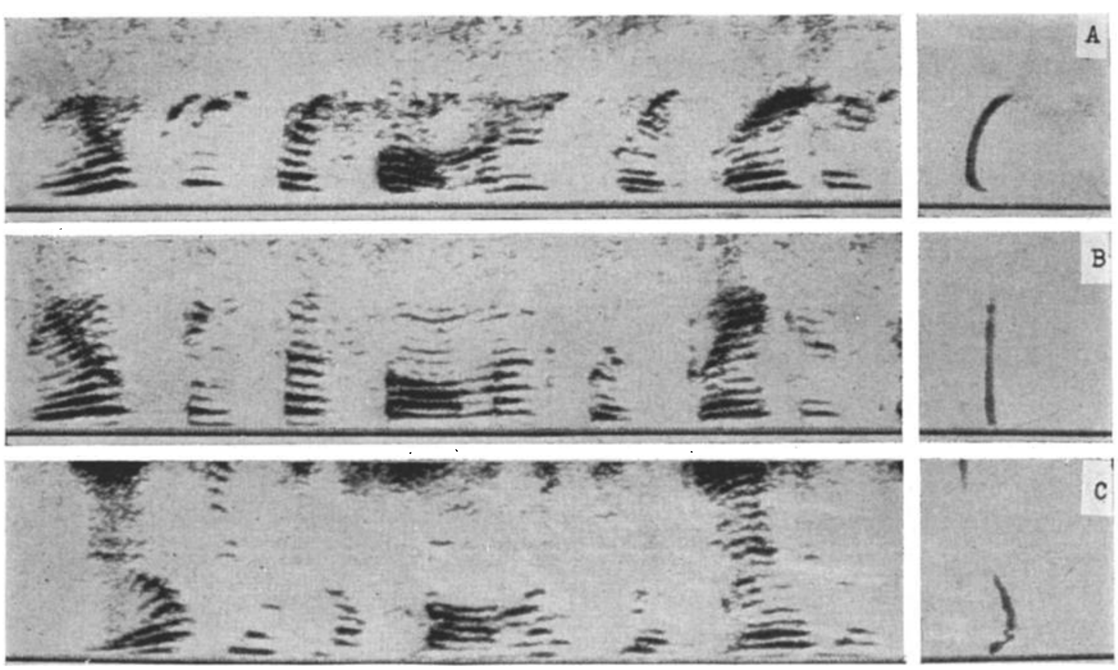

FIG. 30. Illustrating transmission delay in a long loaded line. In Section (A) speech and a click show a cut-off at 1800 cycles, and a curvature which results from the fact that different frequencies were transmitted with different speeds. In Section (B) a delay correction network has been added, restoring the speech and the click to normal. Section (C) shows transmission through the network alone, with the opposite curvature.

