AD NUMBER
ADA800206
CLASSIFICATION CHANGES
TO: unclassified
FROM: secret
LIMITATION CHANGES
TO: Approved for public release, distribution unlimited
FROM: Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; OCT 1944. Other requests shall be referred to Office of Scientific Research and Development, NRDC, Div. 13, Washington, DC.
AUTHORITY
Secretary of Defense memo dtd 2 Aug 1960; Secretary of Defense memo dtd 2 Aug 1960

The U.S. GOVERNMENT

IS ABSOLVED

FROM ANY LITIGATION WHICH MAY ENSUE FROM THE CONTRACTORS IN-FRINGING ON THE FOREIGN PATENT RIGHTS WHICH MAY BE INVOLVED.

WRIGHT FIELD, DAYTON, OHIO

PAGES _____

ARE

MISSING

IN

ORIGINAL

DOCUMENT

SECRET

OSRD 573A

302

NATIONAL DEFENSE RESEARCH COMMITTEE

OFFICE OF SCIENTIFIC RESEARCH AND DEVELOPMENT

DIVISION 13 SECTION 3

OEMsr - 435

FINAL REPORT

ON

PROJECT C - 43

Continuation of Decoding Speech Codes

PART I — SPEECH PRIVACY SYSTEMS —

INTERCEPTION, DIAGNOSIS, DECODING, EVALUATION

O.S.R.D. NO.____ SECTION NO._____ COPY NO. 42 DATE Oct. 12, 1944

BELL TELEPHONE LABORATORIES, INCORPORATED

NEW YORK, N. Y.

SECRET

Communications Division
National Defense Research Committee
of the
Office of Scientific Research and Development
Division 13 Section 8

FINAL REPORT
ON
PROJECT C-43
Continuation of Decoding Speech Codes

PART I - SPEECH PRIVACY SYSTEMS - INTERCEPTION, DIAGNOSIS, DECODING, EVALUATION

October 12, 1944

SECRET

Contract No. : OEMsr-435

Contractor: Western Electric Company Inc.
120 Broadway, New York 5, N. Y.

Project Supervisor: C. H. G. GRAY
Technical Report Prepared by: W. KOENIG

BELL TELEPHONE LABORATORIES, INC.
463 West St., New York 14, N.Y.

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont'd.)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Cont'd.)

HISTORY OF PROJECT

This report covers work carried out for the National Defense Research Committee under Contract No. OEMsr-435 with the Western Electric Company, Inc. by the Bell Telephone Laboratories, Inc.

Early in October, 1940 there was set up in the Communications Division of N.D.R.C. a subcommittee on Speech Secrecy. This group was to consider both the scrambling and unscrambling of telephone signals. It was soon recognized by them that the decoding problem was of primary importance both as a means for evaluating privacy systems for possible use by the Services and for decoding possible enemy signals. Realizing that the ear has very limited capabilities for analyzing scrambled speech Mr. R. K. Potter invented the sound spectrograph to provide speech patterns which could be interpreted by the eye.

Early in 1941 a rough laboratory model of the sound spectrograph became available in Bell Telephone Laboratories Inc. Through Dr. O. E. Buckley, chairman of the Speech Secrecy Section of N.D.R.C. arrangements were made for a demonstration of the sound spectrograph to various N.D.R.C. representatives including Dr. V. Bush.

As a result of the above described demonstration and subsequent Committee action Project C-32, the forerunner of Project C-43, was organized in the fall of 1941 with the immediate objective of producing a sound spectrograph in such form that it would be useful for diagnosing and decoding speech scrambling systems. In Project C-32, "Privacy Cracking", a finished model of the sound spectrograph was constructed and its application to decoding work was successfully demonstrated to representatives of the Army, Navy and N.D.R.C.

Upon the termination of C-32 on February 1, 1942, it was decided that the work initiated under that project should be continued. Accordingly Project C-43, "Continuation of Decoding Speech Codes", was authorized for one year, effective February 2, 1942. The project anticipated some routine decoding, the production of duplicate equipment to be used by the Army and Navy intelligence services and further studies of decoding tools and methods. At that time the Army and Navy military officers were relying almost entirely upon this project to furnish the above services until they could be provided with suitable equipment and could obtain trained personnel. Based on the needs of the military this project was thrice extended.

Under the guidance of the Subsection on Speech Secrecy, Section 13.3 of N.D.R.C., the emphasis was placed at any given time on what was deemed to be most urgent. This is reflected in the subject matter of the Preliminary Reports which were issued from time to time and which form the appendix to this report. In addition to the specific investigations covered by these Preliminary Reports much work was carried on as the basis for more general coverage of the field of interception, diagnosis, decoding and evaluation of speech privacy systems.

In addition to the general studies mentioned above decoding equipment was developed and models furnished to the Army and Navy. This decoding equipment included (1) two models of the sound spectrograph, (2) a variable area pattern machine and (3) equipment for decoding two new enemy privacy systems intercepted by the project personnel at Point Reyes, California. In each case Army and Navy personnel were instructed in the operation and maintenance of these equipments.

Intercept activities of the project personnel included (1) the study of recordings submitted early in the project by the Federal Communications Commission, (2) exploratory work at the Bell Telephone Laboratories experimental radio receiving station at Holmdel, New Jersey and (3) exploratory work and routine interception of radio telephone transmissions at the American Telephone and Telegraph Co. radio receiving station at Point Reyes, California. Reports of the results of the above studies and recordings of intercepted material were submitted directly to the interested military authorities.

Many speech privacy schemes were submitted through N.D.R.C. during the course of this project. These were studied and evaluated. This work led directly to the continued improvements of the sound spectrograph and the development of supplementary decoding tools and techniques.

As the Army and Navy became able to carry on decoding activities themselves with the aid of equipment and information furnished by

N.D.R.C. as the result of work outlined above, the activity on this project gradually decreased.

The technical report which follows, together with the Appendix which includes all of the Preliminary Reports, covers all phases of the work on this project and constitutes a reference work for future studies of speech privacy systems.

INTRODUCTION TO TECHNICAL REPORT

This report summarizes the results of about three years' experience in diagnosing, decoding, and evaluating speech privacy systems submitted for study on this project by the Army, Navy and N.D.R.C. Some of the results of these studies have been described in a series of Preliminary Reports which were issued from time to time to cover specific studies. A great deal of accumulated experience, however, has never been reported in this manner. This final report, therefore, is intended to make available information, both positive and negative, which would have to be accumulated by another group if they were to embark on a similar project.

The immediate pressure behind these studies was caused, of course, by the War. The material here recorded should therefore be of service if a similar emergency should arise in the future. To keep up with the ever-changing art of communication, these studies should be continued under Government auspices during peacetime.

In contrast to a rather extensive literature on code and cipher systems, on cryptanalysis and cryptography, which apply to telegraph types of communication, very little has been written on speech privacy systems or decoding methods applying to them. Two moderately comprehensive articles have been published. One appeared in the Post Office Engineers Journal, October, 1933. The other appeared in the Brown Boveri Review for December, 1941. The latter has been reproduced and discussed in Preliminary Report No. 5. It covers a number of basic types of scrambling systems, and in addition discloses one that is new.

The present report is intended to cover rather completely speech scrambling methods in which the original speech is transmitted with its parts modified, displaced or interchanged. When more detailed technical information is desired, reference may be made to the Preliminary Reports. These are separately bound and form an appendix to this report. All of this material should be helpful in the development of practical, effective privacy systems, and the evaluation of the security which they afford.

CHAPTER I
INTERCEPTION

Speech privacy systems may be used in connection with radio telephone systems or with wire systems. The unauthorized interception of wire communications in wartime, however, is beyond the scope of the present report. This chapter will therefore be confined to radio interception problems and expands the material in Preliminary Report No. 25. The decoding techniques to be described subsequently, of course, apply to wire as well as radio communications.

1. Types of Radio Systems

Radio telephone systems range in size and complexity from high power point to point stations operating over great distances to the low power, short range sets carried by individual soldiers. The high power systems are usually designed to operate between specific points, using specific assigned frequencies. They are equipped with elaborate fixed antennas, which are usually of the directive type. Privacy equipment associated with such terminals may be as large and complex as desired to achieve virtual secrecy. A major consideration in such systems, of course, which adds to size and complexity, is that the privacy must not degrade the quality of the received speech to any appreciable extent.

On the other hand, any one can intercept these highpower signals at great distances, where he can have a well-equipped centralized decoding laboratory, with no limitation on the size and complexity of the decoding equipment he might bring to bear. This laboratory can be adequately manned by a relatively few highly trained decoding specialists not necessarily members of the armed services.

In contrast with this situation, the low power, short range radio sets used in military operations are severely restricted as to size and weight, and these restrictions also apply to privacy equipment. The smallest privacy set submitted to Project C-43 for study was roughly a 10 inch cube, and was designed for mobile applications like tanks, planes, and command cars. While it is difficult to achieve a high degree of inherent privacy in mobile equipment, it should be noted that the very mobility of such systems adds to the security, because the signals can not generally be picked up at great distances, and whatever equipment an interceptor might use to crack the privacy must also be mobile. Furthermore, the decoding equipment must be operated by military personnel, a large number of whom may be required if the enemy is making extensive use of mobile privacy. Whether it is worth while to attempt to decode a large number of small mobile communications is questionable, as discussed in greater detail in Chapter VII.

Intermediate types of radio systems are used for the higher echelons of command. For such applications, the radio equipment is semimobile. It can be transported in trucks and set up very rapidly, and may have a considerable range. For such applications, a high degree of privacy is required, and a truckload of equipment might be justified, because the enemy could afford to devote considerable time, personnel, and equipment to decoding the kind of messages which would be transmitted over such systems.

2. Intercepted Signal Quality

Since most of this report deals with decoding, the material from this point on will be written from the point of view of the unauthorized rather than the authorized listener. It is first of all desirable to get a good signal, as free as possible from interference. There are several reasons for this. First the process which unscrambles the speech also scrambles any noise such as static which has been superposed on the scrambled signal. This changes the time or frequency distribution of the noise, breaks up harmonic relationships, etc., thereby increasing the interfering effect of the noise. Second, the decoding is apt to be less perfectly accomplished than at the authorized terminals, which tends to make the speech harder to understand. Finally, there are usually language differences which still further add to the difficulty of understanding the message. Conversations can be carried on under extremely unfavorable conditions by people speaking their own language, but noise and poor quality rapidly degrade the intelligibility of a language foreign to the listener.

In this connection it might be noted that it is very desirable to be able to hear both

sides of the conversation without interruptions, in order to follow the context. In the case of the point to point systems, this will in general require two receivers because the two directions are transmitted over separate channels at different frequencies. If the two outputs are mixed for listening or recording, however, it should be kept in mind that the noise on the weaker signal will be superposed on the stronger signal and may seriously degrade it. Putting the two signals on two headphones will improve this situation, because noise in one ear does not seriously affect the intelligibility of a signal in the other ear. This problem does not arise in the case of the smaller radio systems, because these are generally operated on the basis of switching between transmitting and receiving conditions on the same carrier frequency.

Methods of obtaining a good signal are the same for the interceptor as for the intended receiver. A few of the important considerations are listed here; further information on any or all of them can be had from radio reference works. (1) Point to point systems usually employ directive antennas; the intercept station should therefore be located along or near the line of the radio beam. (2) In locating stations to intercept radio transmissions in the HF range, account should be taken of the skip distances of the frequencies involved. Better signals will sometimes be obtained by moving farther away from the transmitter rather than closer. (3) The use of directive antennas, directed towards the transmitter being monitored, will improve the signal to noise ratio by discriminating against noise which is non-directional. These antennas of course should be designed for the frequency and polarization of the signal, and properly coupled to the receiving set. (4) Stronger radio signals will be received if the antennas are located in the open, with no trees or other obstructions in the foreground. This is particularly important in the VHF range. (5) Radio signals increase in intensity as the height of the antenna above the immediate foreground is increased, particularly for VHF transmission. Thus better results are obtained with the antennas located on high masts or on hills overlooking the foreground in the direction from which the signal is arriving. If the signal is in the VHF range and other measures are inadequate, it may even be desirable to consider receiving the signal in an airplane and recording it or retransmitting it for decoding. (6) Noise improvement can generally be obtained by keeping the receiving equipment away from sources of man made noise, such as ignition systems and power lines.

3. Receiving Sets

With regard to the receiving sets, a distinction must be made between the various activities of an intercept station. One important activity is searching for possible enemy transmission channels. The object is to determine all the channels in use, the location of their terminals, the type of business transacted, and, most of all, whether any special form of privacy is used on the channel. Some preliminary searches of this type are described in Preliminary Reports Nos. 2 and 23. If no privacy is used, other than the usual commercial types, it is unlikely that information of military importance is transmitted over the channel, and it may not be necessary to monitor it continuously. If a new privacy system is located, however, it is very likely to be worth monitoring and decoding continuously.

For the searching and scanning activities, the ordinary commercial sets of the "communications" type, equipped with a beat frequency oscillator, will serve very well for all types of transmission. Even the suppressed carrier type can be handled very well provided the signal is fairly strong. It may require continual manual adjustment of the local oscillator, but sufficiently good reception can be obtained to determine the nature of the channel. Cases of extreme spread band transmission can also be handled in this manner.

If a particular channel employing suppressed carrier is determined to be worth monitoring continuously, then a single sideband receiver will give improved reception. These receivers are equipped to amplify the partly suppressed carrier, or supply a new one with great stability, and they may provide as much as 15 db improvement in signal to noise ratio in some cases. They also permit selecting either the upper or the lower sideband of double sideband systems, which may be of advantage in cases where interference occurs on one or the other sideband of such systems. However, these receivers are not suitable for searching.

4. Types of Radio Transmission

A knowledge of the types of radio transmission which may be encountered is very important to the personnel of an intercept station. Experience has shown that without such knowledge, the nature of intercepted signals may be completely misinterpreted. It is possible to mistake certain normal types of transmission for new systems, or conversely to fail to recog-

nize new systems which should be monitored at once.

The commonest type of transmission is the double sideband type in which the carrier is transmitted along with the sidebands, which are usually about 3,000 cycles in width, and are located immediately adjacent to the carrier. These are readily demodulated by the ordinary receiver. This is true even if the carrier is rapidly wobbled, provided the wobble does not cover too great a frequency range. Such wobbles are sometimes used in combination with simple inversion, to prevent reinverting with a locally supplied carrier at the edge of one sideband.

In the so-called spread band system, some or all of the sidebands are displaced from the carrier. Demodulated signals of this type will cover an audio frequency range greater than 3,000 cycles, usually as high as 6,000 cycles. It is essential, therefore, that the receiver be capable of handling such a band. To obtain the intelligence, the signals must be further demodulated as described subsequently (B1 in Table I).

In the ordinary transmissions described above, the carrier level is high compared to the speech sidebands. In order to avoid loading up the transmitter with carrier, and thereby permit radiating a higher sideband level, many channels operate on the "suppressed carrier" basis. In this system the carrier is either eliminated completely, or transmitted with greatly reduced level. To demodulate such signals properly, the weak carrier must first be greatly amplified, or a new one supplied locally. If this is not done the signals will demodulate themselves around whichever component in the sideband happens to be predominant, producing thoroughly scrambled speech which can thereafter not be restored. This condition can be recognized by its characteristic sound to the ear, together with wide syllabic fluctuations of the meter which ordinarily indicates the carrier level.

With suppressed carrier systems, usually only one of the speech sidebands is transmitted. However, a second sideband, transmitting a second speech channel, is sometimes added, usually displaced from the carrier by about 3,000 cycles, to avoid crosstalk between the channels. This is called "twin channel" operation, and gives on demodulation an audio signal covering about 6 kc. The two channels must be separated and placed in their normal positions by the methods previously cited under spread band systems.

The above systems are the main types of radio transmission used commercially with amplitude modulation. In addition, in the VHF range and above, there are frequency modulation systems, and also pulse modulation systems, both of which require receivers specially designed to handle their particular types of signals. This is too large a subject to cover here, and reference must again be made to the radio literature.

Finally it should be mentioned that in addition to speech a great deal of telegraph transmission will be found. There are several types of telegraph signals, including hand keyed, such as Morse code, or machine keyed such as Boehme and teletype. Any of these types may be transmitted by keying the carrier, or by keying a tone modulated on the carrier. The marks and spaces may be represented by changing the amplitude (on - off) or by changing the frequency (two-tone). Finally, since telegraph requires a much smaller band than speech, it is often operated on a multichannel basis, that is, a voice channel will be divided into a number of telegraph channels. In addition, there are facsimile transmission systems, which also may be operated on an AM or FM basis. If a new signal is encountered whose nature is in doubt, these possibilities should be kept in mind for further investigation when the need arises.

5. Recording

The same considerations, discussed in section 2 above, which make it desirable to obtain a good intercepted signal, apply also to recording and reproducing scrambled speech. In addition to the requirements as to quality and noise, there is an even more serious one concerning speed regulation. In general, systems designed for a high degree of privacy require a high degree of synchronization, and in many cases ordinary recording methods are not good enough, not only in long time average speed regulation, but in the steadiness of the instantaneous speed. In the case of some of the systems described in Chapter III, for instance, the requirements are so severe that even the best commercial recorders will not meet them.

The best solution of this problem is to decode before recording. This will be possible in many cases, although it may sometimes entail the loss of parts of the message while adjustments are being made or the code is being determined. It happens that some of the systems described in Chapter III which impose the severest requirements on speed regulation, (B3 in Table I) can be handled in this way. When

this method is feasible, even poor quality recorders, such as those designed to record a great deal of material in a small area, may be good enough.

In some of the systems to be described, it will not be possible to decode before recording. It happens, however, that in the case of the only known system for which this is true (F3 in Table I), the requirements as to quality and speed can easily be met by good commercial type recordings.

The matter of convenience or ease of use of the reproducing system is very important in decoding work. In this respect also, the requirements are different for different privacy systems. The recording systems using the embossing process, for instance, are convenient because they produce no thread, and they require little attention. However, they all suffer from poor tracking during reproduction, which can be exceedingly burdensome, especially where the material must be reproduced many times over. Recording magnetically on wire is attractive from the standpoint of convenience and also quality, but back-tracking is very time-consuming and laborious.

The best solution, at the present writing, appears to be disk recording on acetate, with a machine capable of recording at various speeds. Low speeds can be used where quality need not be too good, and a long record is desired. Higher speeds can be used where better quality is needed. Such recording systems are commercially available.

6. Decoding Tools

In addition to the facilities discussed above, an intercept station, if it is to be prepared to diagnose and decode intercepted enemy signals, must be equipped with a considerable variety of special tools. These should include, of course, such well-known devices as oscillographs and oscilloscopes, amplifiers, oscillators, modulators, rectifiers, fixed and variable filters, and a supply of components for constructing special circuits that may be required. Some of the less well-known devices, whose nature and usefulness will be made clear in subsequent chapters, include magnetic tape or wire recording and reproducing equipment in the form of loops with multiple pickups, commutators for sweep or timing circuits, variable speed drive mechanisms, channel shifters, the variable area pattern machine, and the sound spectrograph. There should also be models of the more important types of existing speech privacy systems. Finally, and perhaps most important of all, there should be stationed at the intercept location a group of highly trained technicians, who should be thoroughly familiar with radio transmission problems, radio facilities, cryptanalytic procedures, and diagnosing and decoding methods. If these technicians are not conversant with the language encountered in intercepted communications, interpreters should be continuously available.

Even with all of the special tools and personnel, decoding in many instances will be a difficult problem, and patience and painstaking effort will be required to obtain useful information from scrambled speech. Unless the needs have been anticipated the enemy may have secret communication for a considerable period of time as a direct result of unpreparedness.

CHAPTER II
THE SOUND SPECTROGRAPH

This chapter is devoted to the sound spectrograph -- its history, method of operation, and capabilities. The sound spectrograph analyzes speech in terms of its three basic dimensions, frequency, amplitude and time, and portrays the analysis in the form of spectrograms. These are helpful in understanding the complexities of speech, and what various scrambling methods do to speech to make it unintelligible.

1. History

In March 1941 an early laboratory model of the sound spectrograph was demonstrated to Dr. V. Bush as an instrument that with further development might be useful in studies of telephone privacy. It was appreciated at that time that the need might arise for intercepting communications in scrambled speech and decoding them. It was also appreciated that new scrambling systems might be encountered and that means would be needed for diagnosing such systems. For such a purpose the unaided ear has very limited capabilities. Such things as oscillograms, which show the wave form, also provide little in the way of clues as to the mechanism by which the wave form was changed. Project C-32, the forerunner of C-43, was therefore organized in the Fall of 1941, and its immediate objective was to produce a sound spectrograph in such a form that it would be useful for diagnosing and decoding speech scrambling systems.

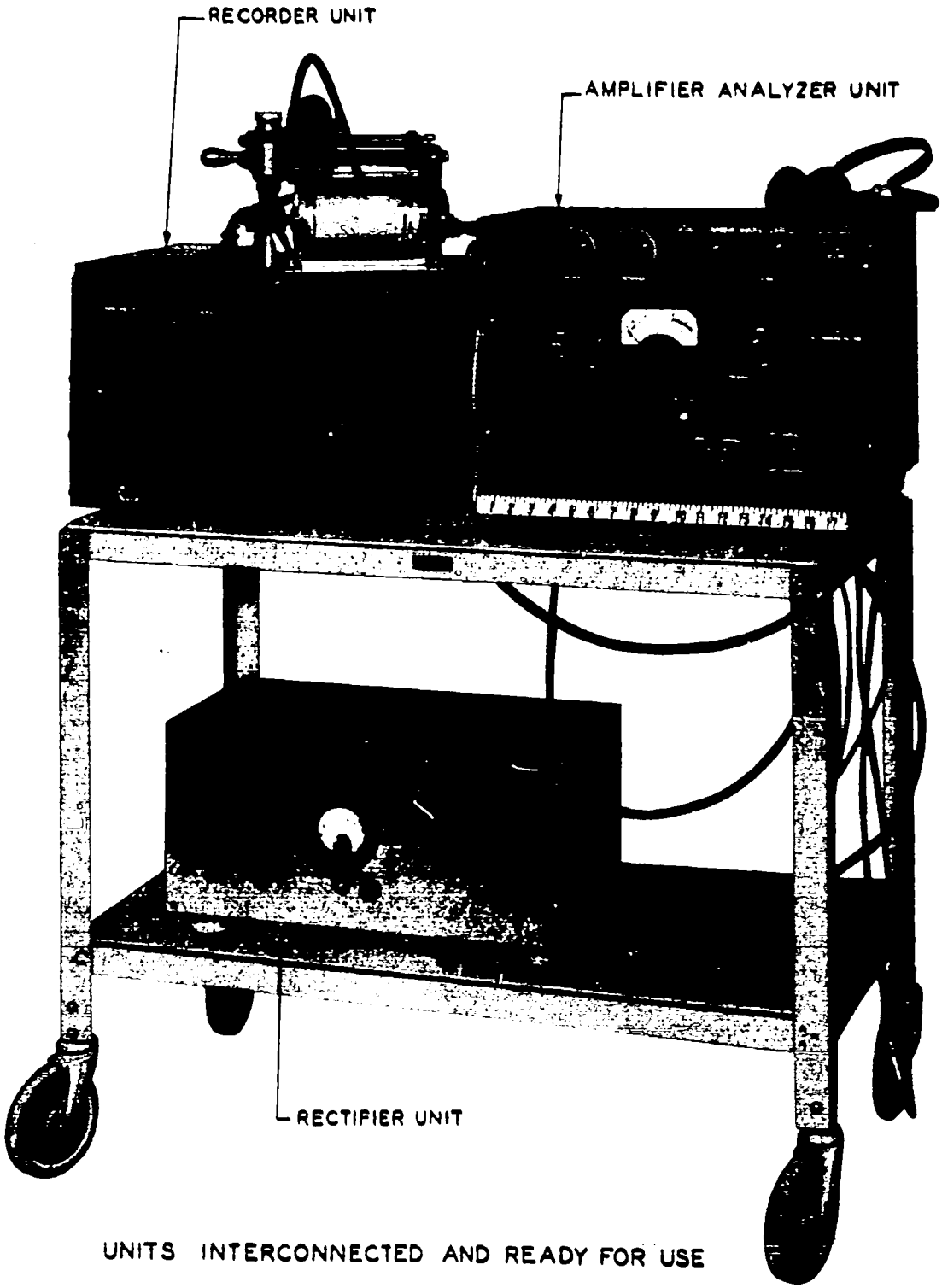
About a month before the attack on Pearl Harbor, patterns that could be used for decoding work were being produced with a breadboard model, and the first finished model of the spectrograph was available by the end of that year. Additional models of the spectrograph have since been built for the use of the armed services, incorporating improvements in operation and in ruggedness. The most recent model is shown in figure 1. The spectrograph has been used in studies of various privacy systems submitted by the Army, Navy and N.D.R.C. for the purpose of evaluating the degree of security which they afforded. In the course of these studies it became evident that improvements in the spectrograph would be useful in this work. Accordingly a calibrating circuit was built into the spectrograph and control circuits were added in the form of an applique unit.