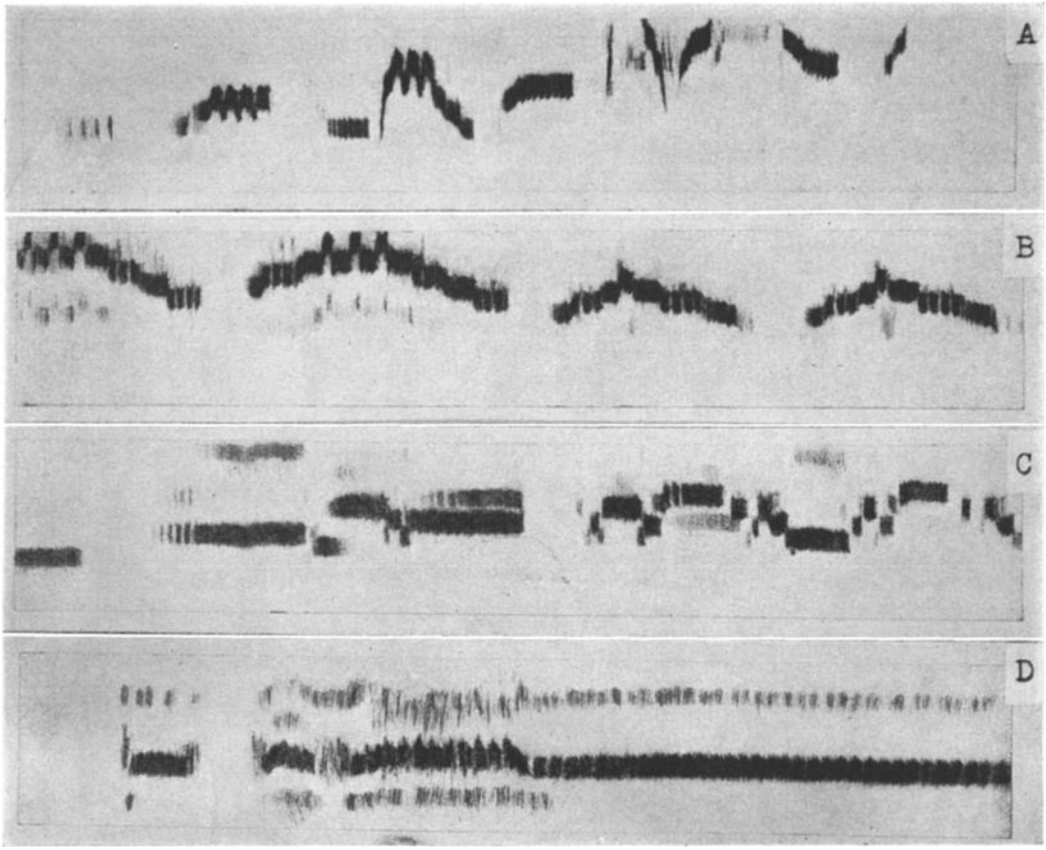

FIG. 31. Some additional bird songs. (A) olive backed thrush. (B) veery. (C) hermit thrush. (D) rooster.

In the present paper all the spectrograms have a linear frequency scale. For some applications, a logarithmic frequency scale would appear to be more logical. In music for instance, the notes are separated by logarithmic intervals. Also, the ear has a natural scale, as shown by differential pitch perception, which is more nearly logarithmic than linear. It might be advantageous to use scales in the spectrograms which approximate those occurring in the ear so that the pictures might be made to look the way they sound. In this connection, the analyzing filters might also be made to simulate more closely the operation of the ear. For instance, a narrow filter might be used in analyzing the low frequencies, with the filter gradually widening towards the high frequencies. The shape of the filter characteristic is also important. In the present spectrograph the narrow filter has extremely steep attenuation characteristics. The wide filter is also very steep with a relatively flat pass band. Different representations can be obtained with filters having different characteristics.

Figure 35 illustrates a few experiments along this line. Sections (A) and (B) were made with the wide and narrow filter, respectively. The subject matter consists of a series of sharp clicks and at the right a calibration tone consisting of all the odd harmonics of 60 cycles. Section (C) was made with a filter which is just about as wide as the one used in Section (B). However, the sides do not slope quite so steeply. It will be noted that in Section (B), the clicks are fuzzy on both the leading and trailing edges. In Section (C), the leading edge is much sharper. In Section (D), the filter is actually narrower than (B) and still retains a sharper leading edge than (B). In Section (E), the pass band is about the same as (B) but the sides slope much more gradually. Here the total width of the clicks is about the

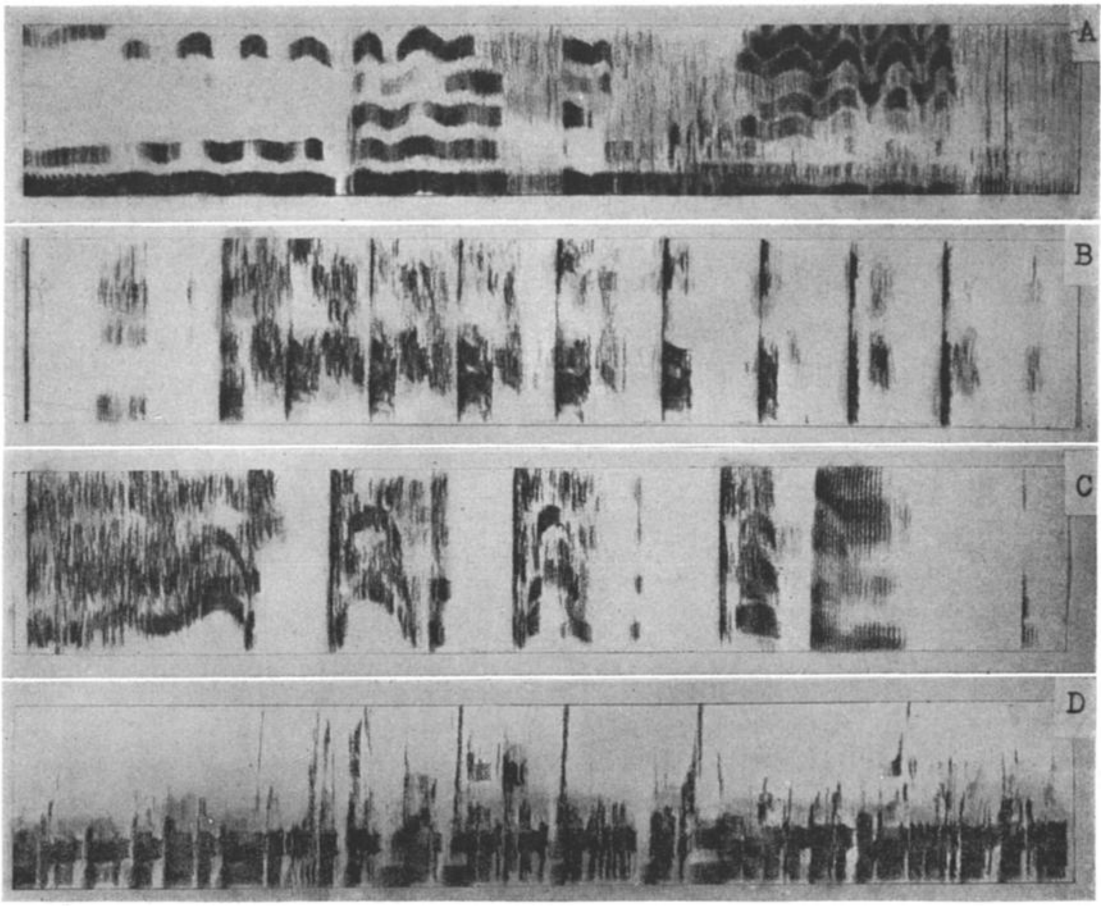

FIG. 32. Some additional voice sounds. (A) trained soprano. (B) laugh. (C) cough. (D) gargling with water.

same as in (B), but the leading edges are about as sharp as with the 300-cycle filter. Obviously, this is a large subject and these illustrations merely indicate the possibilities.

The time requisite for analysis is also subject to wide variation. Time can be saved by reducing the number of lines or by raising the speed-up ratio. If this is carried to sufficiently high frequencies, the patterns might be made substantially instantaneously on a cathode-ray tube.

Various types of signal recording mediums might be used. The magnetic tape seemed desirable because it can be used over and over again. Magnetic recording on wire might be used instead of tape with the advantage that the analyzed samples could be stored. Film recording techniques might be used although this seems rather slow and expensive. On the whole some of the newer types of magnetic recording look most promising for spectrograph applications.

Similarly, various types of facsimile recording mediums can be used. The particular medium used in the present spectrograph is paper with the trade name Teledeltos. This has the advantage of requiring no processing and is permanent. There are other types of facsimile papers in commercial use, including some which are saturated with a chemical which changes color when current passes through it. These require much less power to mark them and would be adaptable to higher speeds. For a wide range of density, photographic film appears to be best. It has, of course, the disadvantage of requiring processing and is rather expensive.