2. Operation

A schematic diagram of the sound spectrograph is given in figure 2, together with a description of the method of operation. There is produced by the operations described in the illustration, a pattern which shows by its light and dark areas how the intensity in the signal varies as a function of both time and frequency. It is the fact that both time and frequency variations are simultaneously displayed which makes spectrograms so valuable for decoding work.

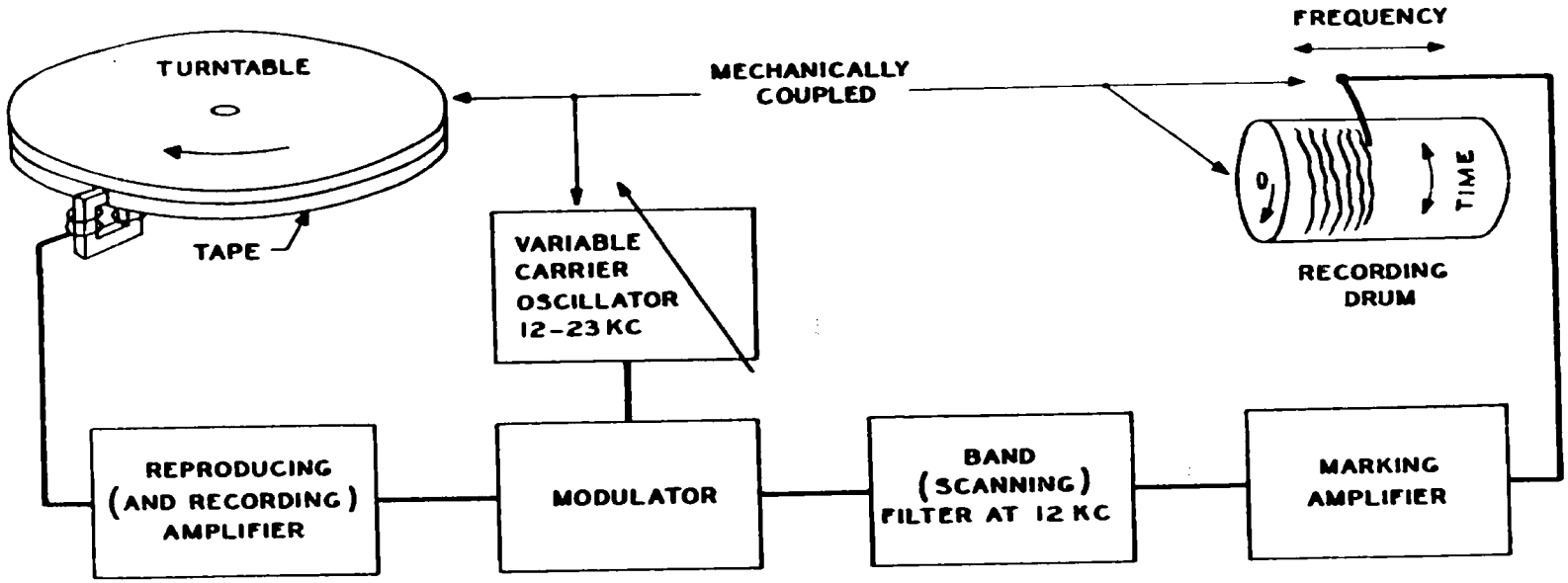
Scanning filters of various widths can be used for different purposes. If the filter is wide, it will give an analysis which is limited in the amount of detail it can portray in the frequency dimension, but it will respond quickly to changes in amplitude with time, and will therefore give sharp time resolution. The narrower the filter the more frequency detail is shown in the spectrograms at the sacrifice, however, of some of the time resolution. With all the filters thus far used, the shift in frequency range from line to line is only a fraction of the width of the filter. Successive lines in the spectrogram, therefore, do not represent successive frequency bands. They represent frequency ranges which overlap by a large fraction of their total width. The density of the patterns, therefore, changes very gradually along the frequency dimension.

The kind of patterns produced by this method of analysis is illustrated in figure 3. The upper spectrogram in the figure was made with a scanning filter about 300 cycles in width. The separate words can be plainly distinguished. The vowels are distinguished by dark bands with vertical striations. The consonants are in general less intense and show a different type of structure. It will be noted that the dark bands are different in the different vowels, and they change not only from one word to the next but also within each word. Analyses of this type, therefore, graphically portray changes in the energy frequency distribution of a complex signal with both time and frequency. It should be emphasized, however, that the relative intensities of the various components of this particular sample of speech, notably the consonants, differ to a far greater extent than would be judged by



UNITS INTERCONNECTED AND READY FOR USE

Figure 1 - D-165529 Sound Spectrograph



This diagram shows the essential elements of the sound spectrograph. The signal to be analyzed is recorded on the loop of magnetic tape, using the reproducing amplifier as a recording amplifier by means of switches not shown. The recording is done at 25 rpm, permitting a sample 2.4 seconds long to be recorded.

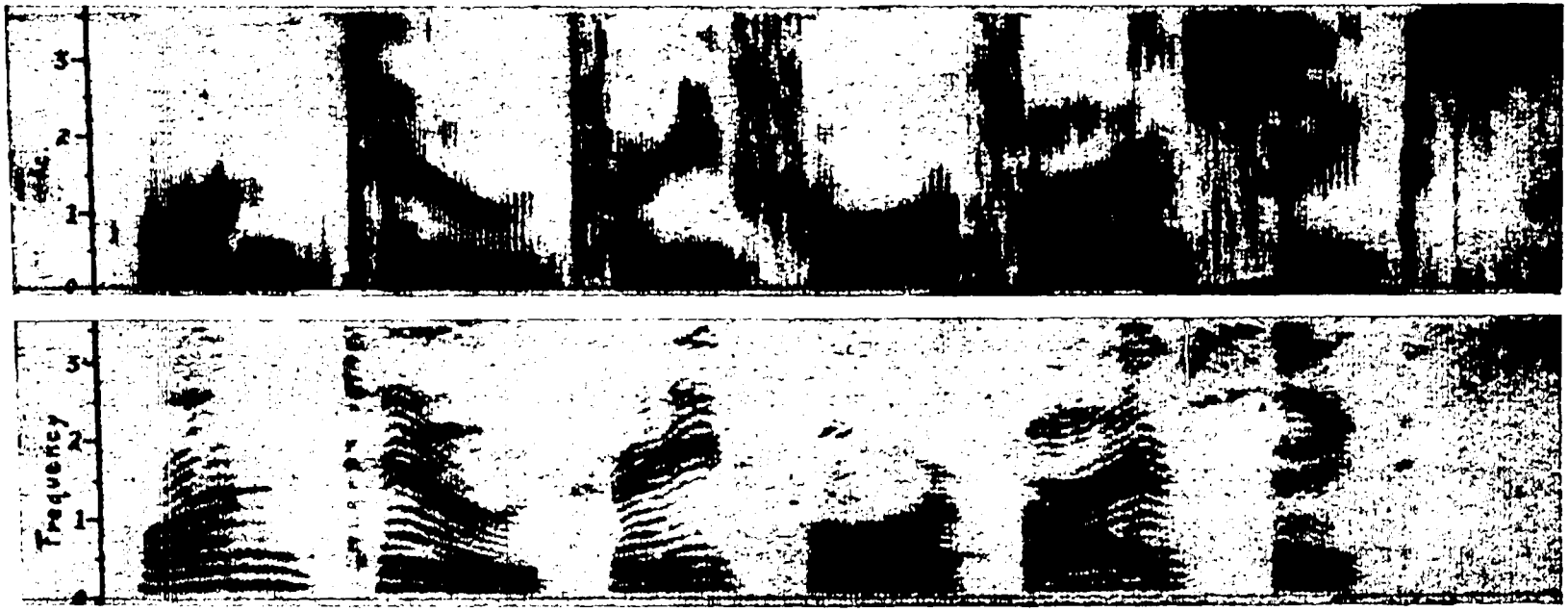
The recorded material is then reproduced at 78 rpm. Because of this speedup, a signal which fills the frequency region from 0 to 3.5 kc now extends to about 11 kc.

The signal is modulated with a carrier which gradually changes in frequency from 23 to 12 kc as the recorded material is reproduced repeatedly. The lower sideband of the resulting signal is thereby gradually scanned by the band filter, which is centered at about 12 kc.

The output of the filter is amplified and is fed to a stylus bearing on facsimile paper, making a trace varying in density with the instantaneous energy passed by the filter. The paper is mounted on a drum which is geared to the magnetic tape. The stylus is gradually moved laterally along the drum as the frequency of the carrier is changed. The spectrogram is thereby built up line by line.

These are the basic elements. Added to these in practice are various means for achieving a high degree of level compression both in the time and in the frequency dimensions. Portraying amplitudes in such a way that they can be interpreted quantitatively also calls for additional equipment in series or shunt with these basic elements.

Figure 2 - Illustrating the Operation of the Sound Spectrograph



These are spectrograms of the words "one, two, three, four, five, six" spoken normally. The time scale is horizontal and the frequency scale is vertical, as indicated. In this illustration the spectrograms are somewhat less than natural size due to photographic reduction and some trimming at the ends.

The upper spectrogram was made with a scanning filter 300 cycles wide. This degree of frequency resolution is sufficient to bring out the resonance bands which characterize the different vowel sounds. Note that these regions of resonance change continually as the oral cavities change their size and shape in forming the different words. This clearly demonstrates how spectrograms portray changes in energy-frequency distribution with time.

The lower spectrogram was made with a scanning filter 45 cycles wide. This degree of frequency resolution is sufficient to show the separate harmonics of which voiced sounds are composed. These are multiples of the voice fundamental, or pitch. Note that they rise and fall in frequency as the voice is inflected. As they rise, they become more widely spaced, and vice versa.

The vertical striations in the upper spectrogram are caused by the fact that the filter is wide enough to pass several harmonics at once. These produce a beat note with a frequency equal to the voice pitch.

A rather high degree of level compression was used in producing these spectrograms. The degree of compression may be varied, depending on how much low level detail may be desired for different purposes.

Figure 3 - Spectrograms of Normal Speech

the relative blackness of their patterns. In other words, a very large amount of level compression is incorporated in these patterns as will be described in the following section.

The lower spectrogram in the figure shows the same words analyzed with a filter only 45 cycles wide. This filter is narrow enough to resolve the individual harmonics of which vowel sounds are composed. The harmonics consist of the fundamental voice pitch together with both odd and even multiples of this frequency. Some of the harmonics are stronger than the others, because they are reinforced by resonance in the oral cavities as the words are formed. It will be noted that the dark areas in these patterns correspond in frequency and in trend with those in the upper spectrogram. The fact that vowel sounds consist of discrete harmonics causes the vertical striations in the patterns made with the wider filter. Whenever the filter is wide enough to pass several harmonics at once, these harmonics beat with each other and form maxima and minima in the output of the filter. The frequency of the beats corresponds exactly to the frequency of the voice pitch.

It will be noted in the 45 cycle spectrogram that the harmonics rise and fall in frequency from moment to moment. This reflects the changing pitch of the voice known as inflection. Inflection is normally used in connected speech, and this fact is of assistance in decoding work, because the spacing and trend of the individual harmonics in spectrograms provide important clues in diagnosing privacy systems as will be demonstrated in subsequent chapters.

3. Level Compression

In normal speech there is a tremendous change in level from moment to moment particularly in the level of consonants as compared to vowels. There is also a considerable difference in the average level at low frequencies as compared to high frequencies. This latter difference can be corrected by predistortion, and present models of the spectrograph contain shaping networks for this purpose. There are, however, changes from moment to moment in the relative levels of high and low frequencies in different speech sounds which cannot be corrected by shaping networks. The facsimile paper on which spectrograms are made has a range of between 10 and 15 db. The range of levels in speech greatly exceeds this value. This means that if the average level is adjusted so that the highest components appear at maximum blackness, the lowest level components will be invisible. Conversely if the level is so adjusted that the low level components appear in the pattern, the high level components will severely overload the recording paper. In order to show both the high and low level components occuring in speech, therefore, it is necessary to compress the instantaneous signal into a much narrower range.

In the earliest models of the spectrograph the marking amplifier shown in figure 2 was given a compressing action by means of a thyrite varistor across the grid of the output stage. Whenever the output of the scanning filter was low the gain of the amplifier was effectively raised from an average condition and whenever the output was high the gain was effectively lowered. This tended to equalize changes in level with both frequency and time. More recently the compressor has been replaced by devices which can exercise certain types of discrimination in controlling the instantaneous gain of the marking amplifier. These devices are known as control circuits. They provide patterns with better resolution in both time and frequency than can be obtained with the compressor. The patterns shown in figure 3 were made with these control circuits in operation. The circuits are described in Preliminary Report No. 27.

4. Possible Improvements

The spectrograph patterns have undergone continual improvement in the course of this project, but they can probably be still further improved. The control circuits thus far produced are by no means the final word. Circuits of this type can be adapted to affect the patterns in various ways, and it is conceivable that different control circuits could be developed for decoding different types of scrambles.

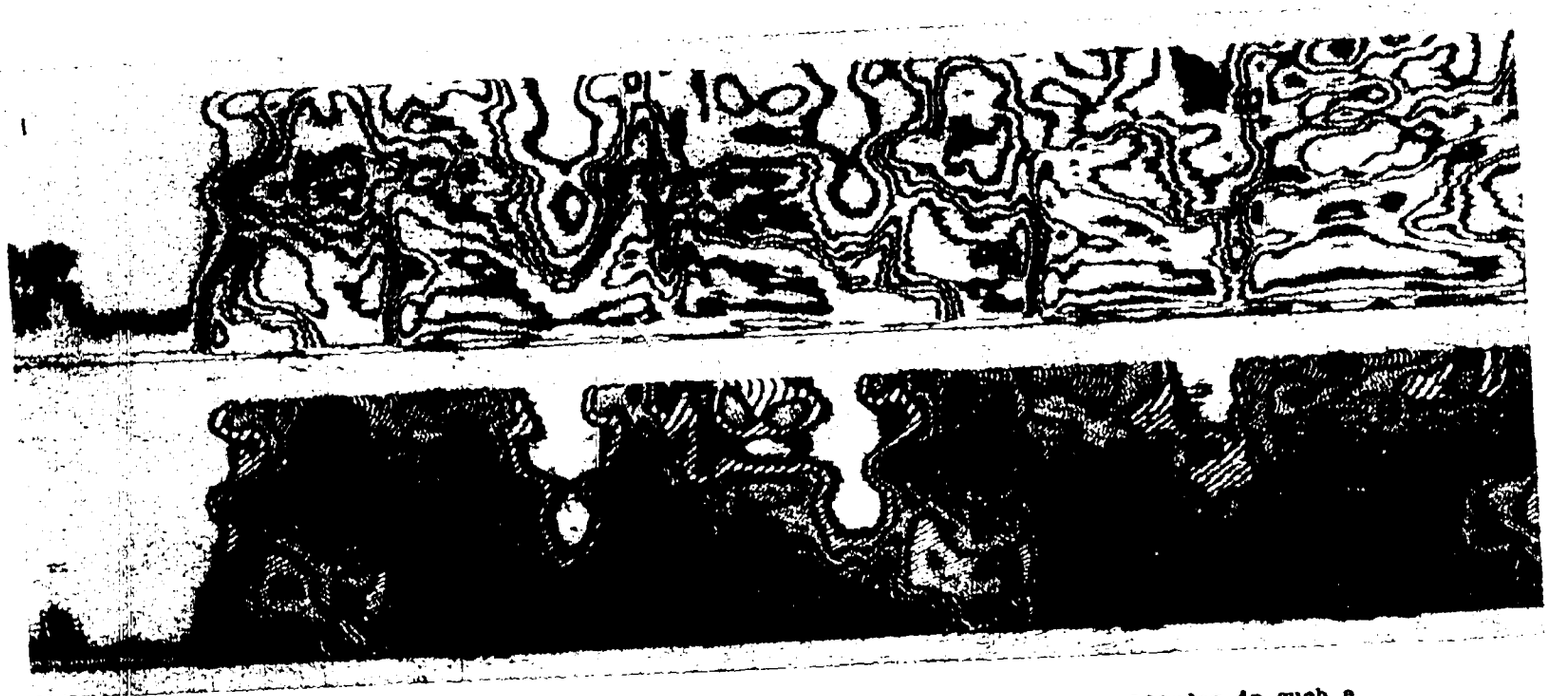
One definite line of improvement concerns the time resolution. Many scrambling methods, as will be seen subsequently, produce sharp discontinuities of the scrambled speech in the time dimension. The process of analyzing the scrambled signal in such a way as to obtain high frequency resolution tends to obscure the signal at these sharp boundaries. This is a basic situation which affects not only the spectrograph, but all types of analyzers. In order to obtain a high degree of frequency resolution, a narrow filter must be used. The narrower the filter, however, the longer its response and decay time, that is, the output of the filter cannot be made to change as rapidly

These spectrograms illustrate one method of representing amplitudes in such a way that they can be interpreted quantitatively.

They are made up of discrete dots, all equally black. The dots are closely spaced in the dark regions, and widely spaced in the light regions. There is a definite quantitative relation between the dot spacing at any point and the level of the signal at that point. The level at any point could therefore be determined by measuring the dot spacing with suitable equipment.

These patterns show the results of exploratory work. Undoubtedly further work would improve the representation.

Figure 4 - Amplitude Representation by Dot Spacing



These spectrograms illustrate one method of representing amplitudes in such a way that they can be interpreted quantitatively.

The representation is similar to that used in contour maps. Every point on any one line represents equal signal level. Successive lines starting from the blank background represent successively higher levels.

In the upper spectrogram, which has only the contour lines, it is not immediately apparent which regions are peaks and which are valleys. In the lower spectrogram, the areas between successive lines have been filled in with various patterns made up of discrete dots. The closer the dot spacing, the higher the level. With this arrangement, it is easy to distinguish peaks from valleys. Furthermore, regions of equal level in different parts of the spectrogram can be recognized by their dot spacing patterns.

These spectrograms show the results of exploratory work. Undoubtedly further work would improve the representation.

Figure 5 - Amplitude Representation by Contours

in level as the instantaneous level of the signal. This causes strong components to "spill over" across the time boundaries. Examples of such spillover can be seen in the upper spectrograms in figures 54, 55, 59 and 60. In general this spillover does not interfere greatly with the recognition of various privacy systems, but it does interfere severely where spectrograms are to be used for decoding work. Several possible remedies for this situation have been devised which are recorded in Chapter VI.

5. Amplitude Representation

In the patterns thus far discussed the instantaneous intensity of the signal is represented by the lightness or darkness of the trace in the spectrograms. This representation is inherently non-linear and practically impossible to make quantitative. For some types of work it would be highly desirable if the amplitudes could be represented in such a way that they could be interpreted quantitatively.

Figure 4 shows a spectrogram which upon close inspection will be seen to be made up of discrete dots. The dots are close together in the dark portions of the spectrogram and farther apart in the light portions. The dots themselves are all of equal blackness. The spacing of the dots is in fact quantitatively related to the instantaneous level of the signal. The level at any point in the spectrogram can, therefore, be measured by measuring the dot spacing with suitable equipment and comparing it with a scale showing dot spacing vs. level.

Another type of representation is shown in figure 5. Here the levels are represented by the type of technique used in representing topographical variations in contour maps. The contour lines each represent regions in which the signal reaches a particular fixed level. The lines may be spaced so as to represent steps of any desired number of db, or any number of volts. In the lower spectrogram the spaces between the contour lines have been filled in with various densities of dot spacing. This permits instant recognition of equality of level in different portions of the signal.

Quantitative amplitude representation may or may not prove useful in decoding work. For certain kinds of signal it should prove useful, because it provides another dimension besides time and frequency which can be used for determining continuity or discontinuity in the signal. In other cases, however, it may prove useless, because changes in level have arbitrarily been introduced into the scramble.

The developments mentioned above emphasize the fact that the sound spectrograph is a highly flexible device and its capabilities along any line can be greatly increased by adding features designed for the specific purpose in mind.

CHAPTER III

SPEECH SCRAMBLING METHODS

In this chapter we shall examine a wide variety of speech scrambling methods in order to become familiar with the devices which might be used alone or in combination to make up speech privacy systems. Some of these systems are in commercial or military use, others exist only on paper, mostly in the form of patents or patent applications. It is not intended to include all the variations of all the different methods but rather to cover basic scrambling methods, with their most important variations, in which the original speech is transmitted with its parts modified, displaced or interchanged.

The two main dimensions of speech which are operated upon to make it unintelligible are the frequency dimension and the time dimension. Scrambling systems usually depend on rearranging the components of speech in either or both of these dimensions. In general it may be said that those that operate on the frequency dimension alone are capable of the best quality in the reproduced speech. A complete list of the systems covered in the discussion is given in Table I together with other data concerning them. Most of them are illustrated by means of spectrograms which will be discussed in Chapter IV.

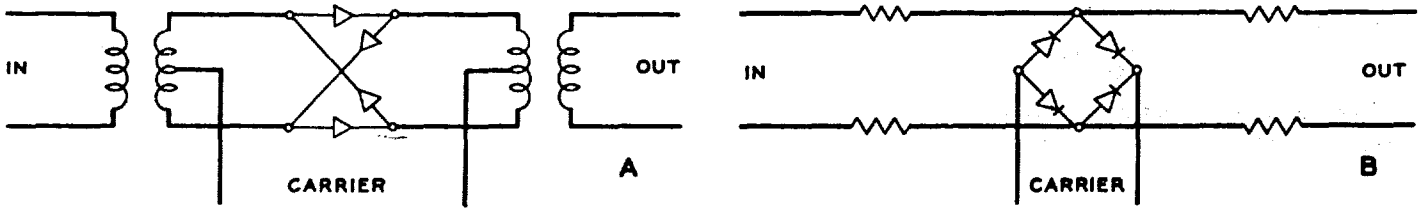


Figure 6 - Modulators

1. Systems Involving Single Modulation

A basic device in privacy systems is the modulator. One form of modulator, shown in figure 6A, consists of four copperoxide varistor units between two balanced coils. The carrier frequency is fed into the midpoints of the coils as shown. In some cases the coils can be omitted as shown in figure 6B.

Figure 7 shows the method of producing simple inversion. In this and in succeeding illustrations the numerical values are not necessarily the best values for practical operation, but they serve to illustrate the manner in which the device operates.

In the system shown in figure 7 the speech band is limited to 3,000 cycles by a low-pass filter. It is then modulated with a frequency of 3,000 cycles. This produces an upper and a lower sideband of which only the latter is passed by the output filter. The system is called inversion, because the high frequencies in the original speech appear as low frequencies in the output and the low frequencies in the original speech appear as high frequencies. At the receiving end the inverted signal, in passing through an identical system in the same direction, is reinverted back to normal speech.

A very commonly proposed variation of this system involves using a variable frequency instead of the steady 3,000 cycle carrier. We might vary the frequency continuously or in discrete steps. It should be noted, however, that the cut-off of the low-pass output filter

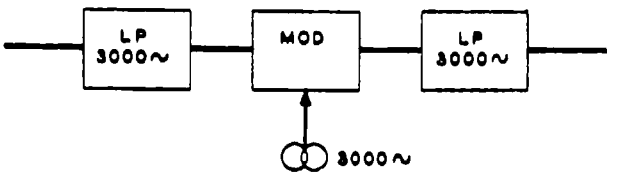


Figure 7 - Single Modulation

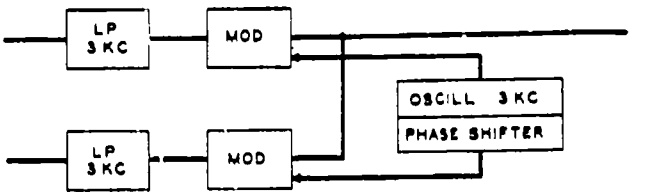


Figure 8 - Split Phase Multiplex

is fixed which limits the variation permissible in the carrier frequency. A wide variation would either permit too much of the upper sideband to get through or would cut off some of the lower sideband. Two proposals of this type are discussed in Preliminary Reports No. 8 and 20.

If the modulator in figure 7 is of the type shown in figure 6A, speech can be scrambled by introducing instead of the 3,000 cycle carrier a square wave whose changes from positive to negative value are irregular in time. Each one of the reversals in the carrier wave causes a reversal of phase in the speech wave. The pattern of these irregular reversals may be arranged so that the speech becomes unintelligible. At the receiving end a coding wave must be introduced which is exactly in step with the one at the receiving end, with proper allowance for any delay there may be in the transmitting channel.

A two channel system using one modulator for each channel, is shown in figure 8. In this system the carrier fed into both modulators is the same in frequency but differs 90 degrees in phase. Two separate speech channels can be transmitted by this method without substantial mutual interference, but both sidebands as well as the carrier must be transmitted. At the receiving end the carrier must be split into two components with the proper phases. Each component will demodulate its own portion of the signal and thereby separate the two speech channels. Naturally one of the channels may consist of noise or spurious speech from a recording or the like, which tends to mask the real message if the signal is demodulated with an ordinary set. This scheme was originally proposed as a multiplex system, but an obvious variation is to divide a single speech band into two halves with filters and then transmit the two halves on carriers differing by 90 degrees in phase.

8. Systems Involving Double Modulation

Figure 9 shows a much more flexible system. Here the signal is modulated twice, with a band-pass filter between the two modulators. With this arrangement the carrier frequency fed into the second modulator can be varied in several ways. In the illustration two carrier frequencies are shown for the second modulator. If the 8 kc value is used the output consists of the speech band right side up but displaced from its normal position by 2,000 cycles. If the 16 kc value is used the output consists of the 3,000 cycle speech band inverted and displaced by 3,000 cycles. We might use these two values alternately at short intervals, or we might have the carrier vary continuously back and forth, say between 13 and 16 kc. Another variation is to use a multiplicity of values, say 500 cycles or 1,000 cycles apart, (not between 10 and 13 kc for this illustration,) and switch between these values in a regular or irregular sequence. A disadvantage of these systems, of course, is that the transmission channel needs to be wider than that usually afforded by radio sets or telephone lines. In all of these systems, the speech is

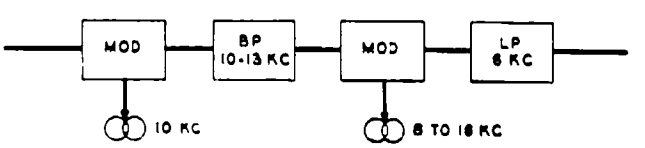


Figure 9 - Double Modulation

restored by passing through identical equipment in the opposite direction.

3. Triple Modulation - Reentrant Band Shift

Going back to figure 7 suppose the carrier frequency were made 4,000 cycles instead of 3,000, but retaining the 3,000 cycle input and output filters. The output would then be an inverted sideband ranging from 1,000 cycles to 3,000 cycles; that portion of the sideband above 3,000 cycles would be cut off by the output filter. Since, however, there is a 1,000 cycle gap at the lower edge of the transmitted band, the portion which would be cut off by the filter might be modulated down and sent along with the rest of the signal in this lower part of the spectrum. In other words the portion of the sideband which would otherwise disappear above the upper edge of the transmitted band might be made to reappear at the bottom.

Figure 10 shows a system of modulators and filters for accomplishing this kind of band shift. The first modulator is followed by a high-pass filter which selects the upper sideband from 3,000 cycles to 6,000 cycles. This is combined with some of the original signal which ranges from 0 to 3,000 cycles. The second modulator is fed with a carrier frequency of say 7 kc which inverts the whole band. This is followed by a band-pass filter passing the range from 3 to 6 kc. A second modulator with its carrier frequency placed at the lower edge of the 3 to 6 kc input band moves the whole band, still inverted, down to the usual range of 0 to 3,000 cycles. The upper sideband of this modulation step is removed by the output low-pass filter. A variation of this arrangement is to allow the 7 kc carrier to vary in discrete steps according to some regular or irregular program or vary it continuously between the limits of 6 to 9 kc. This provides a variable band shifting arrangement without using more than the normal 3,000 cycle transmission channel.

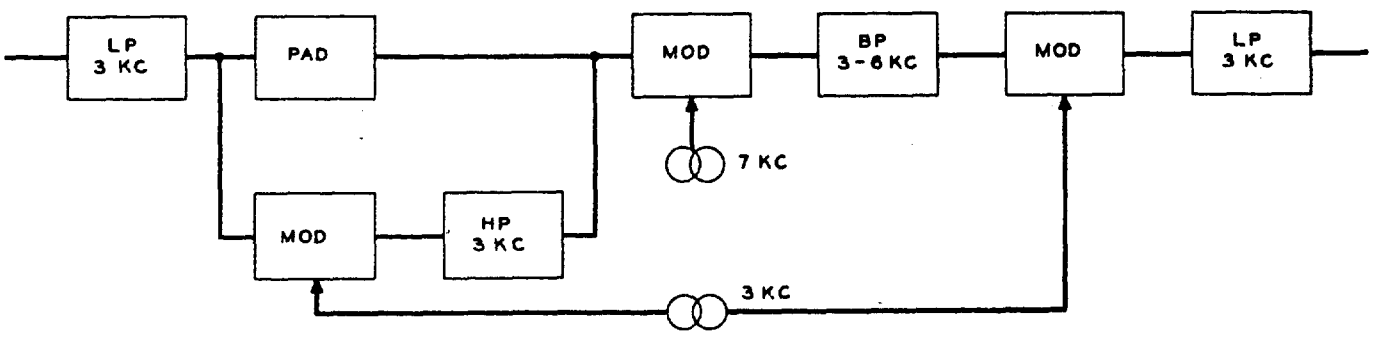


Figure 10 - Reentrant Inversion

4. Band-splitting Systems

A privacy system in wide commercial use, known as the split band system, involves splitting up the whole speech band into a number of subbands and shifting these around out of their normal positions in the frequency spectrum. Figure 11 shows one manner in which this can be accomplished. The numerical values are chosen so that the band from 250 to 3,000 cycles is divided into five subbands each 550 cycles wide.

The speech band is fed to five modulators in parallel. The five band filters following the modulators are all alike, passing the band from 3,100 to 3,650 cycles. It will be seen that the uppermost modulator in figure 11 with its carrier of 6,100 cycles will invert the speech band and displace it by such an amount that the frequency band which originally occupied the space from 2,450 to 3,000 cycles will pass through the filter. In other words this modulator in combination with its band filter selects the uppermost of the five subbands from the input signal. Similarly the lowest modulator in combination with its band filter selects the lowest subband from the input signal. The outputs of the band filters all occupy the same frequency range, but they all came originally

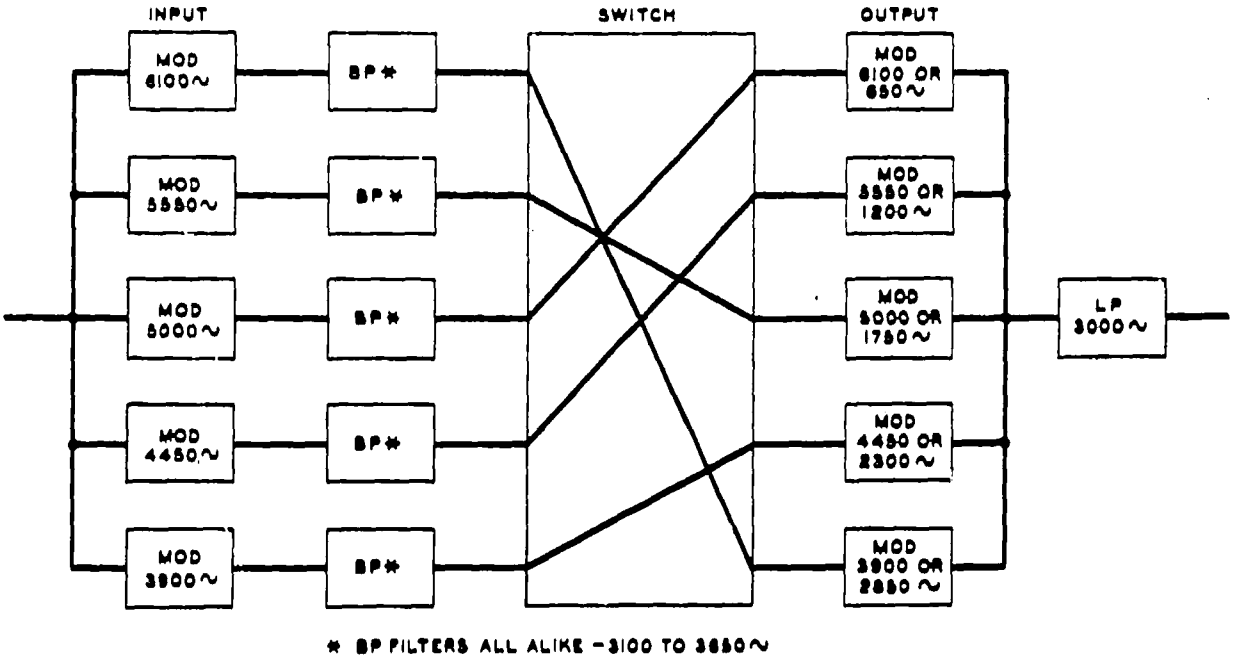


Figure 11 - One Form of Split Band System

from different frequency ranges. Similarly the output modulators are so designed that each one accepts the band from 3,100 to 3,650 cycles and shifts it to a particular band location in the output. The five leads going into the box labeled "SWITCH" may, therefore, be cross-connected in any desired manner with the five output leads. The resulting output will always cover the complete range from 250 to 3,000 cycles and there will be no overlapping subbands.

An additional set of frequencies is indicated in the drawing for the second set of modulators. These frequencies will cause the output subbands to be inverted instead of right side up. One or all of these alternate frequencies may be used as desired.

The "SWITCH" may be changed as often as desired. On an experimental basis the codes have been changed as often as 25 times per second without appreciable distortion in the quality of the received speech showing that it is possible to shift bands as wide as 550 cycles at a rapid rate without generating appreciable distortion products.

5. Time Division Multiplex

Time division multiplex (T.D.M.) is a system in which n separate signals occupying the same frequency range are sent over a single line, each signal being transmitted only 1/nth of the time. This might be illustrated by showing the n signals connected to the n segments of a commutator. A rapidly rotating brush picks up the n signals one after the other. For acceptable quality, however, the brush must make at least as many rotations per second as the highest frequency in the transmitted signal. This means that a mechanical brush is out of the question and is used simply for illustration. This kind of switching, however, can be accomplished with electronic ring circuits.

Since we are interested here in privacy systems rather than multiplex systems, we will confine ourselves to the use of T.D.M. for transmitting a single speech channel. This can be accomplished by dividing the speech band into n number of subbands all occupying the same frequency range, and connecting these to the segments of our hypothetical commutator. Referring to figure 12 which is similar to 11, this can be

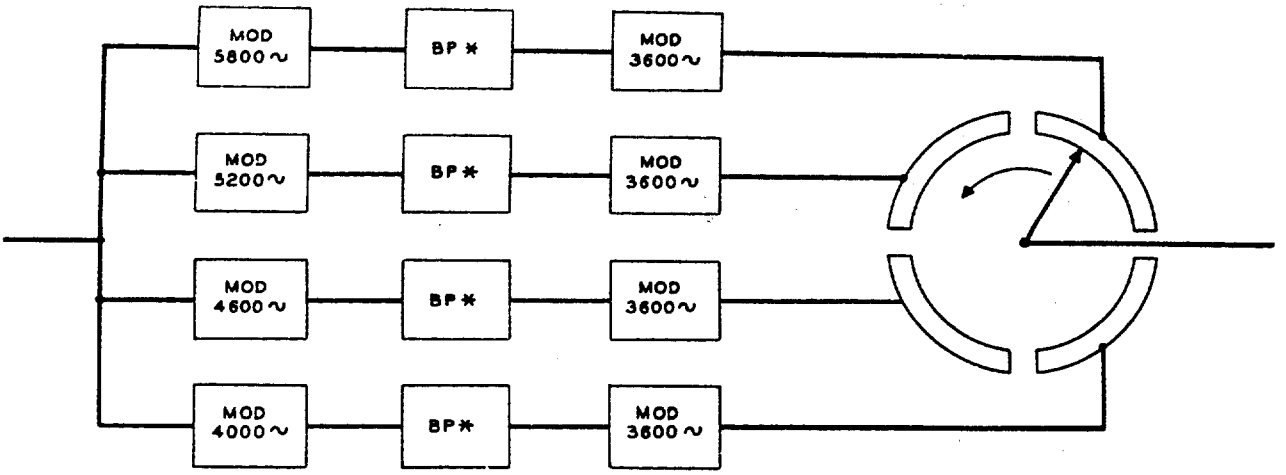


Figure 12 - One Form of Time Division Multiplex

accomplished by feeding all the output modulators with the same carrier, and connecting each modulator to a commutator segment. In this illustration, there are four 600 cycle subbands, covering the range from 400 to 2,800 cycles. It has been shown mathematically that the output of this system consists of sidebands around a frequency corresponding to the rotation of the brush and also sidebands around frequencies corresponding to odd harmonics of the rotation frequency. Each side-band, however, contains components from each of the subbands. It has also been shown that the total channel width required for good transmission need be no greater than that of the original signal.

To increase the privacy of this system one of the subbands may be replaced by a band of noise. This can be filtered out at the receiving end. Obviously this system requires a high degree of synchronism between the two ends.

6. Systems Using Tape Recording

Leaving the frequency substitution systems for the time being, we will introduce a device which permits operating on the time scale. The most versatile device for this purpose is the magnetic tape recording and reproducing system. This takes the form of a tape of magnetic alloy a few mils thick either run as a loop over pulleys or attached firmly to the perimeter of a disk. The recording is done by means of small electromagnetic pole-pieces. The signal is picked up by similar pole-pieces which may be placed at a distance from the recording pole-piece depending on the amount of delay desired. The outstanding advantage of the magnetic tape system for this type of application is that the signal may be erased and the recording medium be used over and over again. The quality of this type of transmission can be made very good with proper design.

Figure 13 shows a rather simple privacy system using magnetic tape. The input signal is passed through a 3-way pad, whereby it is impressed on a band filter and also recorded on the magnetic tape. It is picked up by equally spaced pole-pieces each associated with a different band filter. With the arrangement shown in figure 13 the band from 0 to 1,000 cycles is transmitted without delay. The band from 1,000 to 2,000 cycles is transmitted with say, 100 milliseconds delay and the band from 2,000 to 3,000 is delayed 200 milliseconds. At the receiving end the scrambled signal is passed through an identical system in the same way except that the two extreme band filters are interchanged. In this way the band which received no delay in transmission is given maximum delay in the receiving machine, and the band which re-

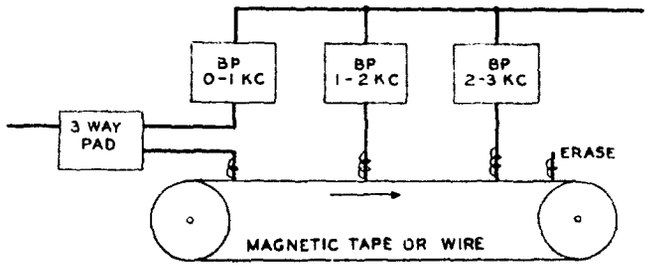


Figure 13 - Variable Subband Delay

ceived maximum delay in transmission is given zero delay in the receiver. In this way all the bands are delayed the same amount and the speech is restored to normal.

This system alone does not provide any high degree of privacy but it can be combined with other systems as we will see subsequently.

An important class of scrambles involving magnetic tape is known as time division scrambling (TDS). A simplified diagram of this system is shown in figure 14. There is a recording pole-piece and a number of pickup pole-pieces. There is also a commutator driven in synchronism with the tape. The length (in time) of each segment of the commutator is, in general, equal to the delay between successive pickup

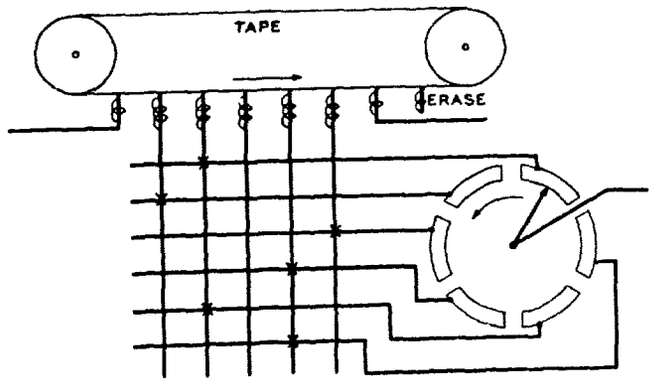


Figure 14 - One Form of Time Division Scramble

pole-pieces. However, the number of segments need not be the same as the number of pole-pieces. A switch is provided whereby any segment may be connected to any pole-piece.

With this system the speech is cut up into time elements corresponding in length to the spacing of the pole-pieces. These time elements are transmitted in a scrambled order. For instance, 6 successive time elements which we might label 1,2, 3, 4, 5, 6 might be transmitted in the order, 2, 4, 1, 3, 6, 5. The possibilities of TDS coding are far too complex to cover here. Analytical discussions are given in Preliminary Reports Nos. 3 and 6. The general requirements for all TDS systems, may be stated as follows: (1) Each element of the original speech must be transmitted once and only once.(2) The sum of the delay in the transmitting machine plus the delay in the receiving machine must be equal for all elements. With these two requirements fulfilled it is obvious that the speech comes out of the receiving machine in its normal order. It is delayed, however, by an amount equal to the sum of the transmitting and receiving delay.

At the receiving end there are several ways of handling the scrambled signal. (1) The pickup pole-pieces can be used as recording pole-pieces and the signal picked up by an additional pole-piece shown at the right in figure 14. With this arrangement the connections between the commutator and the polepieces are the same in the transmitting and receiving machines. (2) The signal can be recorded with the same pole-piece used in the transmitting machine and the connections between the pole-pieces and the segments rearranged for receiving by a push-to-talk relay. (3) The codes can be restricted to a particular class called self-converse codes. These have the property of being self-decoding, that is, the same code which scrambles the speech in the transmitter restores it in the receiver.

An important variation of this system is called "Interlace". In this system the number of segments on the commutator is doubled. The odd segments are connected to the pole-pieces according to one code and the even segments are connected according to a completely independent code. The reason for this device is to increase the difficulty encountered by the the enemy in trying one code after the other to find the right one, particularly if the total number of codes available is small. With the

interlace system the total number of combinations possible is equal to the square of the number of codes.

The rotating commutator shown in figure 14 results in a repeated code, that is, each rotation produces the same scramble. It is possible to substitute for the commutator and switch arrangement shown in figure 14, a more complex arrangement whereby the speech is scrambled in a never-repeating manner. There are several ways of accomplishing this. Perhaps the simplest way to represent it is as a punched tape which permits the pole-pieces to be connected to the output, one at a time, in any desired order permissible under the restrictions outlined above.

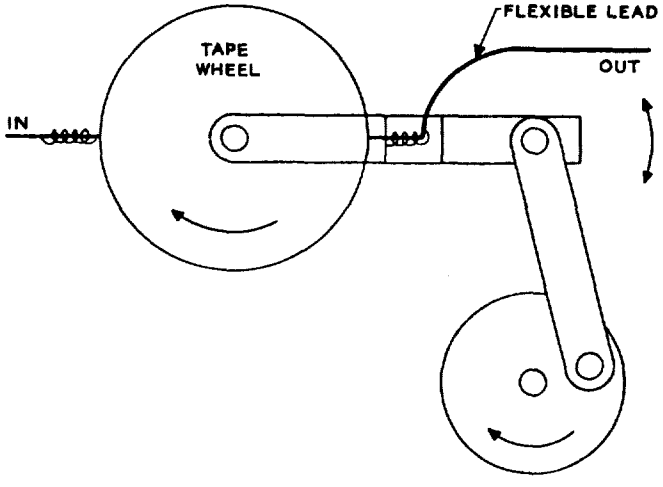


Figure 15 - Speed Wobble

Another way of utilizing magnetic tape to scramble speech is shown in figure 15. Here the pickup pole-piece is oscillated back and forth along the tape mechanically. With this arrangement, or other variations equivalent to speech changes, the speech time scale is alternately compressed and expanded. The frequency scale is correspondingly expanded and compressed, respectively.

With the arrangement shown in figure 16, speech is broken up into time segments each of which is transmitted backwards. The motion of the pickup pole-pieces is twice as great as the motion of the tape and is in the same direction. Therefore, the relative motion of the tape and the pole-pieces is the reverse of that used in recording. This is the same as running the tape backwards for reproduction.

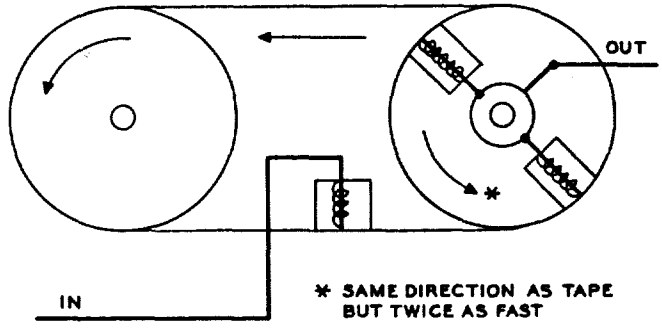


Figure 16 - Time Inversion

7. Combinations of Time and Frequency Scrambling

Obviously the two kinds of systems described in the previous sections can be used together. For instance, some of the time elements of a TDS system might be inverted according to a regular or irregular program. The next more complex step is to combine the band splitting system of figure 11 with the TDS system. The codes of the band splitting system might be fixed or might be switched in synchronism with the TDS elements, the time scale of the scrambled speech not being further broken up. If they are switched nonsynchronously, however, the time dimensions will be further broken up as will be seen subsequently. Combinations of nonrepeated code TDS and rapidly switched split band coding can be made to afford a very high degree of privacy. The two kinds of coding, of course, must not be so interrelated that one furnishes clues for the other. If, for instance, a certain pole-piece were systematically associated with a certain split band code the total privacy of the combination might be impaired rather than enhanced. A coding method for avoiding this difficulty is described in Preliminary Report No. 21.

A very special kind of scramble is produced by a system which consists functionally of figure 11 (rapidly switched) in tandem with

figure 13 (with 5 bands) followed by an additional figure 11. This is not the simplest form of the system, but it serves to illustrate the principle. Two frequency scrambles with a time shift in between produce a particular kind of two-dimensional scramble in which the speech is broken up into both time and frequency elements. Each of these elements may be shifted both in time and in frequency so as to be out of proximity with other elements with which they were originally associated either in time or in frequency. Another way of accomplishing this kind of scramble would be a combination of rapidly switched split band with a separate TDS system in each subband. A two-dimensional system has been described in the Brown Boveri article reproduced in Preliminary Report No. 5 and analyzed in Preliminary Report No. 9. It is capable of a very high degree of privacy.

For the sake of completeness two other systems involving time and frequency shifting will be mentioned, although as far as is known they exist only on paper. Suppose a sample of speech were recorded on tape and then reproduced at twice its normal speed. It would occupy only half the time it took to speak the words, but its frequency range would be twice the normal range. Let the upper half of the expanded frequency range be separated by a filter and modulated down to the normal range and used to fill up the unused time. The directly opposite but analogous system would involve reproducing recorded speech at half its normal speed; the frequency range would then be only half the normal range. Alternate sections, therefore, could be modulated up to fill the unused frequency space, thereby keeping the total transmitting time substantially unchanged. In both of these systems, there would be a delay equal to the length of one time element.