It might be mentioned that all the variables

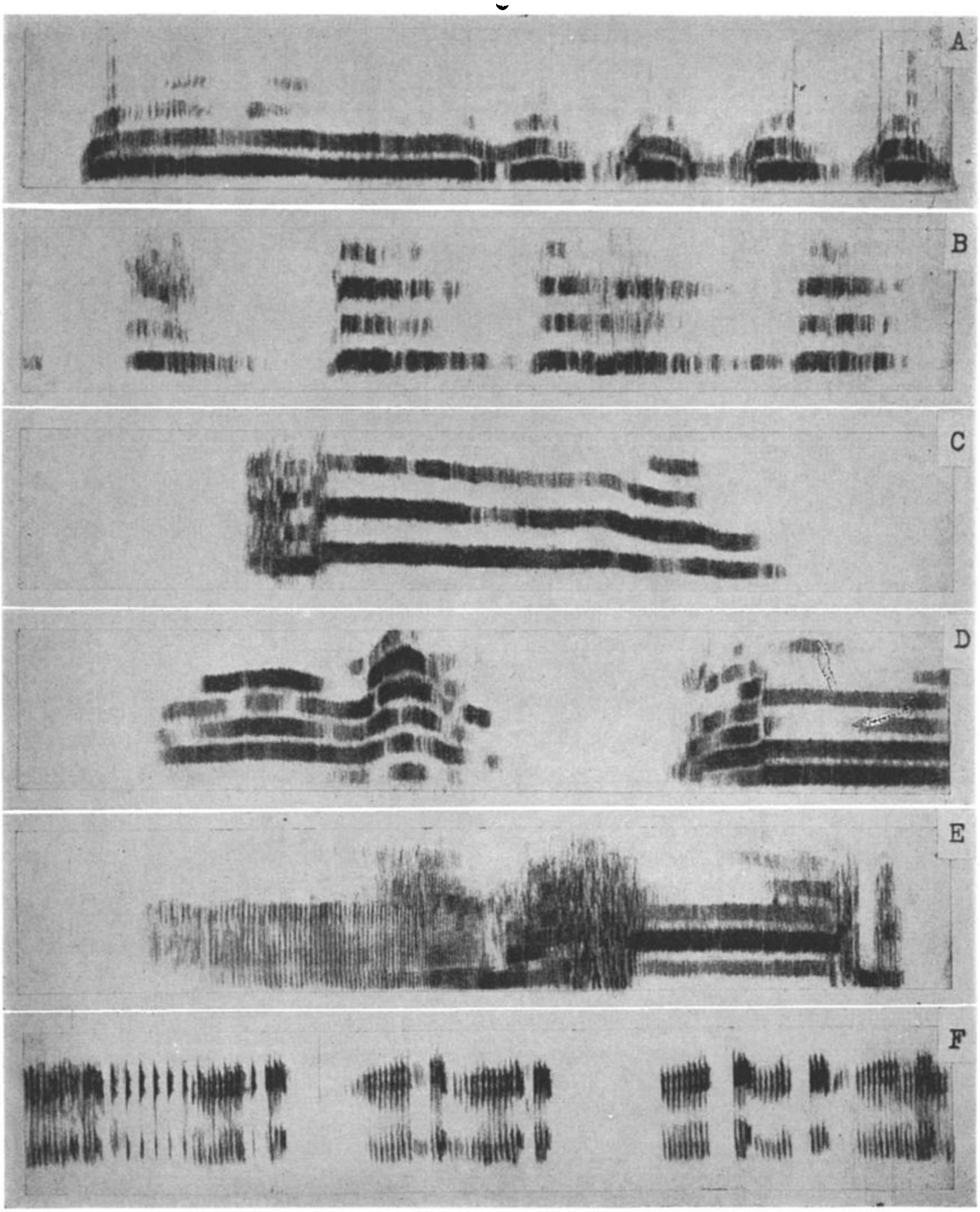

FIG. 33. Some animal sounds. (A) Newfoundland dog. (B) small dog. (C) and (D) wolf. (E) cow. (F) frogs.