8. Wave form Modification

Thus far we have considered systems in which frequency bands were shifted around or time elements were rearranged. There are a few privacy systems which make speech unintelligible by a direct modification of the wave form. One of these is shown diagrammatically in figure 17. It depends upon a process whereby two waves are multiplied together, that is, the instantaneous amplitude of the resulting wave is the product of the amplitudes of the two input waves. One of the input waves to the multiplier is speech. The other is a complex coding wave. If the coding wave is sufficiently complex the resulting scramble is unintelligible. At the receiving end a reciprocal of the coding wave is derived and used as a multiplier, thereby restoring the original speech. Naturally, the coding waves at the two ends of the system must be in close agreement, otherwise there will be considerable background noise in the decoded speech.

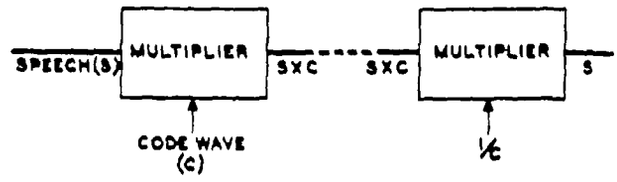


Figure 17 - Wave Multiplication

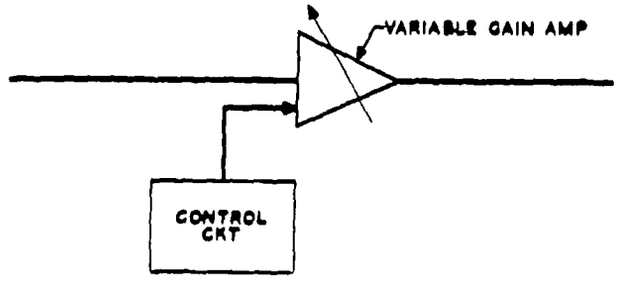


Figure 18 - Level Modulation

Another method for changing the wave form is shown in figure 18. The essential feature of this system is an amplifier whose gain can be varied rapidly with time. Drastic changes in the level of speech, if they occur rapidly enough, will make the speech unintelligible. The level changes might be made according to some program or they might be made to follow the speech wave itself. For instance, extreme compression or expansion could be used. Corresponding gain changes, of course, must be made at the receiving end.

A variation of this system is shown in figure 19. Here the speech band is first divided into sub-bands, and these are individually subjected to level changes according to separate programs.

9. Masking Systems

One of the first schemes which is likely to occur to a person considering how to

make speech private is to add noise or other disturbing signal to the speech and remove it at the other end, in other words, to mask the speech. He will find, however, that it is necessary to use very high levels of masking signal in order to hide the intelligibility. This of course, makes it difficult to subtract out satisfactorily; the difficulties are such that masking systems are more likely to be found on wire lines than on radio. A few speculative masking systems are outlined below.

One form of masking system is shown in figure 20. In this system, two telephone lines are used. At the sending end, noise is added to the speech in a mixing pad and the combination is sent over line 1. The noise alone is sent over a second line and it is used at the receiving end to cancel the noise transmitted with the speech by simple subtraction. This system has the advantage that the noise can be completely random. However, since the enemy might take taps from both lines and thereby be able to make the same subtraction, a variation of this system consists in distorting the noise

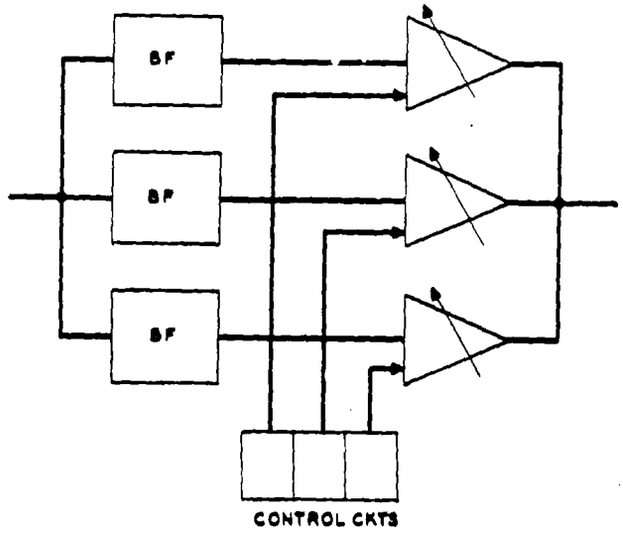


Figure 19 - Subband Level Modulation

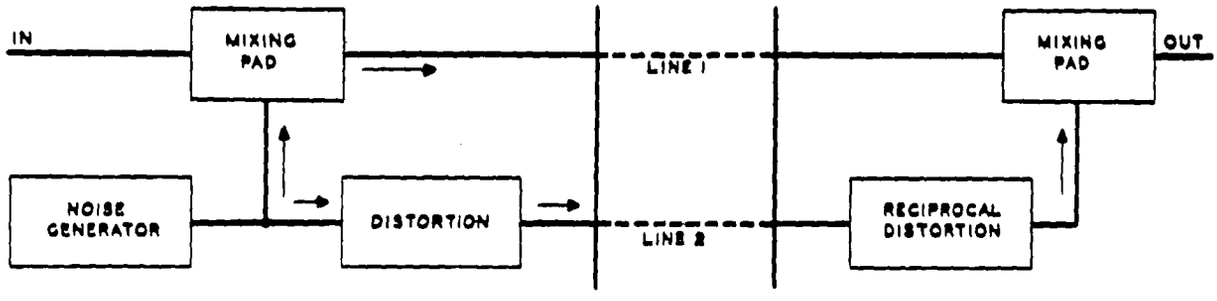


Figure 20 - Noise Masking Using Two Channels

in some predetermined manner before sending it over the second line. At the receiving end, this distortion is first nullified so that the noise may be subtracted. Naturally, the form of distortion must be unknown to the enemy. It can, of course, be varied from moment to moment.

Another masking system is shown in figure 21, which uses only one line. In this system, noise is added to the line at the receiving end instead of at the sending end. Again, the noise can be perfectly random. Since the noise is generated at the receiving end, the process of cancellation can, theoretically, be made very exact. This system, however, cannot be used for radio at all because the level of the

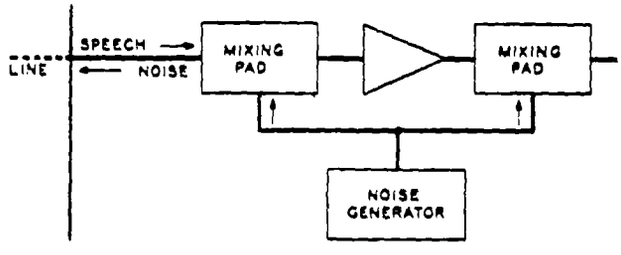


Figure 21 - Masking Noise Applied at Receiving

noise decreases with distance from the receiving station, while the level of the signal increases. The interceptor, therefore, will get good speech signals if he is close to the transmitter. With telephone lines this differential can be kept small.

Another simple masking system is to have a sequence of tones superposed on the signal at the transmitting end. At the receiving end, sharply tuned band elimination networks can be synchronously switched so as to remove the tones from the listener's ear. Similarly, short spurts of noise covering the whole frequency band can be applied at the transmitting end and shorted out at the receiving end. The spurts can be made to occur at irregular intervals according to a never-repeating program. Both of these systems involve the loss of small portions of the speech, either in the time scale or the frequency scale.

A system described in Preliminary Report No. 4 might be classified as a masking system, although it might be better classified as a means of communicating without the enemy's knowledge.

10. Vocoder Systems

The vocoder system which has been described in the Bell System Technical Journal[1] and the Accoustical Society Journal[2] may be made the basis for privacy systems of various kinds. The system is shown schematically in figure 22. At the transmitting end the speech is passed through a series of band filters, the outputs of which are individually rectified to form a fluctuating d-c signal. These signals are individually modulated in such a way that they can all be sent over a single transmission path.

At the receiving end synthetic speech is manufactured in accordance with the signals transmitted over the line. A source of noise which covers the whole frequency range is passed through a set of band filters similar to those at the transmitting end. The output of each of these filters is controlled so that it is the same level as the level of the speech in the corresponding band at the transmitting end. This is accomplished by separating the signals in the various channels, detecting them and using the resulting fluctuating d-c to control the variable gain amplifiers in their respective channels.

The noise is of two types, depending on whether a voiced or unvoiced sound is to be simulated. For an unvoiced sound, it is a hiss like thermal noise. For a voiced sound it is a buzz which consists of a series of harmonics covering the whole frequency range. A separate carrier is used to transmit information for operating this part of the system. At the transmitting end the pitch used by the talker is measured and this information is used to control the pitch of the buzz sound. The absence of a pitch signal switches the hiss sound into the system.

This system by itself, of course, is not private, since the enemy can build a similar system and use the signals to regenerate speech. Privacy must be achieved by operating on the channel signals. One method is to permute the channels at short intervals according to a prearranged program. Another method is to put a TDS system into the line, or into each channel separately. A still more effective method of this type is to apply a two-dimensional scramble, such as was described in an earlier section, to the channels so that signal elements are displaced in both time and frequency.

11. Channel Mixing Systems

Thus far, the methods we have examined apply to a single transmission path. There is another class of privacy system which depends on using a multiplicity of paths. This is, of course, inefficient if only a single message is to be transmitted. However, the method can be applied to cases where a number of channels exist between two points and a number of messages would normally be transmitted over these channels simultaneously.

---

1 - Bell System Technical Journal Volume XIX, Page 495 October 1940
2 - Journal Accoustical Society of America Volume 11, Page 169 October 1939

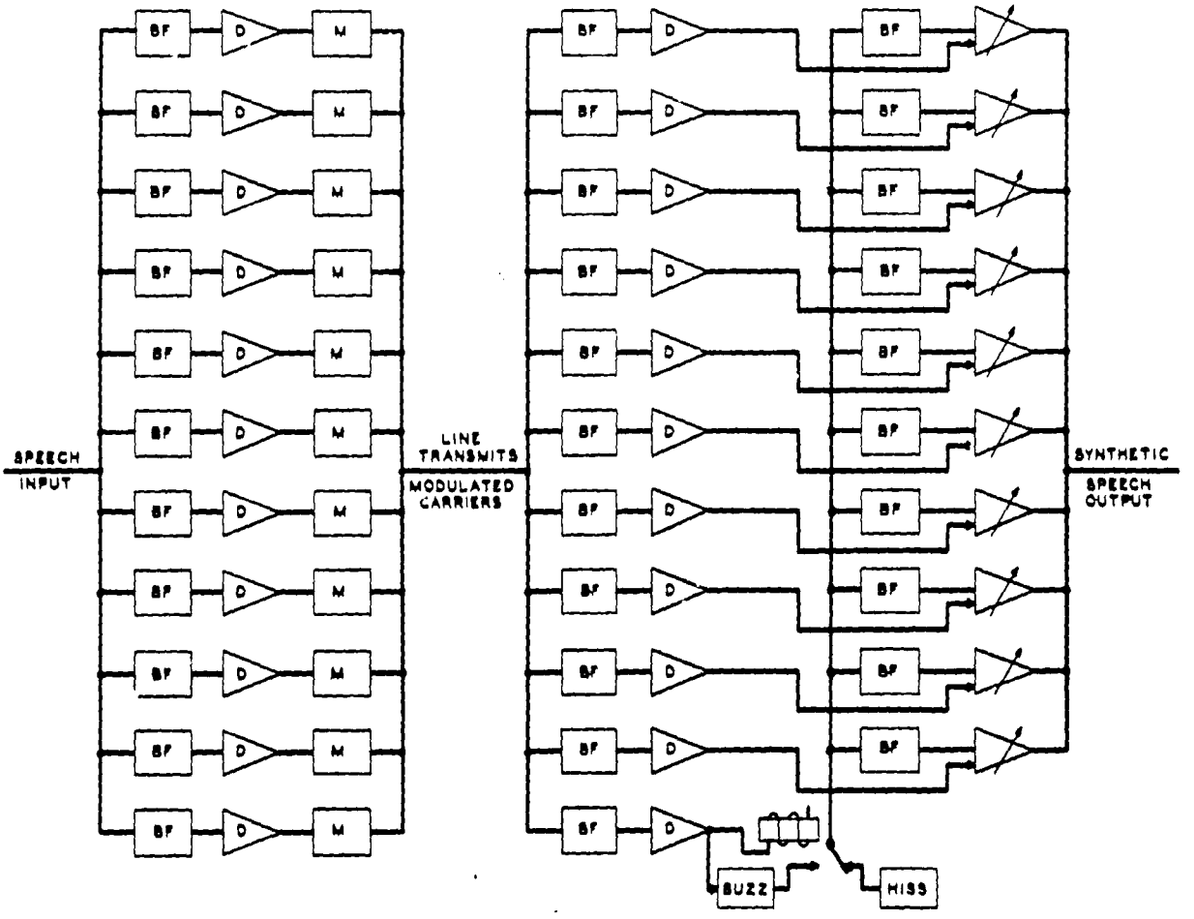


Figure 22 - Vocoder System

Figure 23 shows one form of channel mixing system. Here three channels are shown connected to the three segments of a commutator. Three brushes on this commutator are connected to the outgoing channels which are thereby caused to pick up one channel after the other on a time division basis. Each channel contains parts of messages from all three channels. The commutator, of course, is too simple to be very effective and would, in practice, be replaced by a permuting switch capable of switching according to a more complex program. One or more of the channels may be filled up with noise or spurious speech from a recording or other similar source.

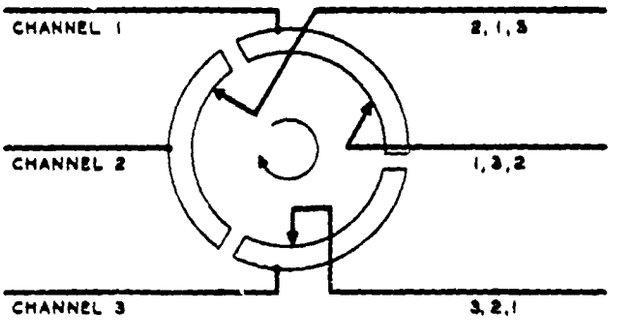


Figure 23 - Channel Mixing

An analogous system which divides the messages on a frequency basis is shown in figure 24. Here each channel is passed through three band filters which divide the speech into subbands. Each of the outgoing channels contains subbands from each of the incoming channels. To increase the privacy, a permuting switch is shown which rearranges the subbands on a time division basis. If only one message is to be transmitted the other channels can be filled in with noise or spurious speech.

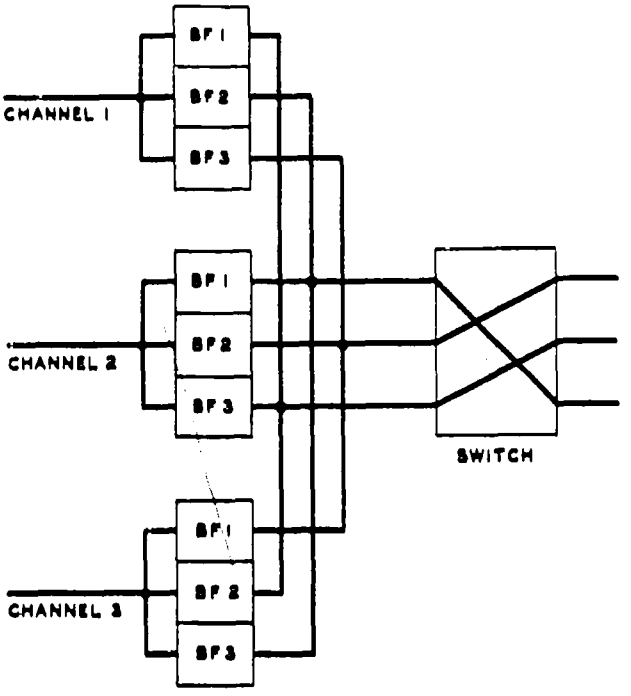


Figure 24 - Subband Channel Mixing

12. Summary

The above examples cover fairly completely the range of schemes that might be used to scramble speech at audio frequencies. In subsequent chapters we will examine each system from the decoding standpoint. To facilitate reference to the various systems, they are summarized in the attached Table I. This table also refers to the places where methods of decoding the various systems are discussed.

CHAPTER IV

DIAGNOSIS OF UNKNOWN SYSTEMS

Before discussing the diagnosis of speech privacy systems it should be pointed out that facts concerning the origin of unknown signals are often very necessary to their correct interpretation. Such things as the frequency, strength, and direction of the signals, the location and type of receiver, and the manner in which the signals were recorded, can be very important pieces of data. That is why it was stated in Chapter I that interceptors should be equipped with complete knowledge of the various kinds of radio systems and transmissions used by both sides, including jamming and radar signals as well as telegraph and facsimile signals. Some of these signals, particularly if transmitted with suppressed carrier, can give extremely puzzling results if demodulated with an ordinary radio set. These possibilities should be taken into account if signals are found which do not seem to fit into the classes discussed below.

As stated in Chapter II, the spectrograph is of tremendous assistance in recognizing the nature of an unknown scrambling system. The ear can usually recognize the presence of time discontinuities. It can also usually recognize the peculiar quality which results from band shifting systems. The exact nature of the scramble, however, is usually impossible to establish with the ear. Even scrutiny of the wave form may yield no clue. The strikingly graphic analysis provided by the spectrograph, however, usually takes the mystery out of the scrambling method immediately.

Speech privacy systems having frequency sub-bands will show horizontal discontinuities or boundaries in their spectrograms. Similarly systems employing time division will show vertical boundaries. A considerable variety of systems display both horizontal and vertical boundaries. How to tell these different scrambling systems apart is the subject of the discussion and illustrations in this chapter.

1. Measurements on Spectrograms

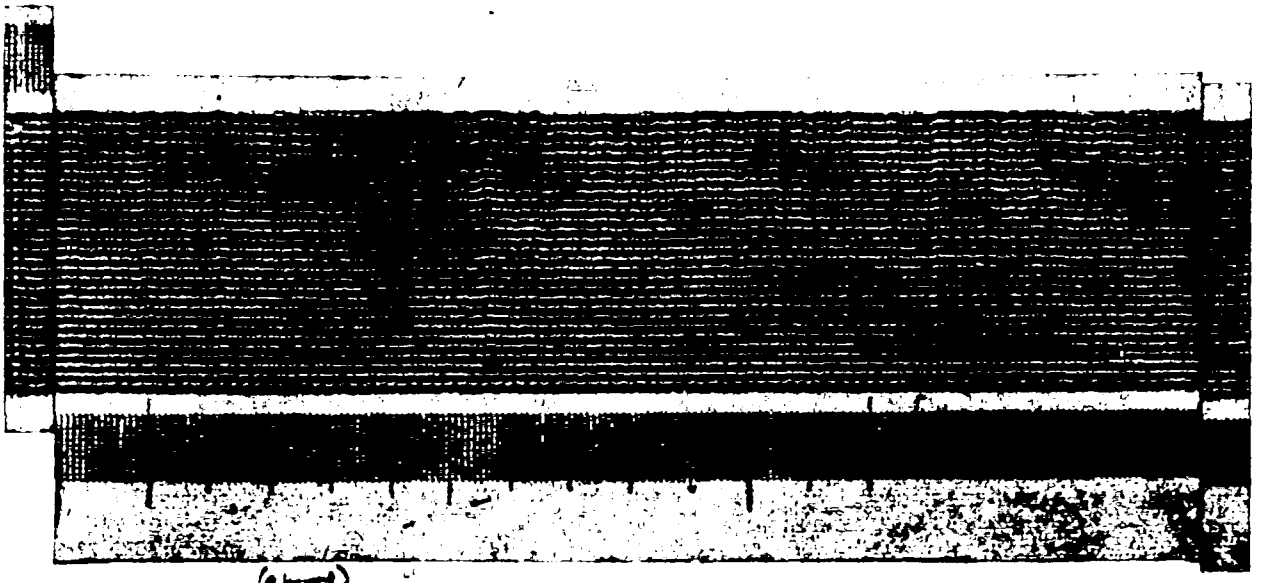
Since an important part of the diagnosis procedure consists in determining the length of time elements and the location of frequency boundaries, let us first examine the procedures whereby the time and frequency scales of the spectrograph can be established. The spectrograph is equipped with a calibrating device which consists of means for producing a complex wave rich in harmonics from the 60 cycle power supply. Spectrograms of this wave made with both the 45 cycle filter and the 300 cycle filter are shown in figure 25. If the power frequency is known, the horizontal and vertical striations in these patterns provide the time and frequency scales. If the power frequency is not known the scales may be established by the formulas given in the figure. This involves additional measurements with a stop watch.

The application of this method to 11 kc spectrograms is not explicitly stated in the figure. A value of K for this condition can be found by the same formulas. This establishes the time scale for the 11 kc spectrograms. For the frequency scale the same pattern is used as for the 3.5 kc spectrograms. However, each horizontal striation is labeled with a frequency obtained by multiplying the normal frequencies by the ratio of the two K's.

Figure 26 shows how these scales can be used to measure the time and frequency boundaries in a scramble. It will be noted that for measuring the time elements spectrograms made with the 300 cycle filter are best because they have sharper time boundaries. For measuring frequency boundaries the same filter must be used as was used in obtaining the scale. It may be noted here that in present models of the spectrograph, the wide filter has a different absolute location than the narrow filter and therefore should not be used to estimate the frequency of components or boundaries.

2. Illustrations of Scrambled Speech

Spectrograms illustrating a large number of privacy system scrambles are shown in figures 47 to 67 which are segregated at the back of this report. In so far as possible, these spectrograms were obtained with actual working models or systems. In some cases they were made with a laboratory setup simulating the systems under scrutiny. In a few cases also the illustrations were made by cutting up spectrograms and rearranging the parts. It should be noted in these latter cases that the boundaries are



The spectrogram made with the narrow filter shows all the odd harmonics of the 60-cycle input to a special harmonic generator.

Below this is a portion of a spectrogram made with the wide filter. The striations represent a beat note of 120 cycles.

At the left is a portion cut off and inverted. The fact that the harmonics can be lined up in this as well as other shifted positions, illustrates the linearity of the frequency scale.