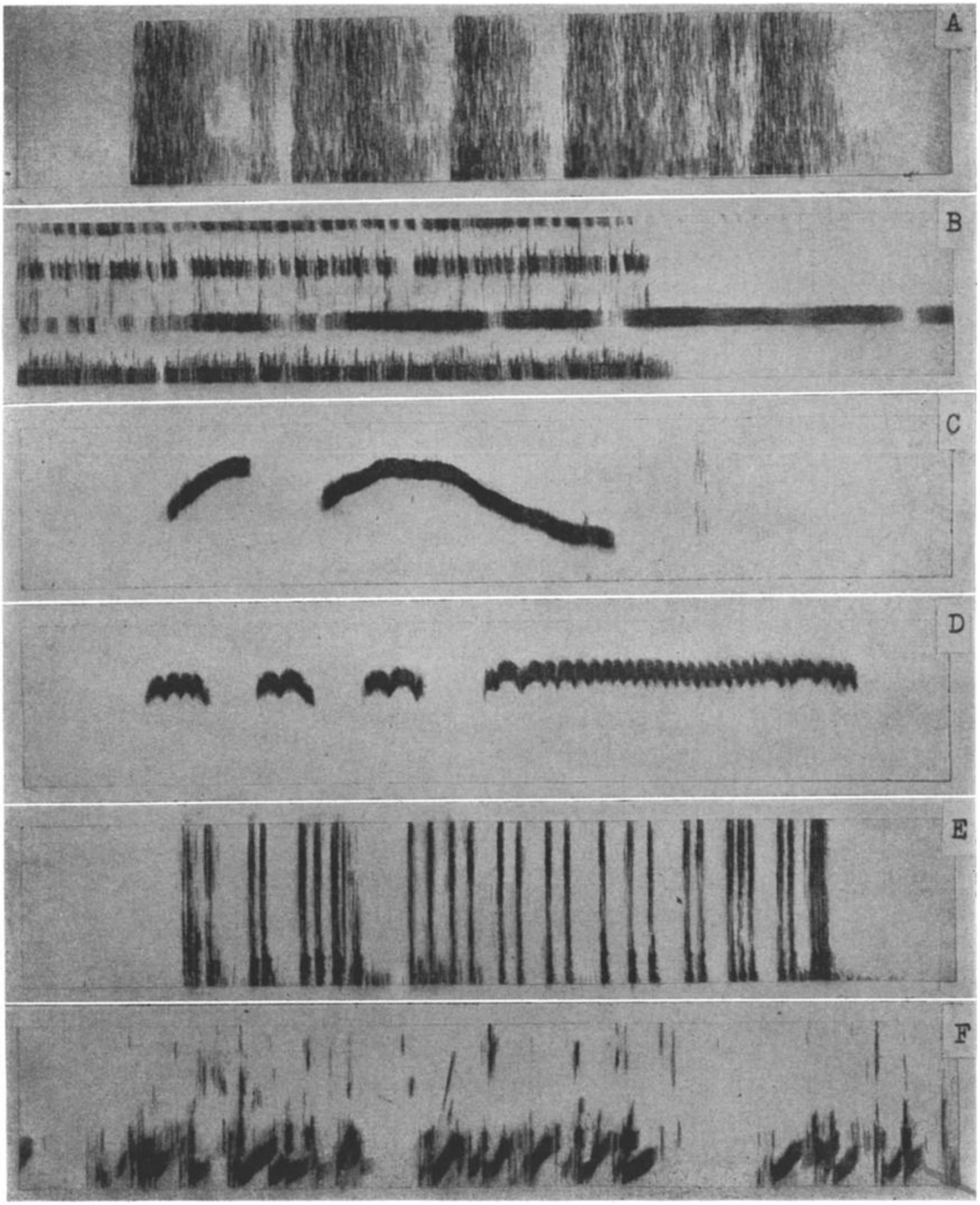

FIG. 34. Some familiar sounds. (A) snare drum. (B) telephone bell. (C) man whistling. (D) police whistle. (E) riffling cards. (F) bubbles blown through water.