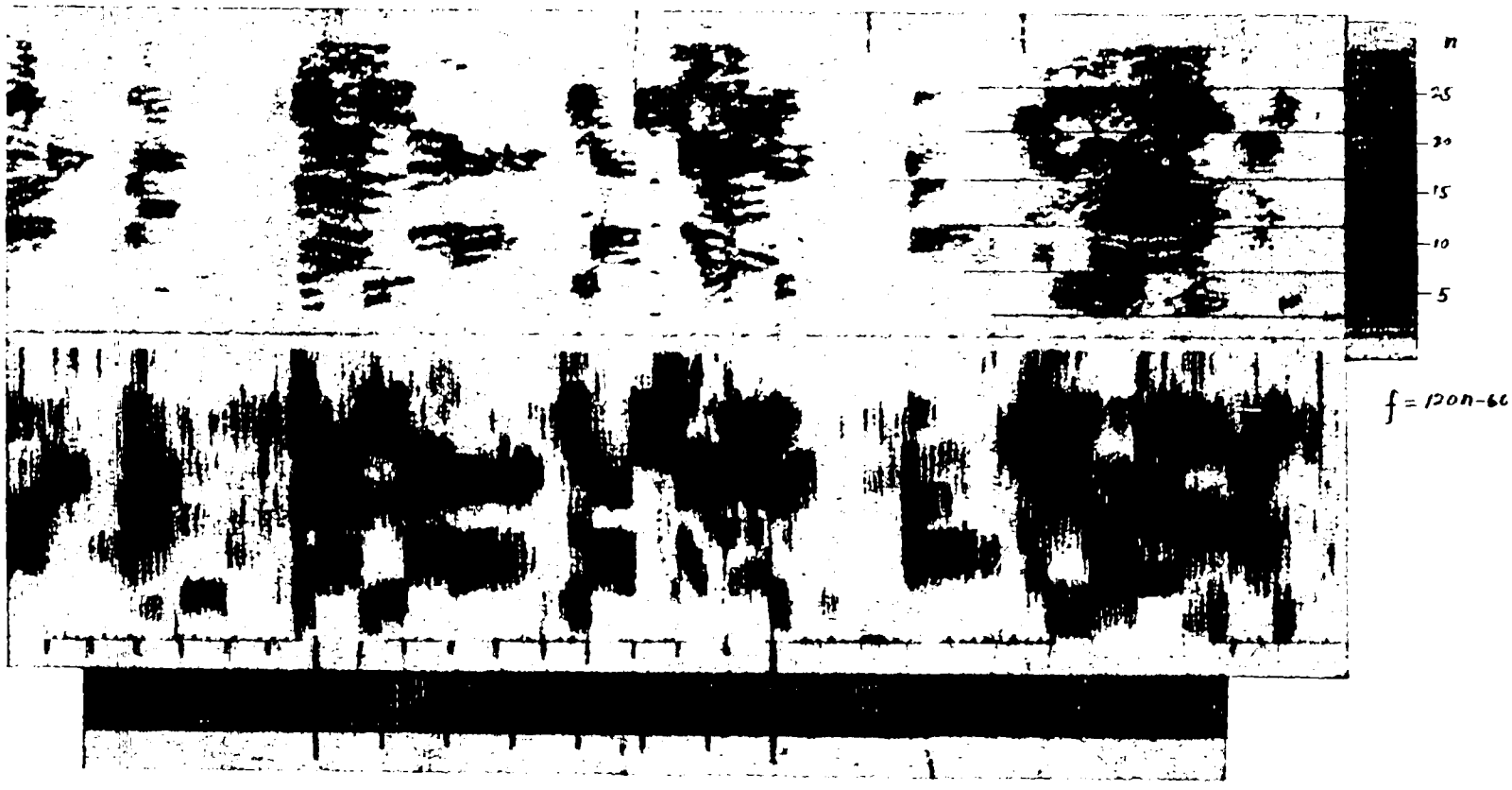
At the right is a portion cut off and shifted downward by one component. Since they are odd harmonics, the base line will fall exactly between two harmonics if it represents exactly zero frequency.

If the power frequency is exactly known, both the time scale and the frequency scale are determined by the two patterns above.

If the power frequency is not known, the time scale factor can be determined by the equation $K = \frac{L R}{T}$ inches per second, and the frequency by $F = \frac{K S}{N}$.

L = Total length of the spectrogram (circumference of the recording drum).
R = Number of rotations of the drum in T seconds.
S = Number of striations in N inches.

Figure 25 - Calibration of Spectrograph Scales



Upper Spectrogram, 45 cycle Filter; Lower Spectrogram, 300 cycle Filter.

The frequency boundaries are determined by comparing them with the harmonics of the calibrating wave. These are all 120 cycles apart, but the lowest is only 60 cycles from the base line. The frequencies of the harmonics, therefore, are given by the above formula.

The element length is best determined by using the 300 cycle filter which gives sharp time boundaries, comparing them with the striations obtained by making a spectrogram of the calibrating wave with the 300 cycle filter. Each one represents 1/120th of a second. Ten of the above elements cover 70 striations. The length of each element is therefore 1/10 x 70/120 seconds.

Figure 26 - Time and Frequency Measurements

unnaturally clear and sharp because in practice any discontinuity causes a transient which tends to obscure the true speech along the boundaries.

It will be noted that some of the spectrograms in the illustrations were made with the 45 cycle filter and some with the 300 cycle filter depending on what features were to be brought out most clearly. The spectrograph with which these illustrations were made was equipped with the control circuits mentioned in Chapter II and described in greater detail in Preliminary Report No. 27. The illustrations are therefore clearer and sharper than those included in Preliminary Report No. 25. Furthermore, a larger number of privacy systems are included than in the latter report.

Each of the illustrations contains not only reproductions of spectrograms, but also written material describing the features whereby the different scrambling systems can be recognized. It was intended that these illustrations should be self-contained in so far as possible for easy reference.

It will be noted that in some cases the spectrograms alone are not sufficient to determine the exact nature of the scramble. Certain systems completely destroy the typical harmonic structure of speech leaving structureless patterns which cannot be interpreted. This indicates a distortion of the wave form. One of these systems, which had a repeating code and a synchronizing pulse, could be resolved by the oscilloscope as shown in Chapter VI (figure 46). No general rules, however, can be given for diagnosing this type of system.

3. Systems Not Illustrated

Examination of Table I shows that there are a few scrambling systems which are not represented in the illustrations. These will be discussed in the following paragraphs. In most cases, the appearance of the spectrogram pattern which would result can be visualized by analogy with other systems.

The phase reversal system (A4) will produce a scramble indistinguishable from the multiplication system (H1) provided the phase reversals occur at irregular intervals and about as rapidly as the crossovers in the coding wave involved in H1. It is believed, but not known for certain, that they would have to occur about that often in order to make speech unintelligible.

The split phase system (A5) involving carriers 90 degrees apart was tried out in the laboratory. The output appears just the same as if two speech channels, or a speech channel and an interfering noise, were simply superposed and then modulated with a single carrier.

The stepped displacement system (B2) is rather easy to visualize. There will be time boundaries, with two or more discrete conditions of displacement. Obviously, there are a great number of possible sequences, including the possibility of some of the conditions consisting of inverted displacement.

The irregular wobbled displacement (B4) will of course be similar to B3 except that the wobble pattern will not be as simple.

The continuously varied reentrant displacement (C2) is practically impossible to simulate artificially as was done with C1. If C1 is thoroughly understood, however, the appearance of a wobbled instead of stepped reentrant condition is not difficult to visualize.

Non-repeated code TDS (F3) will have the same general appearance as repeated code TDS. It may or may not have the synchronizing pulse. There will of course be no regularity in the patterns such as was pointed out in F2.

TDS plus inversion (G1) is not difficult to visualize. Some or all of the elements might be inverted, as in A3.

The systems listed in G5 and G6 will both show equally spaced time boundaries corresponding to the length of the elements. In G5, the harmonics would be spaced much farther apart than in normal speech, and show greater slopes and curvatures. Alternate elements would show rather consistent differences in frequency distribution and in the degree of slope or curvature. In G6, the harmonics would be spaced abnormally closely, and show very little slope or curvature. Words and spaces would be very long. There would be a horizontal boundary in the middle of the band, and the patterns in each half would appear like complete spectrograms, with vowel and consonant structures apparent. In both of these systems, if the elements were cut apart, they could be rearranged to form continuous speech with the time and frequency scales compressed or expanded from the normal condition (see Chapter VI, Section 2).

Level modulations (H2 and 3) would hardly show up in spectrograms because of the level compression incorporated in the spectrograph. This has been verified experimentally.

In J1 and 2, if the noise were sufficient to mask the speech effectively, the speech could not be seen in spectrograms. J3 is easy to visualize, as is also J4. If the noise spurts are sufficiently close together, however, they may produce a pattern like H1. J5, as far as is known, exists only on paper.

In vocoder types of scrambling systems the spectrograph would show only the channel signals, which might be either amplitude or frequency modulated. For this type of scramble, oscillograms of the wave form of each separate channel signal provide the best means for diagnosis and for decoding. A sample of such oscillograms, which was obtained from an actual vocoder system, is shown in figure 42 in Chapter VI on decoding methods. The various methods of scrambling such signals (K1, 2, 3, 4) will produce discontinuities in these traces which are easy to visualize. A sample of K5 has not been available.

Channel mixing (L3) can be done in various ways and at various speeds. It will not be very easy to recognize if done rapidly. No actual systems are in use, as far as is known.

It is felt that the above illustrations and discussions cover the known scrambling methods fairly thoroughly. It is hoped that with their help any system which might be encountered in the future can be recognized. Decoding, of course, is another matter, which forms the subject of the next two chapters.

CHAPTER V

NONCRYPTOGRAPHIC TOOLS AND METHODS

Beginners in the study of privacy systems never fail to be amazed at the difficulty of scrambling speech sufficiently to destroy the intelligence. The ear can tolerate or even ignore surprising amounts of noise, nonlinearity, frequency distortion, misplaced components, gaps, superpositions, and other forms of interference. We can therefore very often obtain partial or even complete intelligence from a privacy system by partial or imperfect decoding, and this in turn can often be accomplished by operating on the scramble in some way which the designer did not contemplate.

Incidentally, the fact that the ear is such a good decoding tool in combination with these noncryptographic methods makes the production of privacy systems very difficult. Scrambling systems which look very effective on paper sometimes turn out on trial to degrade the intelligibility very little, although the scrambled speech usually sounds unpleasant. Most methods if they are pushed to the point where they do succeed in hiding the intelligibility are impossible to restore with good quality. There are in fact very few speech privacy systems which achieve a high degree of privacy with acceptable quality.

These noncryptographic methods are very important, because they may reduce the delay in obtaining the intelligence substantially to zero. Furthermore, they may render completely futile the most elaborately irregular code changing systems which could be handled only with the greatest difficulty by straight cryptographic methods. A number of noncryptographic methods are given below. Some of them of course, result in poor quality, but the saving of time, labor and equipment may be very great. Each of the noncryptographic methods has been given a designation which appears in Tables I and II and at the beginning of the following paragraphs in which they are discussed. These designations should not be confused with the designations appearing in the text which denote privacy methods.

1. Captured Set or Functional Equivalent

a. With many privacy systems all that we need in order to listen in is a captured set or its functional equivalent built from knowledge of the scrambling method. An extreme example of this is simple inversion. In this case the scrambled speech is quite unintelligible to direct listening, but if we know it is inversion, we can find the inversion frequency very quickly by trial. Another example is the split phase system (A5). The phase shifting network in the captured set could readily be adjusted to demodulate either of the two overlapping sidebands.

b. Slightly more complicated systems are those with a simple program. Again with a captured set or its equivalent it is usually easy to find the program by trial. The only possible difficulty is in keeping step with the sending end, particularly if there is no synchronizing pulse. An example of this is a wobbled band displacement (B3). If for instance, the wobble is sinusoidal, with the frequency and the sweep limits known, the problem is to keep in synchronism. In this connection a device might be mentioned which is familiar in gun-fire control circles, namely "aided tracking". With this system changes in both frequency and phase are made simultaneously. This is illustrated in figure 27. Suppose we find ourselves slightly out of step with the signal. By rotating the adjusting handle forwards or backwards we can get back into step. Suppose this adjustment was in the forward direction. The fact that we had to catch up is an indication that the motor is slow. Therefore, some of the motion of the handle required for catching up is used by means of gearing to change the frequency driving the motor. The gear ratios are chosen to suit the particular problem. With this method it is possible to get

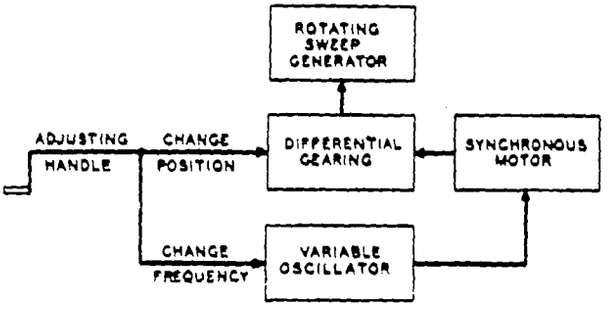


Figure 27 - Illustrating Aided Tracking

into step with and stay in step with systems such as alternate displacements and regular wobbles.

2. Compromise Decoding Methods

The methods outlined in this section have all been tried, at least in the laboratory. Their success, however, naturally depends to some extent on the switching rates and similar variables. It is possible, therefore, that a method might prove unsuccessful against a scrambling system which seems to be in the same general class as the one that was tried in the laboratory.

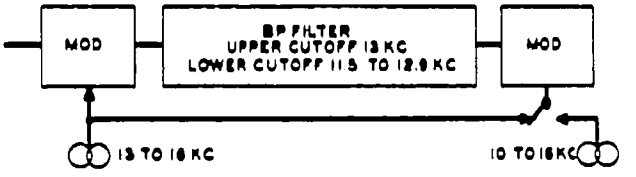


Figure 28 - Band Shift Filter

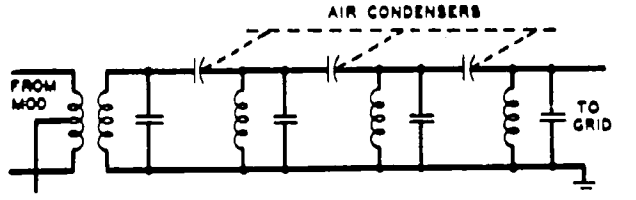


Figure 29 - Variable Band Pass Filter

a. Take for example a system (A2) which involves inversion about a number of frequencies in succession. If these frequencies are not too far apart we can choose a single frequency somewhere in the middle range and demodulate the whole signal with this one frequency. The resulting band will be right side up, but displaced by varying amounts not exceeding half the total range. This has been found to be quite intelligible, provided the switching rate is not too high or the range of frequencies too wide.

b. With some systems it is expedient to listen to only a portion of the frequency range rather than the whole range. An outstanding example of this is the system in which the subbands are variously delayed (F1). Conceivably, these delays could constantly be changed with time according to a never repeating program. This, however, would be futile because with a band filter, we need only listen to one of the bands, disregarding the others. Unless this band is very narrow the intelligibility may be practically complete. Similarly in band splitting systems if the switching is not rapid (D1) we can follow one of the bands around the frequency range. The lowest or second lowest band is usually the best. Another example is the tone sequence (J3); instead of trying to filter out one tone at a time as it occurs, we can leave all the filters in all of the time and still have enough speech coming through to yield the intelligence.

A special case in which the rejection of a part of the frequency band of the scramble makes decoding easier concerns those systems such as A5 which depend on carrier phase to mix and then separate components. There is no phase requirement imposed on the demodulating carrier unless both sidebands are transmitted. Therefore, either sideband of such a system may be suppressed with a filter, and the remaining sideband demodulated with a carrier of any phase. The two signals in the sideband will then be simply superposed.

For purposes such as those outlined a valuable tool is the band shift filter illustrated in figure 28. With this device a band of adjustable width can be taken from any portion of the signal frequency range (0-3000) and relocated in any other portion of the same frequency range either straight or inverted. One form of band shift filter is described in Preliminary Report No. 11. It consists essentially of a double modulator, such as was described in Chapter III, but with a band filter of variable width. If the frequency location of the band is not to be changed, the switch in figure 28 should be in the left-hand position. One form of variable band filter is shown in figure 29. This tool has also proved useful in certain other systems such as the multiplication system (H1) and the TDM system (E1).

g. Sometimes it is expedient to listen to a scramble only part of the time. Some of the simpler coding programs can sometimes be broken down in this manner by trial. For instance, if a coding cycle has N elements we can listen to every Nth element and make whatever adjustments are needed to make this sound natural. We can then listen to the next adjacent element and

adjust the sytem so that these elements blend properly. This attack applies for instance to a system in which several different displacements are used (B2). A captured set, of course, is the easiest way of selecting every Nth element because it is usually easy to make the other time elements inoperative.

g. Another useful device is the limiter, or peak chopper. In this same class is the compressor. These are illustrated in figures 30 A, B, C. They all tend to equalize the successive lobes of a complex wave. The peak chopper simply chops off any peak which exceeds a certain instantaneous voltage. The compressor operates more gradually and leaves the waves well rounded. If straight speech is put through any of these devices, distortion products are generated because the wave form is radically modified. It is found, however, that this kind of distortion damages the intelligibility very little. These devices should be useful against any privacy system in which sudden changes of level occur. A good example is the subband level modulation system (H3). A separate limiter or compressor in each of the sub-bands will tend to smooth out the level variations and make the speech intelligible.

g. Another nonlinear device is the rectifier. Two forms are shown in figures 31 A and B. The rectifier as used here should not be confused with the detector. The latter device also rectifies but then it has a time constant incorporated in the output circuit which tends to smooth the output and give the envelope wave. The rectifying action which we want here simply takes all the negative lobes of the signal and turns them over. As in the case of the limiter, straight speech put through a rectifier of this type is about 95 per cent intelligible.

In the privacy system designated A4 the phase of the speech signal is reversed at short irregular intervals. If this signal is now rectified, all the negative lobes will be made positive and the resulting wave will be indistinguishable from rectified straight speech except for slight discontinuities at the points where the reversals occurred in the privacy system. This is illustrated in figure 32. Therefore, a simple phase reversal system, no matter

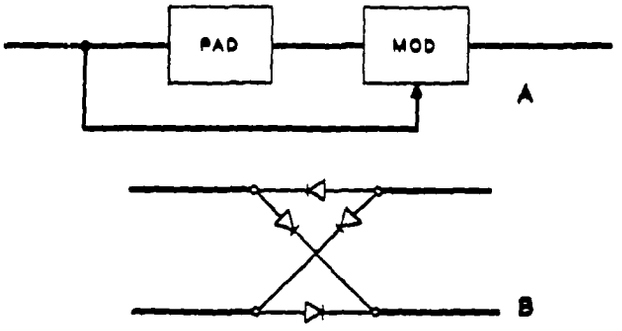


Figure 31 - Rectifiers

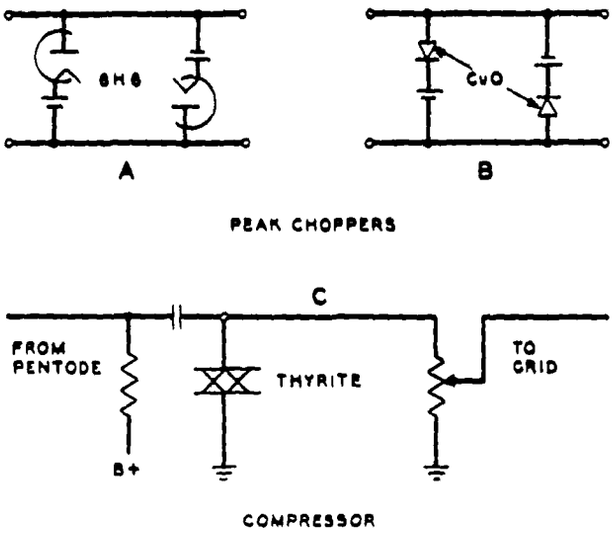


Figure 30 - Peak Choppers and Compressor

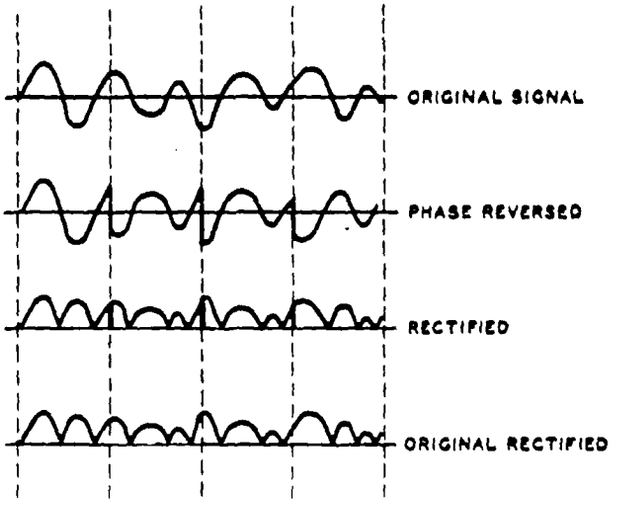


Figure 32 - Illustrating Action of Rectifier

how irregular, should yield to rectification except that distortion in the transmission process tends to change the wave form and thereby degrade the quality of the resulting speech. It should be noted that the multiplication process (H1) also results in a phase reversal every time the coding wave passes through zero. It has been found that rectification tends to make this kind of scramble more intelligible also.

f. A very useful noncryptographic device is superposition. For instance, suppose we had a three-channel mixing system such as L1 or 2. If we simply listen to all three channels simultaneously we will hear three conversations at once, or possibly one conversation with two noises superposed. Experience has shown that under such conditions it is usually easy to concentrate on the desired channel and ignore the others.

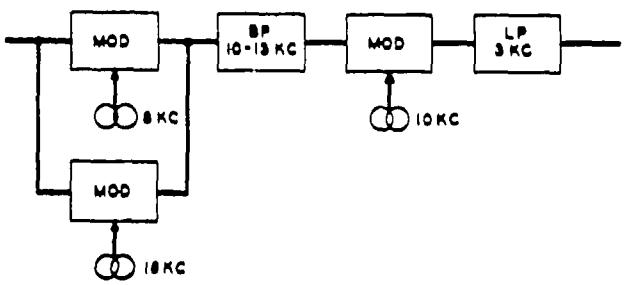


Figure 33 - One Form of Superposition Decoding

Another form of superposition is illustrated by the following: Consider a split band system (D2) in which 6 different codes are used in a never-repeating sequence. This would be rather difficult to handle by cryptographic means. Suppose, however, we had 6 separate decoding units, each set to decode one of the 6 codes. If the scrambled signal were fed into all 6 of these decoders simultaneously, one of them would always have straight speech in its output. The other 5 would be scrambled. If these 6 outputs are all superposed, we will hear straight speech with 5 scrambles superposed. This straight speech can be understood quite easily. It will be noted that the unwanted components in this kind of superposition are derived from the wanted components, and always vary in level simultaneously with the wanted components; it appears that under these conditions they do not do much damage.

The split band equipment illustrated in figure 11 of Chapter III is adapted for this kind of superposition. A multiplicity of cross-connections is made from each of the bandpass filters to the output modulators whereby each of the bands in the signal is placed in the desired bands in the output. Steps should be taken of course, that these cross-connections do not interfere with each other. An amplifier after each band filter, for instance, will perform this function. Figure 33 illustrates a simpler case of superposition applied to a system using 2 band shifts (B2).

It may be noted here that superposing time-displaced elements does not appear to be successful. For instance, if all the segments of the commutator in a TDS machine are connected to all the pole-pieces, the output will be straight speech with several scrambles superposed. This has been found to be completely unintelligible.

g. In certain cases which have been met in Project C43 the privacy sets are equipped with dials or similar means which were intended to provide an easy method for obtaining a large number of different codes. In some cases the different codes may not be sufficiently different to be mutually private. That is, while there may be literally millions of different combinations, it sometimes happens that there are thousands of combinations which will decode material scrambled with one of the combinations. Various degrees of quality, of course, will result from these partial or incorrect decoding operations. However, as long as intelligibility can be extracted the codes cannot be considered mutually private. In such cases it is possible with a captured machine to simply manipulate the dials systematically or unsystematically and listen to the result. When the speech begins to sound somewhat natural, systematic trials of each dial in turn will sometimes steadily improve the quality. Something of this sort could be done with simple TDS systems also, except that the use of interlaced codes makes this somewhat more difficult.

h. In certain cases where there are a large number of codes possible but only a few of these codes are good codes from the standpoint of direct listening, it would seem reasonable that any code applied to the scrambled signal should turn the good code into a poor code. In the 5-band split band system for instance, there are some 3840 possible codes but only 12 or so are considered really good. Any code in the decoding machine, therefore, should decrease the privacy for direct listening. This has been tried in the laboratory but has not been pushed to the point of determining whether it could

compete with the superposition method. It is mentioned because the idea may possibly apply to other systems which may be encountered.

i. A very specialized device, which applies to wire line communication only, should be mentioned here because it is not very well known. It distinguishes between the two directions of transmission over wires. In the masking privacy system J2, for instance, the clear signal in one direction is masked by noise sent in the other direction. The device illustrated in figure 34, however, discriminates against the noise, allowing the speech to be heard. It requires a small series resistance, which is built up by a step-up transformer to the line impedance. The secondary is connected to the other side of the line. The direction of discrimination depends on the phasing of the transformer windings.

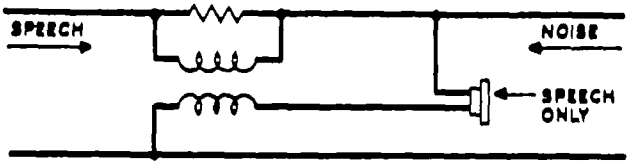


Figure 34 - Directional Discrimination

3. Automatic Decoding

Whether speech is intelligible or unintelligible is a purely subjective matter. However, the method of making speech unintelligible involves making physical changes in the speech wave. Certain kinds of physical changes can be detected quite readily by objective means and utilized to undo the scramble automatically. Obviously, the most elaborately irregular code program is completely futile if this kind of decoding can be applied.

a. A very simple example of this is shown in figure 35. Suppose the system consists of short spurts of noise applied in an irregular manner. It has been pointed out that the noise must be high in level compared to the speech in order to mask the speech. Therefore, if the signal is applied to an amplifier detector, connected to a relay (or electronic switch) the relay can be so biased that it operates only on the noise spurts. The receiver is momentarily disconnected from the line whenever a noise spurt

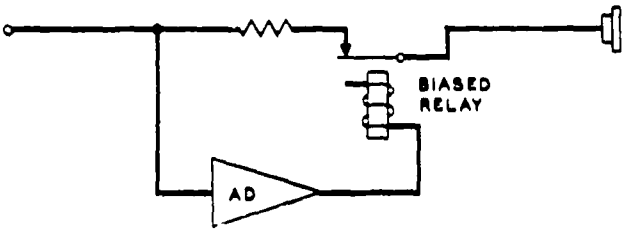


Figure 35 - Automatic Decoding - Total Energy

occurs. The same method can be used for level modulation systems (H2 and H3). Instead of disconnecting the receiver, the high level portions of the signal cause the receiver to be connected to a parallel path containing the required amount of loss to equalize the levels. In the case of subband level modulation, (H3) of course, a separate device of this type must be used in each subband.

b. The system just described operates on a total energy basis. Sometimes it is possible to obtain a switching signal on the basis of energy frequency distribution. Consider for instance, the system using two different displacements (B2). The alternate positions of the speech band are illustrated in figure 36. In one position, the band is right side up and occupies the range from 2 to 5 kc. The alternate position is inverted occupying the range from 3 to 6 kc. Since most of the energy in the speech band is concentrated in the low frequency part of the original spectrum most of the time, the system illustrated in figure 37 can be used to decode this material automatically. The signal is applied to two band filters, one covering the range 2 to 3 kc, the other passing 5 to 6 kc. The outputs of these band filters are rectified individually and fed to the two windings of a

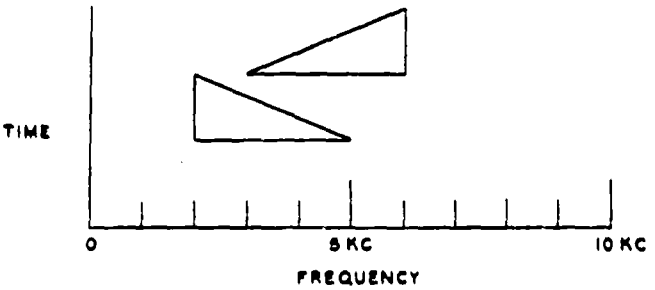


Figure 36 - Sidebands in Two Position Displacement System

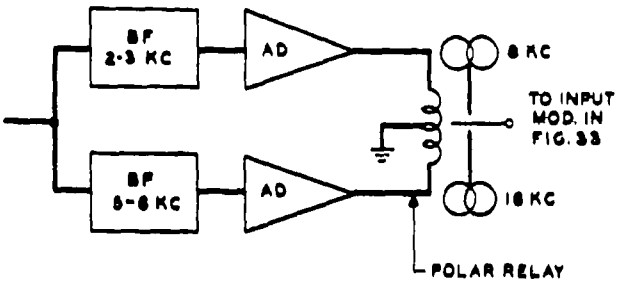


Figure 37 - Automatic Decoding-Energy, Frequency Distribution

polar relay. Obviously, the relative energy in the 2 band filters will be different for the 2 displacements and the relay in figure 37 will be operated alternately in the 2 directions thereby automatically connecting the proper carrier to the input modulator in figure 33 to put the speech band in its normal position. Obviously, this will not be infallible but with displacements as different as the ones used in the illustration, it should operate sufficiently well to yield most of the intelligence of the message. Naturally, the smaller the physical difference between the 2 positions being distinguished, the more false operations there will be. However, this method is instantaneous even with an irregularly switched system, whereas cryptographic methods would be very time-consuming.

g. Another variation of this general technique might be mentioned for the sake of completeness although it is somewhat more speculative. Consider a privacy system which depends on speed changes (F4). Changes in speed cause changes in the pitch of the voice. Suppose we apply this signal to a circuit which measures the voice pitch. This technique has been worked out in connection with the vocoder. The output of this circuit, which is a varying frequency, is used to change the speed of a motor. The motor is part of the drive of a magnetic tape recording and reproducing system through which the signal is passed. As the motor speed is made to change, the tape speed changes in such a direction as to tend to keep the derived frequency constant. This takes out not only the speed variations, but also the voice inflections. However, a monotone is quite intelligible.

h. The following method, which has not been tried out, is intended to apply to irregular band displacements or wobbles (B4), which would be exceedingly difficult to handle any other way. Consider a system in which the band is kept right side up, but is wobbled over a range sufficient so that demodulation with some intermediate carrier frequency will not give an intelligible signal. Suppose the wobble follows an irregular, nonrepeating program. The following decoding method is proposed.

The signal is impressed on a network having a very steep rising loss characteristic. If the speech band were not wobbled, this network would simply tend to make the lowest harmonic of all voiced sounds the strongest component. With the wobble, the same thing will be true except that the level of this component will undergo severe fluctuations. Therefore, the resulting signal is subjected to some form of automatic volume control and also a limiting action, tending to derive a single frequency. Forgetting voice inflections for the moment, this derived frequency would fluctuate up and down (in frequency) in step with the band wobble. In fact, it could be used as a subcarrier in a double modulation decoder to demodulate the signal to approximately the correct position in the frequency range. It will be in error, however, by an amount equal to the instantaneous voice pitch. Possibly this amount of error will not prevent the signal from being intelligible (we know that this amount of displacement does not destroy the intelligence of otherwise normal speech).

If it is desired to correct for this error, two methods suggest themselves. One possible method is to subtract from the derived frequency, by a modulation process, an amount equal to the average pitch of the voice being monitored. This will leave a small fluctuating error. Another possibility is to derive the actual instantaneous voice pitch, by difference tone methods as in the previous section, and subtract this amount from the derived subcarrier frequency.

If the displaced band is inverted instead of right side up, a similar procedure can be used, with a network of opposite loss characteristics. Obviously, this method in either case will correctly demodulate only the voiced sounds, but experience suggests that this is sufficient. If not, some kind of carryover effect might be incorporated in the system to prevent sudden changes in the subcarrier frequency, and thereby tend to hold over correct demodulation for short unvoiced sounds also. As mentioned above, this method has not been tried, but is felt to be worth recording because of the great difficulty of handling irregularly wobbled systems by any other method.

e. Another rather speculative automatic method might be mentioned because some form of the method might prove useful against certain multiplication systems such as H1. The code wave in the particular case encountered, was repeated many times per second, and there was a synchronizing pulse ahead of each cycle,. If the signal is applied to a synchronized cathode ray oscilloscope with a highly persistent screen, a definite pattern appears because the coding wave always passes through zero at the same time. Also, the speech energy tends to average out after a few cycles so that the pattern reflects the amplitude of the coding wave. It is quite conceivable that this pattern on the screen could be scanned optically and used to generate a decoding wave for automatically unscrambling the signal. Obviously, if the coding wave is changed periodically, a new decoding wave is automatically produced. The only requirement is that the coding wave persist long enough to form an average pattern on the screen.

f. Another variation of automatic decoding methods might be termed "parallel automatic" because two or more complete decoding units are used in parallel but only the correct one is applied to the listening receiver. To emphasize the difference between this method and the one previously discussed, we will use the same example, namely, the system with two band displacements. Referring to figure 38, suppose instead of the parallel modulators, there were two complete units in parallel including the band filter, the second modulator and the output filter. One of the units is fed with the 8 kc carrier, the other with the 16 kc carrier. Each unit will have straight speech in its output half the time, and the other half of the time will have inverted speech displaced by 1,000 cycles. A 1,000 cycle low pass filter can then be used in a device similar to figure 37 to switch the listener to whichever one of the decoding units has the straight speech. For the particular system used in the illustration, there does not appear to be any particular advantage of one method over the other. However, the latter system can be applied in cases where the other method might not be feasible.

g. The parallel automatic method can be made to give a different type of switching signal. For instance, we might make use of the harmonic relationship between the components of speech when the speech band is in its normal position. If the voice pitch happens to be 100 cycles say, then all the harmonics will be multiples of 100 cycles. If this speech is put through a suitable nonlinear system such as a rectifier or limiter, difference tones will be generated which will also be multiples of 100 cycles. If, however, the speech band is displaced from its normal position in any way, the difference tones will not coincide with the speech components. If, for instance, the whole band has been displaced by 50 cycles, then the speech components will be 150, 250, 350, etc. The difference tones generated by a nonlinear system, will be 100, 200, 300, etc. If we now take a second difference between the output of the nonlinear system and the original components, there will be generated multiples of 50 cycles. The lowest component of this series will be lower than the pitch of the voice. This will be true regardless of how far the original band has been shifted, except for the special case where the shift happens to be an exact multiple of the voice pitch. Since, however, the pitch is constantly varying, this coincidence is of very brief duration. Theoretically, at least, a low pass filter with a cut-off lower than the normal range of voice pitch can be used as a clue to determine whether a speech band is in its proper location. The method then would consist in having several decoders in parallel but listen only to the one which did not generate a component in the low pass filter.

The above illustrations will serve to show the possibilities of noncryptographic types of attack on privacy systems. When a new system is encountered, this type of attack should be given serious consideration because of the saving in time and equipment. Naturally, as pointed out above, straightforward cryptographic attack can be made to yield a better quality signal. However, experience has shown that the ear can become familiar with certain kinds of distortion and learn to extract the intelligence more and more readily with practice.

In general, noncryptographic methods require that the signal, as received, be of fairly good quality. In some cases, the saving in time, labor, and equipment would be so great that if the signal, as received, is too poor to permit noncryptographic attack, the most reasonable thing to do is to move the intercept station to get a better signal.

In Table I, there is listed for each privacy system, the type of noncryptographic attack which might apply. It should be emphasized once more, however, that the method which succeeds at one switching speed may fail at another. The list, therefore, should be taken only as a recomendation of systems which should be considered. The noncryptographic decoding methods are summarized in Table II.

CHAPTER VI

CRYPTOGRAPHIC TOOLS AND METHODS

A cryptographic decoding method involves: (1) actually determining a code which will undo the scramble, (2) restoring the speech by means of this code. In the case of repeated codes, this can sometimes be done rather quickly. An example is the repeated code TDS system. The actual codes used can be found in about 15 minutes. Having found the code, we can set it into our receiving machine and thereafter listen to the speech directly. In the case of nonrepeated codes, every bit of the message must be handled individually. It may take a thousand or even a hundred thousand times as long to decode as it did to speak the words. It may take hours or even days to obtain the intelligence from a short message; meanwhile other messages will have been sent and it is obvious that we get farther and farther behind. The only way we could avoid this is to have approximately as many teams working in parallel as the ratio of decoding time to message time, which of course is impractical if the ratio is large.

1. Program Determination

The simplest cases to handle are those involving a program which can be determined directly from spectrograms by inspection or measurements. The reentrant inversion system (C1) might be used as an example. Suppose a multiplicity of displacements were used in some irregular sequence. Discontinuities marking the inversion frequencies appear in the spectrograms and once they have been determined by measurements on a large number of spectrograms, the program can thereafter be determined quite readily by using a template. This template can be marked directly with the settings of the decoding machine which will restore the speech to its normal position.

Another example involving a program would be one like B2, in which two different displacements were used alternately, with the intervals irregular in duration. Here the time boundaries will be quite apparent and they can be measured with a suitable time scale.

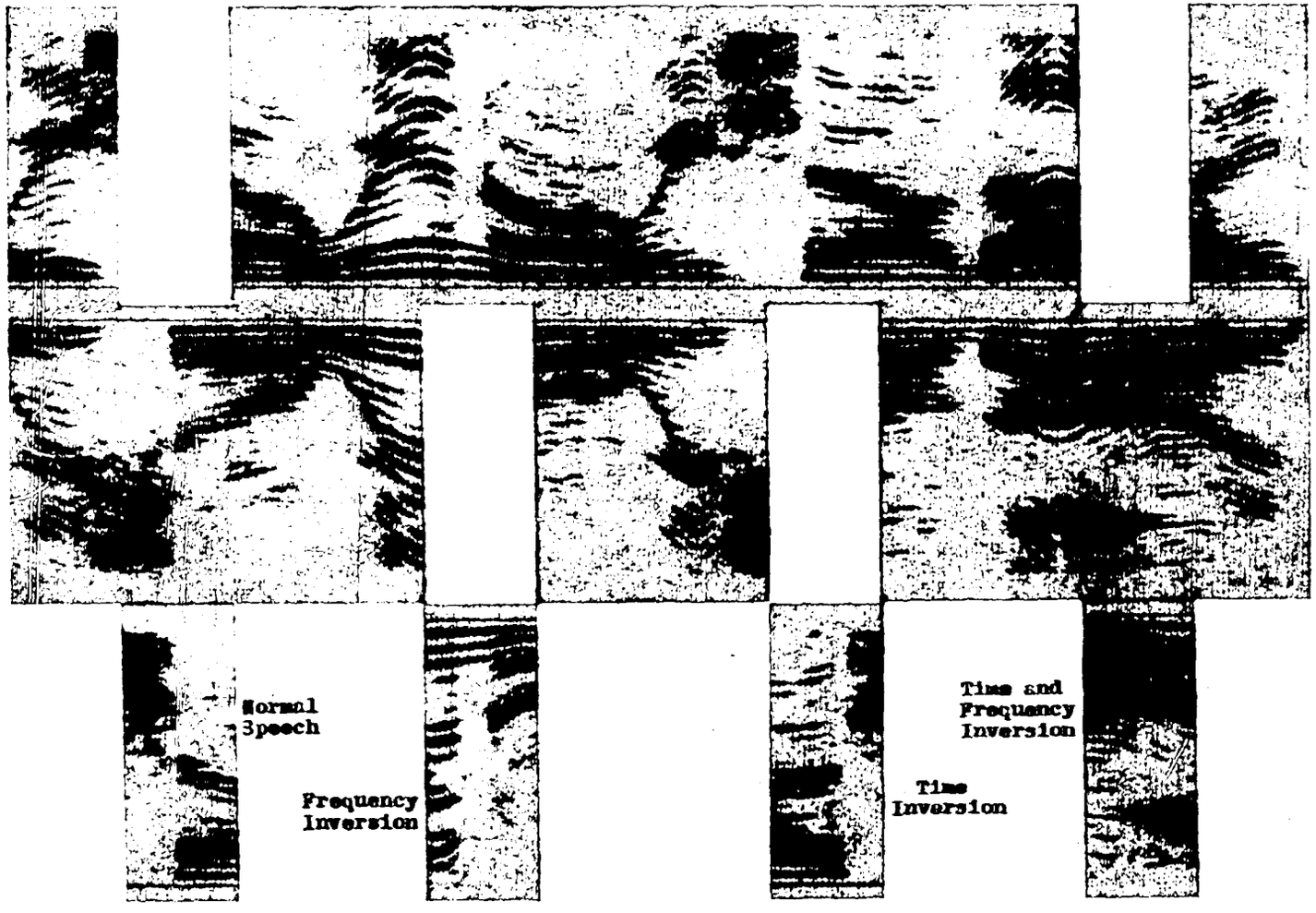
In all likelihood changes of the above types will occur in discrete steps for practical reasons. The use of a program involving continuous changes with time presents formidable technical difficulties at the authorized as well as the unauthorized terminals.

2. Matching Spectrograms

In cases where the scrambling system involves rearrangement of the speech elements in time or in both time and frequency, the basic method for determining the codes involves cutting up spectrograms along the element boundaries and rearranging the elements so as to restore the straight speech. An example is shown in figure 38. The criterion for rearranging the elements is that there should be continuity at the boundaries. This continuity includes the position and direction of the individual harmonics, the position and direction of the resonance areas, and, in general, the amplitude as represented by the darkness or lightness of the patterns. The pieces are numbered before the matching process begins and when the matching has been completed, the numbers on the pieces determine the code.

If the scrambling process involves inversion of the time or frequency scales, straight speech can be restored for matching purposes by making two spectrograms as shown in figure 39. Present models of the spectrograph include means for making a mechanically inverted pattern as well as a normal pattern. The spectrogram at the top of figure 39 shows a normal pattern. Directly below it is an inverted pattern of the same material. A mechanically inverted pattern is indistinguishable from a pattern produced by electrical inversion of the speech. Similarly if the whole inverted spectrogram is turned through 180°, so that the base line is at the bottom and towards the observer, the result is indistinguishable from the case in which the speech is transmitted backwards. Therefore, if an element in the scramble is inverted, it may be recovered as straight speech for matching purposes by cutting the element from the mechanically inverted pattern. If an element has been transmitted backwards, it can be restored to normal by cutting it from the inverted pattern and rotating it 180° as described above. If it is both backwards and inverted, it may be restored by cutting it from the regular pattern and turning it around.

Figure 38 - Matching Spectrograph Patterns of Nonrepeated Code TDS



If the scramble contains inverted elements, these will appear right side up in a mechanically inverted spectrogram.

The time scale may be inverted by rotating the elements 180 degrees. Note the positions of the base lines in the examples above.

Figure 39 - Illustrating Inversion of Time and Frequency Scales in Spectrograms

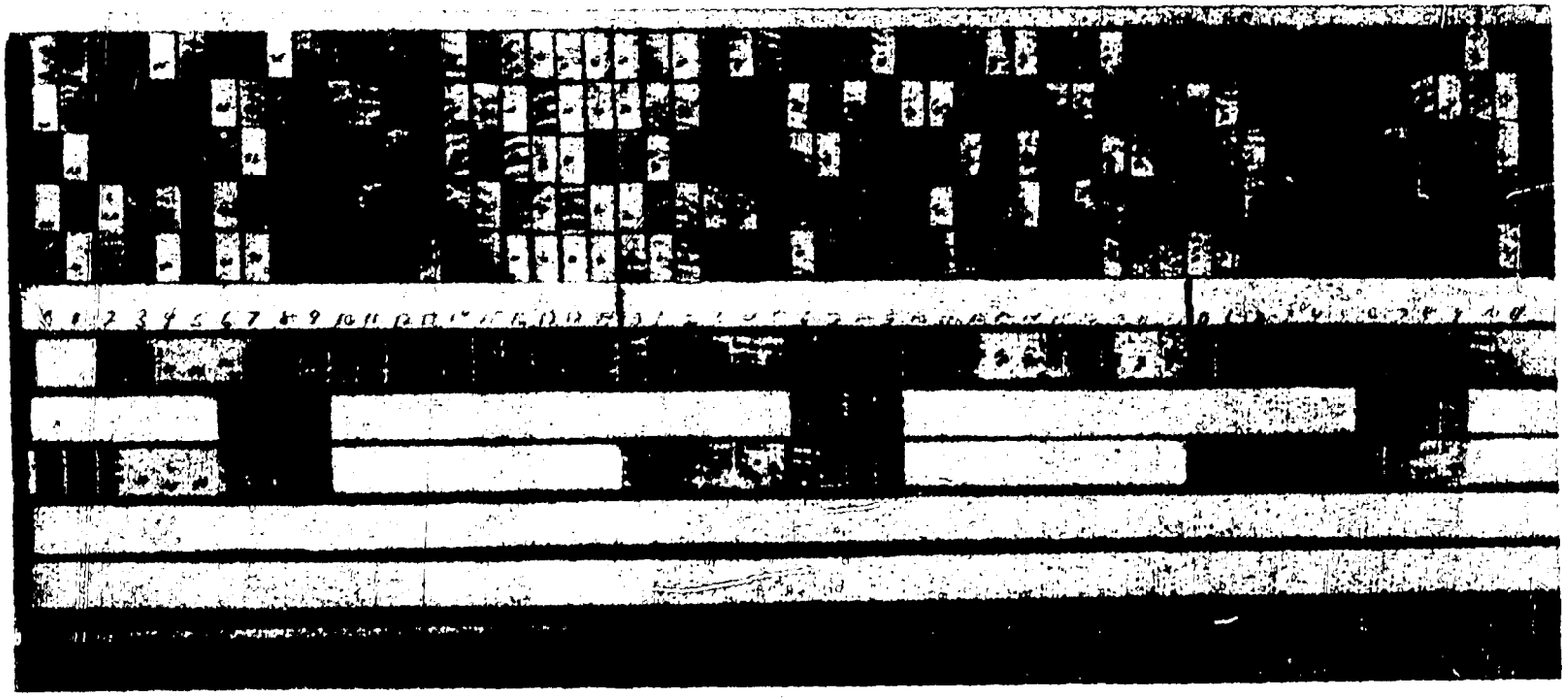


Figure 40 - Matching Spectrograph Patterns of Two-dimensional Scramble

It has been found from experience that matching is facilitated by enlarging the spectrograms by a factor of about 2 to 1. Not only is the increased size easier to handle, but the heavy photographic paper is much better to handle than the facsimile paper. The latter is delicate in texture and its surface is easily stained. In this connection it should be noted that the process of enlarging the spectrograms does not appreciably affect the decoding time in the case of nonrepeated code systems. There will, of course, be an initial delay, but in general, the matching time will be controlling. Spectrograms can be made, enlarged, and cut up faster than they can be matched. If it is found necessary, however, to use spectrograms for matching purposes regularly, then it might pay to adopt the technique described in Preliminary Report No. 13 for producing large spectrograms photographically.

To facilitate matching, appropriate means should be used for handling the elements. It has been found that a slightly adhesive surface is advantageous. In the illustration of figure 38 this surface was provided by coating the boards and also the backs of the elements with ordinary rubber cement. This is also the case in figure 40. This latter example shows a two-dimensional scramble. Horizontal strips of rubberized Bristol board were provided for matching along the time axis.

Once a system has been thoroughly diagnosed certain numerical properties of the coding process will be known. Advantage should be taken of this knowledge to supplement and check the matching process. Examples are given in Preliminary Reports Nos. 10, 14, 22 and 26.

The two examples thus far cited of spectrogram matching were artificially produced by cutting up spectrograms of straight speech, and the boundaries are therefore clear and sharp. In practice the time and frequency boundaries will be obscured by transients as may be seen in the illustrations accompanying Chapter IV. Frequency boundaries are filter cutoffs, and they are marred by overlap or underlap and by phase distortion. This, however, is not as serious as the transients occurring at the time boundaries. There is a basic difficulty here, arising from the desire to obtain a high degree of frequency resolution, which entails the use of a narrow scanning filter. The response and decay time of such a filter is appreciable in comparison with the element length in many scrambling systems. The decay time produces the more serious of the two effects. It causes energy from a strong element to spill over into the adjacent following element in the spectrogram. This difficulty is unlikely to cause trouble in any application of the spectrograph except decoding. Therefore, it is felt that means for alleviating this difficulty should be recorded here. A small amount of exploratory work has been done along these lines, but the embodiment of this improvement in a spectrograph has not been accomplished because the need was not sufficiently pressing in Project C-43.

The basic idea for avoiding the obscuring effects of spillover is to permit the spillover to take place in such a way as to be subsequently removable. For instance, suppose a sample of TDS were recorded on the tape and suppose the spectrograph were equipped with a suitable switching arrangement such that only every alternate element was reproduced. The spillover from each element would then occur in a blank area, and it could subsequently be trimmed off, leaving a sharp, clear boundary. A second spectrogram could then be made of the alternate elements, again trimming off the spillover.

A logical extension of this idea, which would save some time, would be to have two scanning filters and use them alternately by suitable switching means. Both the inputs and the outputs of the filters would have to be switched, and the two switches should be separated by the appropriate time delay to take account of the transmission time through the filter.

A third variation of this idea which requires less equipment, is to make one spectrogram in the usual manner and then make a second spectrogram with the machine running backwards. The spillover always occurs into the leading edges of the elements in spectrograms. Cutting the first spectrogram in the proper places will result in clear, sharp right-hand edges on each element, but each left-hand edge will be obscured by spillover. Cutting the second spectrogram in the proper places will give clear left-hand edges on each element. Matches could then be made between elements from the normal and backwards spectrograms, in such a way as to utilize the good edges of the elements.