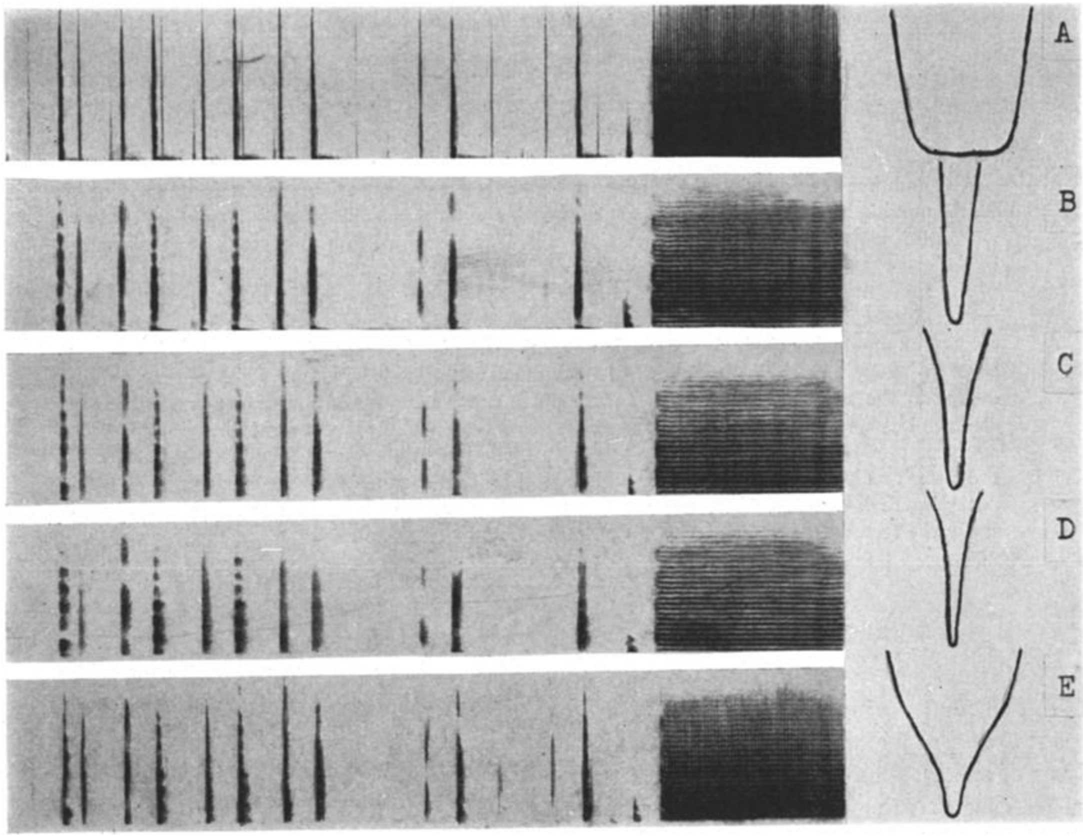

FIG. 35. Illustrating the effect of different filter response characteristics. (A) and (B) the "normal" 300- and 45-cycle filters, respectively. (C), (D), (E) narrow filters with different characteristics. Filters with gradual slopes give sharper pictures of clicks than steep filters.

mentioned in this section have been tried at least in exploratory experiments.

One very important variable should be noted in conclusion. This is the matter of amplitude representation. In the spectrograms presented above, the time and frequency scales are reproducible and subject to precise measurement. The density or amplitude scale, however, is nonlinear and not reproducible. The total range of photometric density is not very great, and furthermore there are many variables such as ambient temperature and humidity, stylus pressure, stylus width, the condition of the drum surface, and so forth, which affect the density scale. In some applications, it is necessary to know the amplitudes of the components quantitatively. Various methods have therefore been devised for representing amplitudes in spectrograms in such a way that they can be interpreted quantitatively. It is planned to cover this subject in a subsequent paper.