In other respects it is to be expected that the patterns produced by the spectrograph can be improved. For instance, studies have been made which show that amplitudes can be represented in such a way that they can be inter-

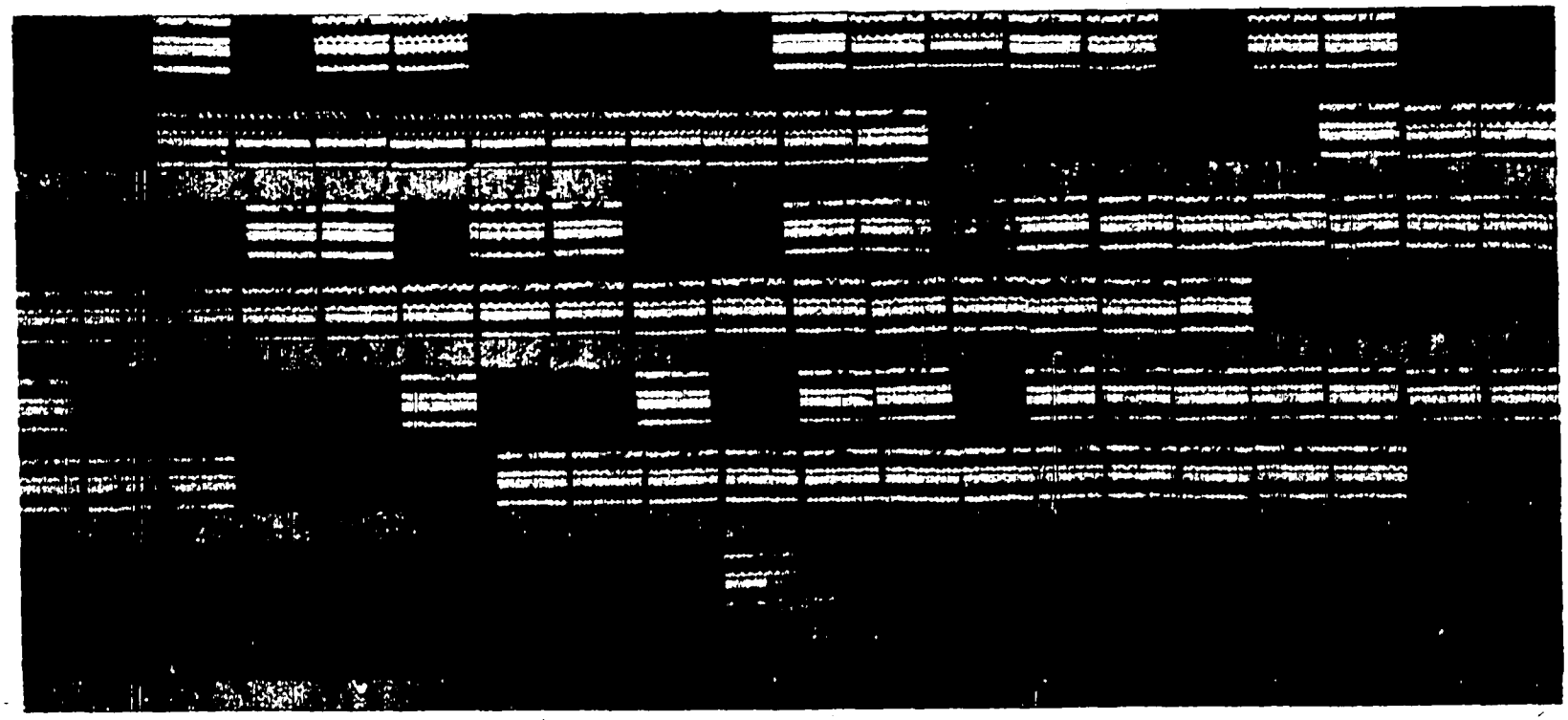


Figure 41 - Matching Variable Area Patterns of Nonrepeated Code TDS

preted quantitatively. This is an improvement over the rather indefinite shades of gray in the usual spectrograms. It would provide another criterion for matching. In some cases, however, this might be a handicap. For instance, in TDS systems the pole pieces are not all of equal efficiency. The amplitudes of adjacent speech elements are affected by this change in efficiency and they might not appear to match when they really should. This condition, of course, might be aggravated intentionally as part of the privacy feature of the system. On the whole, however, it looks as though amplitude representation should be an improvement in decoding work.

3. Matching Variable Area Patterns

For some purposes it has been found that wave form patterns offer certain advantages over spectrograms. They can be made more rapidly and they can be played back directly to reproduce the original speech. Intrinsically, wave form patterns are not as good as spectrograms for diagnosing frequency shifts and the like. However, they present the time scales more graphically and they are not subject to transients at time discontinuities such as the spillover effects discussed in the last section.

The particular type of wave form pattern found most useful was a variable area pattern similar to the sound track used in moving pictures. Variable area patterns are more distinctive to the eye than oscillographic traces. They form geometric designs that catch the eye and facilitate matching. The manner of producing and playing back these patterns is described in Preliminary Reports Nos. 1, 7, and 12. An example of variable area patterns in process of matching is shown in figure 41 taken from Preliminary Report No. 26.

Variable area patterns of this type have been found particularly good for decoding TDS systems, especially repeated code systems. It will be noted that amplitudes are clearly represented in these patterns. By matching a multiplicity of cycles of a repeating code system simultaneously, it is possible to take advantage of this amplitude representation even though the wave form itself might be obscured by other features of the privacy system. For instance, the use of split band coding was once proposed to increase the privacy of TDS systems. This combination would be much more private than plain TDS if judged on the basis of matching spectrograms, particularly if the split band codes were rapidly switched at intervals not simply related to the TDS elements. No difficulty, however, was found in matching the variable area patterns to find the TDS code. This is described in Preliminary Report No. 19. This report also describes a scheme for nullifying the effect of split band coding on the wave form. This consists of modulating all the frequency bands down into one frequency band. Changes in the split band code will then have no effect on the wave form of patterns produced in this manner.

It was also proposed at one time that the use of a whisper or monotone instead of normally inflected speech would increase the privacy of TDS systems. Again this is true in terms of spectrograms, but it was found that variable area patterns could be matched almost as easily for whispered speech as for normal speech, and with the monotone it was actually easier. This is described in Preliminary Report No. 16.

Another feature of the variable area patterns which might be useful is that the patterns have characteristic shapes. Usually they look like a series of damped oscillations with the highest amplitude at the beginning of each fundamental period. This should enable the recognition of cases in which speech is transmitted backwards. The characteristic periodicity of the patterns might also be used to recognize whether a frequency band is in its proper location.

Toward the end of project C-43 it came to be recognized that there would be considerable advantage in using a compressor in the production of variable area patterns. This tends to bring out low level sounds. The distortion of the wave forms resulting from instantaneous compression is immaterial if they are to be used only for matching. This kind of compression, however, should be sharply distinguished from automatic volume control action. The latter is relatively slow acting and it is obvious that in TDS systems, for instance, it would alter the amplitudes of certain elements in such a way as to make matches impossible.

4. Matching Oscillograms

It was stated in the previous section that oscillographic traces could be used instead of variable area patterns, although in general there will be a disadvantage. There is one type of privacy system, however, for which oscillographic traces are required, namely, vocoder systems. The signals in vocoder channels are

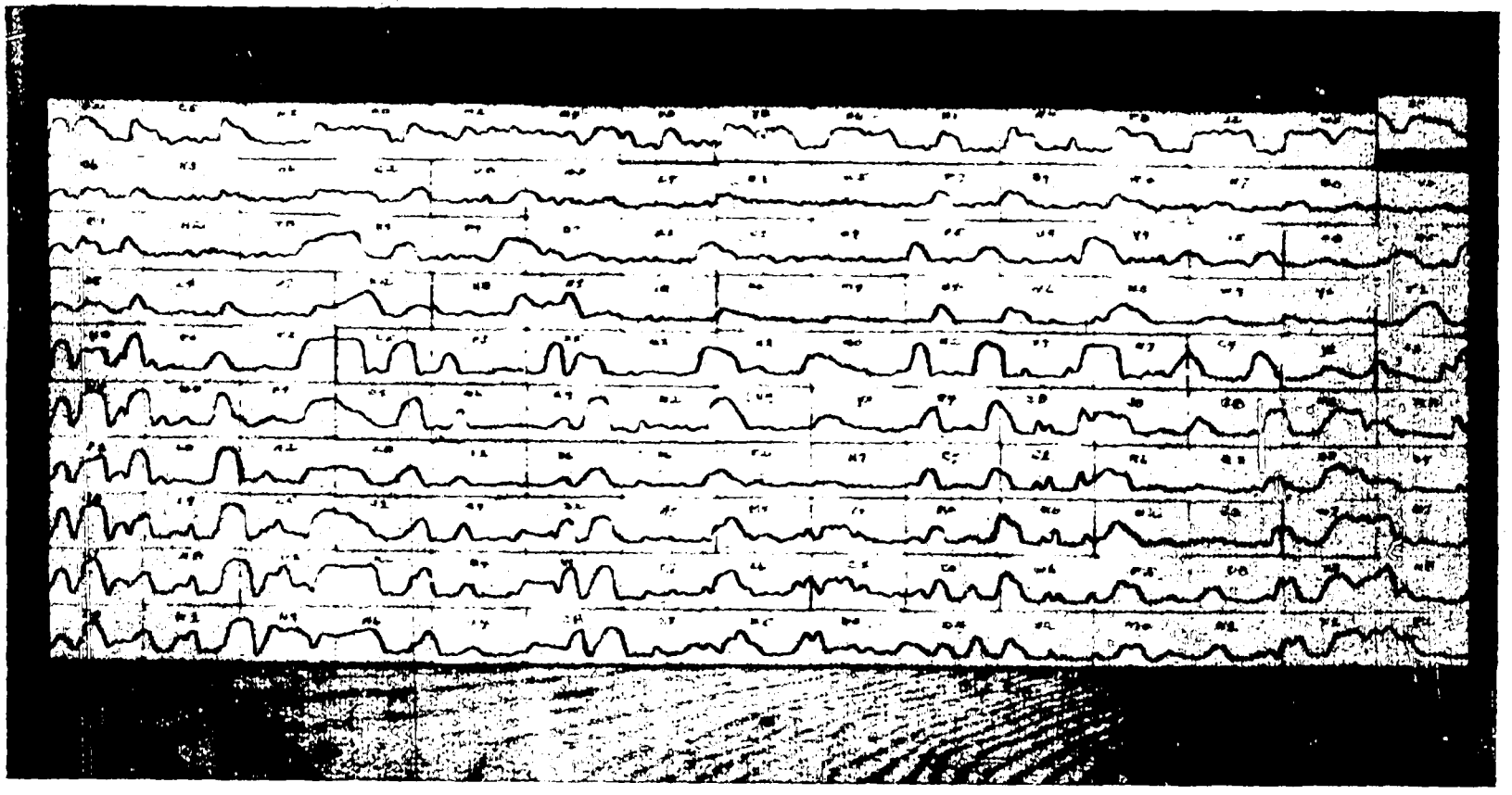


Figure 42 - Oscillographic Traces of Vocoder Channel Signals

essentially fluctuating d-c signals after they are modulated down to their normal frequency location. They can best be examined in the form of oscillographic traces. Figure 42 shows a set of undistorted vocoder channel signals. It will be noted that there is a tendency for the amplitudes to vary simultaneously in the several tracks. It has been found that if the signals in the various channels are permuted, even with the sharp edges resulting from artificially produced scrambles, the number of mismatches tends to be about 40 percent. This means that a vocoder system with its channels permuted at short intervals provides a rather difficult privacy system to decode.

It has been found that compression enhances the value of oscillographic traces of this type. Without compression the lower amplitudes are obscured by the width of the traces. Instantaneous compression makes changes in the magnitude or direction of the traces apparent in the lower level sounds. The patterns shown in figure 42 were produced in this manner.

5. Indicator Methods

In the following methods a visual indication is obtained denoting which of several possible choices puts the speech elements in their proper order. These methods, of course, are applicable only to cases where the possible number of choices is not overwhelmingly great. A natural example of a visual indication occurs in the illustration of TDS in figure 39. Whenever two originally adjacent speech elements remain adjacent in the scramble the two elements are not separated by a time boundary in the spectrogram. Elements which do not belong in adjacent positions have a boundary resulting from discontinuities in the harmonics and from spillover effects. The absence of a time boundary can be taken as an indication that the two adjacent elements belong together. To make use of this effect the following procedure is suggested. Record a sample of the scramble on a loop of tape. Reproduce this sample through a TDS machine and make a spectrogram, noting any adjacencies which occur. Change the code in the TDS machine and make another spectrogram again noting adjacencies. A systematic set of codes should be worked out in advance which explore all the possible combinations of elements. At the end of such a cycle of operations it should be possible to place a large percentage of the elements correctly. This can be applied to non-repeated or repeated code TDS.

A variation of this method, which has been suggested but not tried and which should be much faster, is as follows: Reproduce the recorded sample through a low-pass filter, say 2500 cycles. Pass it through a TDS machine and then through a high-pass filter with the same cut-off. View the output of the high-pass filter on a cathode ray oscilloscope whose sweep is synchronized with the TDS cycle. Transients will occur at the boundaries of elements which do not belong together. These will generate frequencies higher than the cut-off of the high-pass filter and will appear as pulses on the scope. The absence of a transient will indicate either that the elements belong together or that no energy was present. Again a systematic cycle of codes should place most of the elements correctly.

Another example of the indicator method is the following: Suppose in a split band D2 system 6 known codes are used in an irregular sequence, and it is desired to determine the sequence. The following procedure is suggested: Record a sample and reproduce it through a decoding machine equipped with one of the proper decodes, and make a spectrogram. Certain elements in the spectrogram will be seen to be normal speech. These elements, of course, are the ones to which the particular code applies. It is much easier to determine whether a particular element consists of straight or scrambled speech, than to determine which particular code was used. Repeat this procedure with each of the other five codes. Each element can thereby be identified with a particular code.

A variation of this procedure, which should give more positive results, is as follows. The output of the decoding machine used as above is rectified before making the spectrogram. Rectifying normal speech does not add inharmonic components, whereas rectifying speech which contains band shifts results in inharmonic components. This is illustrated in figure 43. The upper spectrogram shows rectified straight speech. This looks perfectly normal except that the frequency range is somewhat more completely covered with harmonics than is the case in normal speech. The second spectrogram shows a sequence of split band scrambles. The third spectrogram shows a similar sample rectified, with none of the elements decoded. Rectifying the undecoded elements results in a complete smear in the spectrogram compared to the rectified straight speech. Properly decoded elements will stand out more clearly against the background of rectified scrambled speech.

1 - Straight Speech Rectified

2 - Six Code Split Band Scramble

3 - Similar to No. 2, Rectified

Figure 43 - Showing Effect of Rectification on Normal and Band Shifted Speech

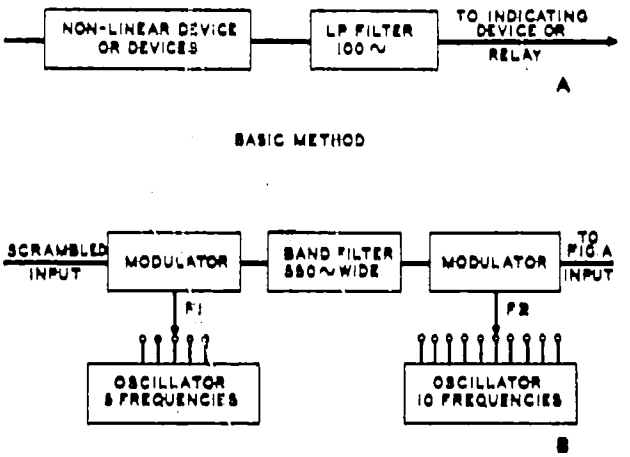


Figure 44 - Band Shift Detector

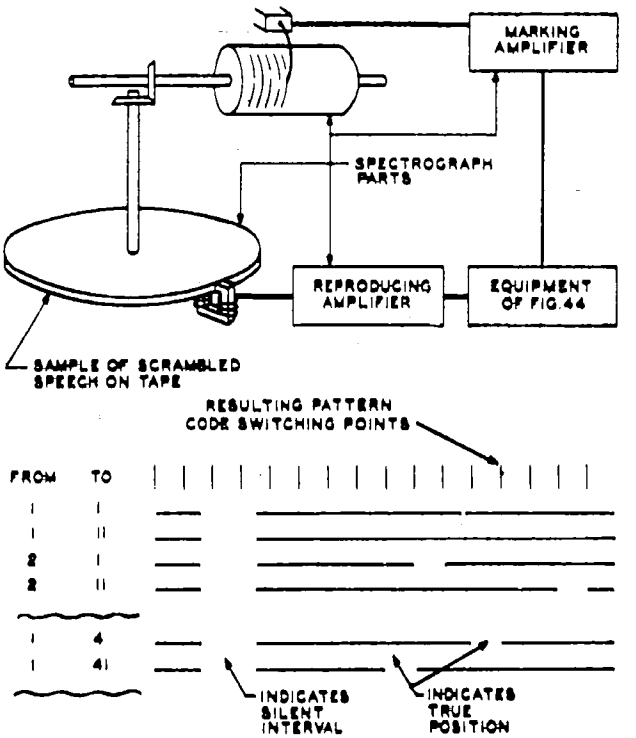


Figure 45 - Adaptation of Spectrograph for Decoding Switched Split Band Scramble

Another variation of the indicator method was touched upon in Chapter V. It consists in subjecting the scrambled speech to a nonlinear device or devices in such a way as to obtain difference tones between the components. In normal speech, in which all components are harmonically related, there will be no difference tone lower than the pitch of the voice. In scrambled speech the components are not harmonically related and there will be difference tones lower than the pitch of the voice. The output of a 100-cycle low-pass filter therefore, can be used to indicate whether a band of speech is in its proper frequency location or not. This is illustrated in figure 44A. The importance of this method lies in the fact that each frequency band can be examined separately. It might therefore be used to determine for each element in a 2 dimensional scramble which frequency band it came from.

Figure 44B shows how each band can be lifted out of the scramble and placed in each of the 5 possible positions either straight or inverted. The spectrograph might be used to speed up the analysis process as illustrated in figure 45. The output of the low-pass filter is fed to the marking amplifier. Whenever the output of the low-pass filter is zero there will be no mark produced on the paper. Whenever there is an output a mark will be produced. The procedure would then be as follows: Set oscillator F1 at one value and then set oscillator F2 successively at each of its 10 values (or 5 if inversion is not required). Repeat with oscillator F1 at each of its other 5 values. For each of these 50 settings allow the spectrograph drum to rotate 2 or 3 times with a few blank rotations between each setting. The traces on the drum will then look something like the drawing. The time axis is as usual disposed lengthwise. If all the traces in a given time interval are blank it is presumed that this represents a silent interval. Single blank intervals in otherwise continuous marks indicate that these settings were the correct ones. If none of the marks for a particular element are blank the indications are that at that particular moment a consonant occurred which of course is composed of inharmonic components. This system has not actually been tried in this complete form but enough work has been done to show that

it is possible to make use of the presence of inharmonic components in some such manner. It appears therefore that a substantial fraction of the elements in a 2-dimensional scramble might be identified as to frequency location.

One other possibility of this type might be mentioned. Variable area patterns of vowel sounds have characteristic configurations. These configurations depend on their harmonic structure, and a disturbance of this structure should change these patterns in a recognizable manner. For instance, if the components are inharmonic there will be no periodicity at the fundamental pitch rate. It might therefore be possible to use variable area patterns, which can be produced much more rapidly than spectrograms, as indicators along the lines of the above discussion.

6. Application to Table I

In this section we will examine the application of cryptographic methods to the specific scrambling systems listed in Table I. In this table the systems which might require cryptographic attack are indicated by a reference to a page in this section. The following paragraph numbers refer to privacy systems in Table I.

A4. Among the systems listed under single modulation the only one that might require cryptographic treatment is the phase reversal system. This system is a special case of the multiplication system which will be treated later.

B4, C2. Among the double and triple modulation systems, the irregular continuous displacements have not been handled by non-cryptographic methods. It might be necessary to make a continuous series of spectrograms to determine the displacements as a function of time. This might some day be done continuously and instantaneously, in which case compensating frequency changes might be made continuously by hand to decode the material.

D1, D2. Among the band splitting systems, the fixed or slowly switched codes can be solved by inspection as discussed in Chapter IV on diagnosis. If the code is rapidly switched, however, single elements seldom contain sufficient information to determine the codes. If the switching sequence is a repeated sequence, it may be worthwhile for the sake of quality to determine the sequence and get in step with it. In this case the methods described in Section 5 above should be of assistance. If the switching sequence is never repeated the indicated non-cryptographic methods appear most reasonable.

F2, F3. TDS systems yield very poorly to non-cryptographic attack. For repeated code systems, however, the code can readily be determined by matching either spectrograms or variable area patterns, taking advantage of the numerical properties of the codes. These methods are covered in Preliminary Report No. 14. Nonrepeated code systems, however, have thus far been found exceedingly difficult to handle, although the methods of Sections 2, 3 and 5 above apply. Efforts in this direction are described in Preliminary Report No. 26.

F4. Speed variations, according to some preliminary laboratory tests, are rather ineffective in masking the intelligence of speech unless the variations are exceedingly wide and rapid. Technical difficulties then become so great that this appears to be an unlikely privacy system by itself. Small variations in speed, however, might be used to make spectrograms of TDS systems more difficult to match. In this case, however, it will be unnecessary to determine the speed variation program if the TDS scramble can be removed.

G1, G2, G3. Combinations of TDS and frequency scrambles are interesting from the cryptographic standpoint. Since repeated code TDS systems were found easy to break, it was proposed to add various forms of split band scrambles. It was argued that the continuously changing frequency scrambles would alter the shapes of variable area patterns so that they could not be matched. Furthermore the changing frequency scrambles would make spectrograms unsuitable for matching, especially if the split band codes were switched nonsynchronously compared with the TDS boundaries. Each time the frequency code was switched a new vertical boundary would appear in the spectrogram, and in combination with the TDS boundaries the spectrograms would be very severely broken up in the time scale. It was found, however, as discussed in Preliminary Report No. 19, that if the TDS code is a repeated code the frequency scrambles can be practically ignored in matching variable area patterns. Having found and removed the TDS code the remaining frequency scramble can be solved by noncryptographic methods.

In the case of nonrepeated TDS, however, the addition of split band coding would increase the difficulty considerably, provided that the two coding systems do not provide clues to each other. The most promising method for

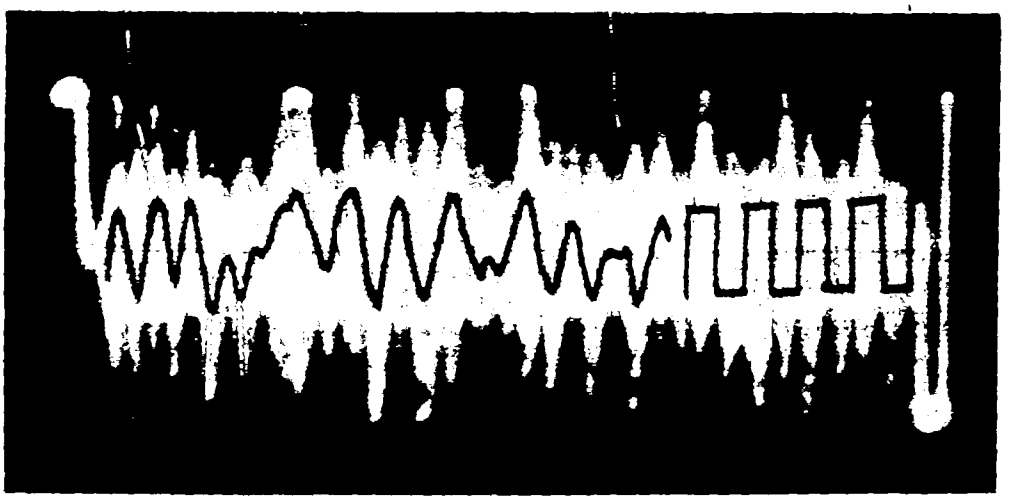


Figure 46 - Illustrating Repeated Code Multiplication System

handling this system appears to be to determine the split band codes first by the methods of Section 5 above. If the split band codes are then removed the remaining scramble can be handled as straight TDS. Another possible method is to make variable area patterns with all the decodes superposed as described in Section 2f in Chapter V. The resulting patterns, however, will not be as satisfactory for matching as patterns of straight speech.

G4. The two-dimensional scramble can be handled by matching spectrograms if a repeated code is used. Experiments along these lines are described in Preliminary Report No. 22. If the code is nonrepeated, however, it would be exceedingly difficult and time consuming to handle by unaided matching. It would help considerably if the original frequency location of each element in the scramble could be determined. This might be accomplished by the methods described in Section 5.

H1. Determining the code for multiplication or phase reversal systems can be accomplished quite readily if the code is repeated at sufficiently short intervals. In the one system which was met in Project C-43 (Preliminary Report No. 18) the code wave was repeated 100 times per second. In this case the scrambled signal could be applied to the vertical plates of an oscilloscope with the horizontal sweep synchronized with the code cycle. It is obvious that every time the coding wave passes through zero the scrambled signal also passes through zero regardless of the value of speech signal at the moment. If several cycles of scrambled speech material are superposed, therefore, they have the appearance shown in the photograph, figure 46. The superposed traces show a definite pattern, with regions of high and regions of low amplitude, and also sharp indentations. These latter are the crossover points of the code wave. There is also a marked tendency for the peaks to occur alternately above and below the center line, but the amplitudes of the peaks are not all alike. Since the speech amplitudes tend to average out over a number of cycles, the amplitudes of the superposed peaks reflect pretty accurately the amplitudes of the coding wave at those points. The probable shape of the coding wave based on this evidence, has been partly traced in.

It has been found experimentally that if only the crossovers of the coding wave are reproduced the speech will be intelligibly decoded. The decoding wave need not be the reciprocal of the coding wave. It can be like the one drawn in at the right in the photograph. It is only necessary, therefore, to generate a wave having its crossovers at the indicated points, and reverse the phase of the scrambled signal of these points.

H2, H3. Level modulations by themselves are not private, but they might very well be used in combination with other systems in an attempt to foil the matching of speech patterns. The level modulations themselves, however, need not be solved cryptographically.

J1. There appears to be no method either cryptographic or noncryptographic for breaking the noise masking method if the noise is predistorted, random and sufficiently high in level to really mask the speech. These requirements, however, make the technical difficulties for system operation very great and it is unlikely that this method can be used over radio channels. Cracking this system therefore becomes a matter of solving the noise distorting system. Project C43 has had no experience along these lines.

K1, K2, K3, K4. Scrambled vocoder channels can theoretically be solved by matching oscillograms. Actually as mentioned in Section 4 above this procedure is very difficult because the channels look so much alike.

L1, L2, L3. Channel mixing systems would be exceedingly difficult to handle cryptographically if a sufficient number of channels were involved so that non-cryptographic methods were inapplicable. The only possible method of attack appears to be matching spectrograms. Since, however, about 25 percent of normal speech consists of pauses, many of the switch points will occur in these pauses and it will therefore be difficult to establish continuity by matching.

7. Determination of the Message

The objective of decoding work is usually not to determine the codes used, but to learn the intelligence which was transmitted under these codes. In the case of repeated code systems, the procedure for obtaining intelligence is obvious once the code has been determined by the methods outlined above. It is only necessary to set this code into a machine similar to that used at the receiving end of the system being monitored, and listen directly to the transmitted speech. If the material has been recorded while the code was being determined, the recorded material can in general be decoded in the same way.

In the case of nonrepeated code systems, the determination of the code sequence leaves us in general a long way from the determination of the message. Obviously, all the material must be recorded in scrambled form. It is necessary during the process to establish time reference points in the scramble, perhaps by superposing clicks or spurts of tone during the recording process, and referring the code sequences to these points. A decoding machine must be available, such as the one described in Preliminary Report No. 15, which is adaptable to a variety of coding systems. The code sequence must be set into this machine perhaps in the form of a punched tape. The scrambled material must then be reproduced and fed into the machine, maintaining proper synchronism between the reproducing and decoding systems. Obviously, this is a very formidable job.

There are some alternative possibilities which may apply in special cases. In the case of nonrepeated code TDS, for instance, the process of matching variable area patterns has actually restored the speech in reproducible form. Variable area patterns can be played back just like the sound tracks used with motion pictures.

A playback machine of this type is described in Preliminary Report No. 12. The rearranged elements are mounted on a strip of adhesive, and scanned with a light slit and photocell. Considerable noise is caused by the joints between the separate elements, but this could be largely eliminated by a specially designed squelch circuit, perhaps controlled by a separate light beam and photocell to cut off the output wherever a joint is passing under the scanning beam. The first attempt to use this decoding method was unsuccessful, as discussed in Preliminary Report No. 26. However, there is nothing basically wrong with the method; it simply needs better execution than it received in the first attempt.

If the solution of the coding system requires spectrograms rather than variable area patterns, it is still theoretically possible to play back the rearranged pieces. A playback machine for spectrograms is described in Preliminary Report No. 17. This first model requires a negative transparency of the spectrograms, to be scanned by a light slit and photocell, with a multi-frequency light chopper interposed ahead of the photocell. Again the method is basically sound. The experimental machine described in the report needs considerable improvement before it will yield adequate quality for the purpose described above, in order to overcome the degradation of quality caused by the joints, by slight misplacements of the elements, by "spill-over" at the boundaries, etc. Furthermore, in order to get good patterns for matching, the signal must be subjected to a very high degree of compression, which distorts both the time and the frequency distribution of energy. It may be necessary to make one kind of pattern for matching, and another kind for playing back, as was done with the variable area patterns described in Preliminary Report No. 26.

As a final alternative, it is possible to learn to read speech spectrograms by visual inspection. Theoretically, therefore, the rearranged spectrograms might yield the message directly. Here again, however, the boundary distortion will increase the difficulty of reading the patterns. It has also been found that the best patterns for matching are not the best for reading, and it may be necessary to make two sets of patterns. However, since spectrograms have been continually improving, the possibility of visually determining the intelligence from rearranged spectrograms must be listed as a distinct possibility, and one which, if it is feasible, is the most general of all methods since the basic procedure is the same for all of the scrambling methods which can be handled in this manner.

CHAPTER VII

PRACTICAL EVALUATION OF PRIVACY SYSTEMS

The material in the foregoing chapters is intended to be useful not only for possible interception and decoding of scrambled messages, but also to aid in the production of new privacy systems, and to estimate the degree of security which we might expect to obtain from these or other systems. Experience has shown that there is a strong tendency to underestimate the security or military value of a given privacy system as soon as laboratory studies have indicated that the system can be cracked. In this chapter, therefore, an attempt will be made to point out the great difference between what might be termed theoretical or laboratory evaluation and practical or field evaluation. In order to balance the effect of the previous chapters this one is written from the standpoint, not of the man interested in decoding a system, but of the man interested in getting a practical privacy system into use in the field.

1. Cracking Time

The objective of a laboratory study of a privacy system is to obtain some kind of quantitative measure of the time or effort required to decode the system. The questions are , "How long does it take to determine the code, and how much equipment and how many people are required?" In general the procedure is to acquire a pair of actual models of the system under scrutiny. The coding and decoding processes are studied in detail possibly with the aid of mathematical analysis to determine whether there are any weaknesses or any characteristics of the coding process of which advantage might be taken to assist in the cracking process. Possibly a noncryptographic method will be found to apply. In this case the cracking time reduces substantially to zero. If noncryptographic methods are not applicable, available cryptographic tools and methods are brought to bear. Usually a new scrambling system will require modifications or changes in the existing tools or techniques. Possibly the basic methods can be improved for use against this particular system, or possibly new methods can be devised. Presumably after all this development work the project personnel will have become skilled in the art of decoding this particular system. The cracking time can then be determined quantitatively, perhaps with estimates as to how far this may be reduced by further skill.

In the case of repeated code systems, the cracking time determined in the above way substantially represents the total decoding time, because, as mentioned previously, this code can be set into a receiving machine and the message obtained directly. Some additional time might be added, however, for determining what was said during the time that the code was being determined.

The procedure outlined above is very well illustrated in the series of Preliminary Reports covering the development of cracking methods for the repeated code TDS system. They include mathematical analysis (Nos. 3 and 6), the development of a new decoding tool (Nos. 1 and 7), and the reduction of the decoding technique to a routine (No. 14). In the case of the multiplication system, the chronological steps are all listed in one report (No. 18).

Too often the cracking time, as determined above, is quoted without qualification to describe the security of a system. It is, of course, usually understood that the use of this figure involves the following assumptions: (1) that the enemy knows all about the coding system (2) that he is equipped with an adequate supply of the machines (our own models may still be far from the production stage) (3) that he has developed the same decoding tools and techniques that we have (some of our tools may be entirely new and secret) (4) he is equipped with an adequate supply of the decoding tools, (5) he has trained men in their use, and (6) he is in a position to receive a good signal free of interference. Such assumptions certainly represent an extreme possibility. Experience has shown that there is a strong tendency to forget just how extreme a condition such assumptions represent. Even if the assumptions are valid there are still other factors which affect the military value of a privacy system as will be discussed in a subsequent section.

2. Nonrepeated Code Systems

If the code is changed periodically it may be necessary to have several decoding teams working in parallel in order to keep up with the transmitted material, as was mentioned in the previous chapter. The number of teams which will be required depends on the relation

between the intervals of the code changes and the cracking time. No particular difficulty presents itself in expressing the decoding effort under these conditions in terms of man hours. The evaluation is complicated, however, by the necessity for additional equipment, not only for decoding but for recording.

In the case of nonrepeated code systems, the cracking time for any given portion of a message will in general be long compared to the duration of that portion of the message. Every portion of the message must be cracked individually, and the decoding effort can be expressed as a ratio of decoding time to message time. This ratio may be 1,000 or 100,000 to 1, that is, each second of message will take 1,000 or 100,000 seconds to decode. Conversely it would require 1,000 or 100,000 teams to keep up with the messages as they are spoken.

This kind of evaluation is somewhat unsatisfactory, because the length of time it will take the enemy to determine the intelligence in a particular sentence which might carry military information will depend on whether or not he is at the moment working on this sentence or whether he is wasting his time decoding previous material which might contain no information of value to him. It has in fact been proposed that the security of such high-privacy systems could be materially enhanced by keeping the circuit 100 per cent busy with all kinds of material, possibly even from recordings, making certain that the enemy has no way of determining when the circuit is being used for passing important information. As in the case of nonrepeated code systems, it seems a bit unrealistic in evaluating such a system to assume that the enemy will seize upon the few seconds of message time which are important, and to compute the length of time it will take him to decode that portion of the message.

3. Code Analysis

Many schemes have been proposed for generating everchanging codes by a combination of short cycles geared together in such a way that the number of elements in the cycle is the product of the number of elements in the individual cycles. One scheme is to use odd ratios, such as 99 to 100, so that the code cycle will not repeat until the smaller wheel has made 100 revolutions. In other words there are 9,900 steps in the code cycle before it repeats. Another scheme is the cyclometer type in which one wheel rotates one step for each revolution of another wheel. Again the total cycle is the product of the number of steps on the individual wheels.

Such schemes should be distinguished from truly nonrepeating codes, because wherever cyclic processes are used, they are subject to analysis. This is a matter pertinent to the field of cryptanalysis and will not be discussed here. In general it may be said that the difficulty of solving such long cycles is not determined by the total length but rather by the length of the individual subcycles.

Systems designed to produce a long code sequence usually contain provision for readjusting or realigning certain elements periodically or from day to day. Assuming that we know all about the system except the momentary settings, estimates can usually be made of the length of time and the number of people it would take to determine the unknown settings by analyzing a given sample of the code sequence. The analyst requires a knowledge of the code for a long sequence of scrambled speech before he can begin the work aimed at determining the unknown settings. He must obtain and solve a sufficiently long sample of the scramble by the methods outlined in the previous chapters and then analyze this sequence to obtain the settings. Too frequently the evaluation of a coding system is based on the analyzing time alone whereas the time required for solving or unscrambling a long sequence of scrambled speech may be overwhelmingly greater than the analyzing time. In fact if there is no way of solving the code sequence from the scramble alone, then the analyst can contribute nothing, and the system is still secret regardless of any inherent weakness of the cyclic coding system.

4. Field Evaluation

The continuously changing military situations of modern warfare require rapid means of communication in order that the required military actions can be taken. Obviously a perfectly secure speech privacy system is of no military value if it requires so much time for encoding and decoding that it slows up the communication system to the point where appropriate steps cannot be taken when needed. Obviously also a cracking system is of no value if it is too slow to permit counter measures to be taken according to the intercepted intelligence. For certain purposes 15 minutes or even 5 minutes cracking time is much too slow. Where this is true, a privacy system giving 15 minutes or 5 minutes privacy is just as good as one with an hour's security. This is important because systems

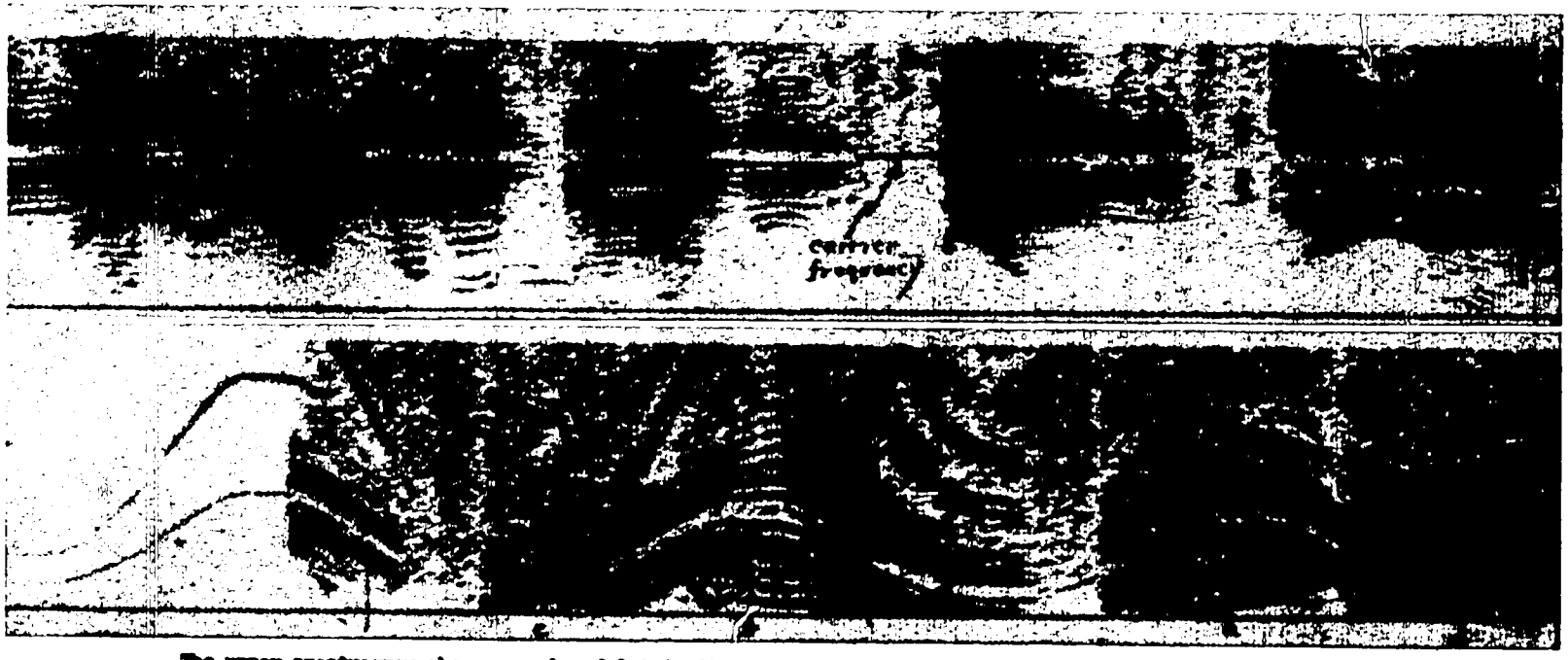
which afford a few minutes of privacy can be produced in portable form, whereas those providing longer privacy, at the present writing cannot.

Consider also the equipment and trained personnel required for decoding intercepted communications. As a specific example, the small TDS unit required about 15 minutes for decoding but it required a van-load of highly specialized equipment as described in Preliminary Report No. 24. Suppose the small portable TDS unit were used in many planes and tanks and other mobile equipment that required some privacy. Suppose also that different codes were used within different groups of units and that the codes were changed at some reasonable interval. Would it be worth the enemy's while to provide enough decoding equipment and enough trained personnel to follow these units around and decode their messages? If it is not worth his while, then units rated as low in privacy may provide high grade privacy under such conditions.

Obviously the foregoing does not apply if the units are used to convey messages between the higher echelons of command. In such cases the messages have a longer term significance to the enemy, and he can afford to devote considerable time and equipment to intercepting and decoding them.

Advantage might also be taken of the element of surprise. Suppose we suddenly introduce in the field a low-grade privacy system in large quantities. How long would it take before the enemy diagnosed the system, developed a decoding method, manufactured receiving sets of the proper type and also decoding equipment, distributed these where needed and organized and trained personnel to use them? Until he has done these things the units provide complete secrecy. A different kind of system might then be introduced which would again provide secrecy for a time.

It is intended in the foregoing simply to point out that there are other considerations in the evaluation of privacy systems than the time it takes a highly specialized group, such as the personnel of C-43, to decode the system under the ideal conditions of a laboratory. The decoding time alone is so often quoted, because it is the only element which can be described quantitatively. While there is always theoretical agreement about the existence of the other considerations apparently they cannot be pointed out too often or too strongly.



The upper spectrogram shows speech modulated with a carrier of 2000 cycles. Note the symmetry of the pattern around this frequency. Each harmonic, and each resonance area, is duplicated on both sides of the carrier. This clearly shows that the two sidebands are exactly alike, except that one is inverted. The recording from which these spectrograms were taken has somewhat attenuated the frequencies near the base line.

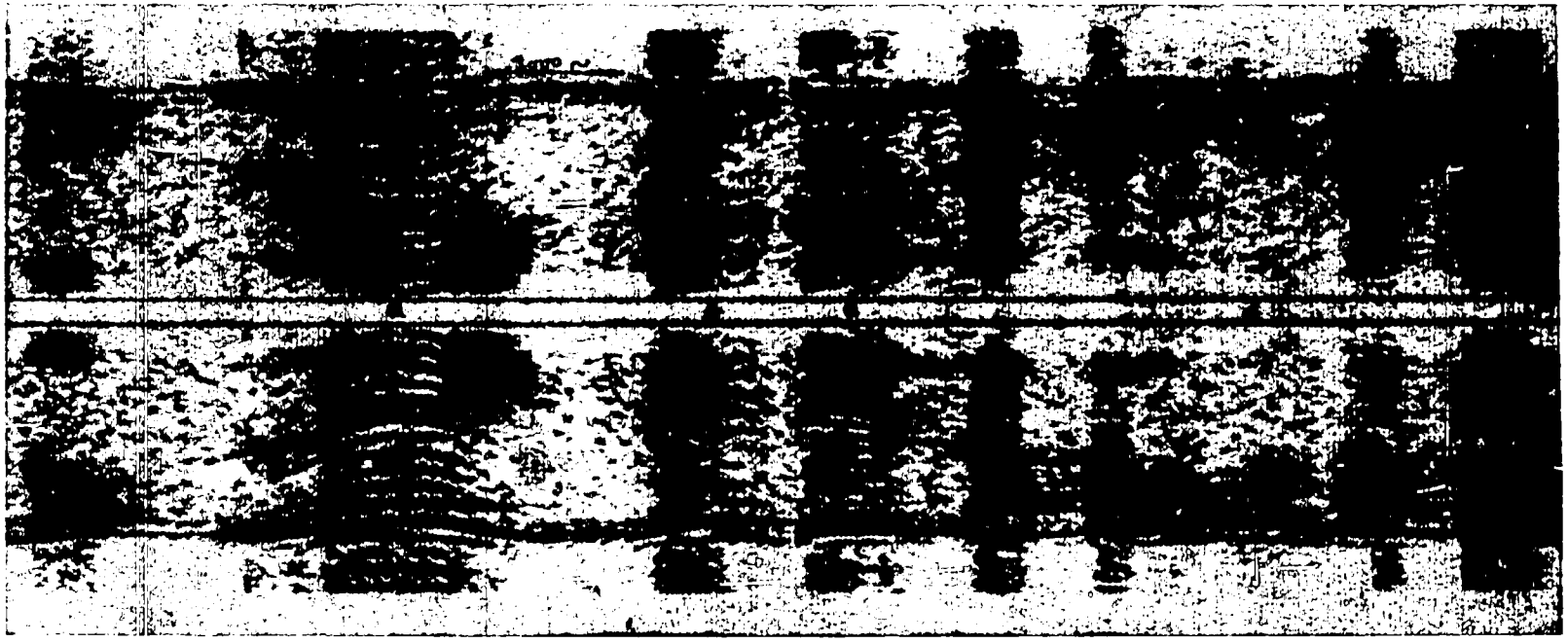
The carrier itself is largely suppressed by the double balanced modulator; high level sounds, however, occasionally unbalance it sufficiently for the carrier to show through.

In the lower spectrogram, the carrier frequency has been wobbled at a rather slow rate. Note that as the sidebands move up and down with the carrier, the harmonics remain parallel, as at point a, except when marked voice inflections occur, as at point b.

In the clear space at the left, the wobbling carrier can be seen, together with its second harmonic. In the upper part of the spectrogram at point a, the second harmonic of the carrier can also be seen, with its own set of sidebands.

When the carrier frequency is low, as at points c and d, the lower sideband can be seen folding back; the folded back portion is right side up, and overlaps the regular sidebands.

Figure 47 - Illustrating Modulation Sidebands



The upper spectrogram was made in the normal manner, showing inverted speech; the lower one was made with reversed oscillator sweep, producing a mechanically inverted spectrogram in which the speech appears right side up.

This sample contains harmonics with marked curvatures. These are voice inflections, and their occurrence can easily be recognized by ear. In general, samples with such voice inflections should be captured because they are most useful for diagnosing scrambling systems.

In the upper spectrogram, at points a and d, note how the curvature of the harmonics is least at the highest frequencies and progressively greater toward the lowest frequencies. Similarly, at points b and c, the slopes of the harmonics are least at the top and greatest at the bottom of the spectrogram. This is directly the reverse of normal speech, and definitely indicates inversion. The lower spectrogram illustrates the normal slopes and curvatures.

There is obviously a low pass filter in the system, at about 3000 cycles, as indicated by the rather abrupt change in intensity. Such a filter is normally used to cut off the upper sideband. It is usually also designed to cut off the carrier. In this case, its cutoff frequency is lower than the carrier frequency. This shows up at points a and b in the harmonics fading out before they completely flatten out. However, the inversion frequency is not far from 3000 cycles, because the slopes are substantially zero as they approach this frequency.

Figure 48 - Simple Inversion

Here the speech is divided up into sections by sharp vertical boundaries. Examining the individual elements, it will be noted that a, b, and c show unmistakable signs of inversion. Elements d, e, f are definitely not inverted.

The regularity of the dark areas in the lower spectrogram suggests that only two conditions are involved. Note also that in general, where slopes can be clearly discerned, the harmonics slope in opposite directions in adjacent elements. These indications point to alternate straight and inverted transmission.

This diagnosis could be confirmed by making a mechanically inverted spectrogram, and matching together alternate pieces from the two spectrograms.

The switching intervals are irregular, with no repetition apparent within the time covered by this spectrogram. Additional spectrograms, covering a longer period, might show a repeated cycle.

Figure 49 - Alternate Inversion

In the upper spectrogram, the speech band has been displaced from its normal location by 1000 cycles, in the lower one by 2000 cycles.

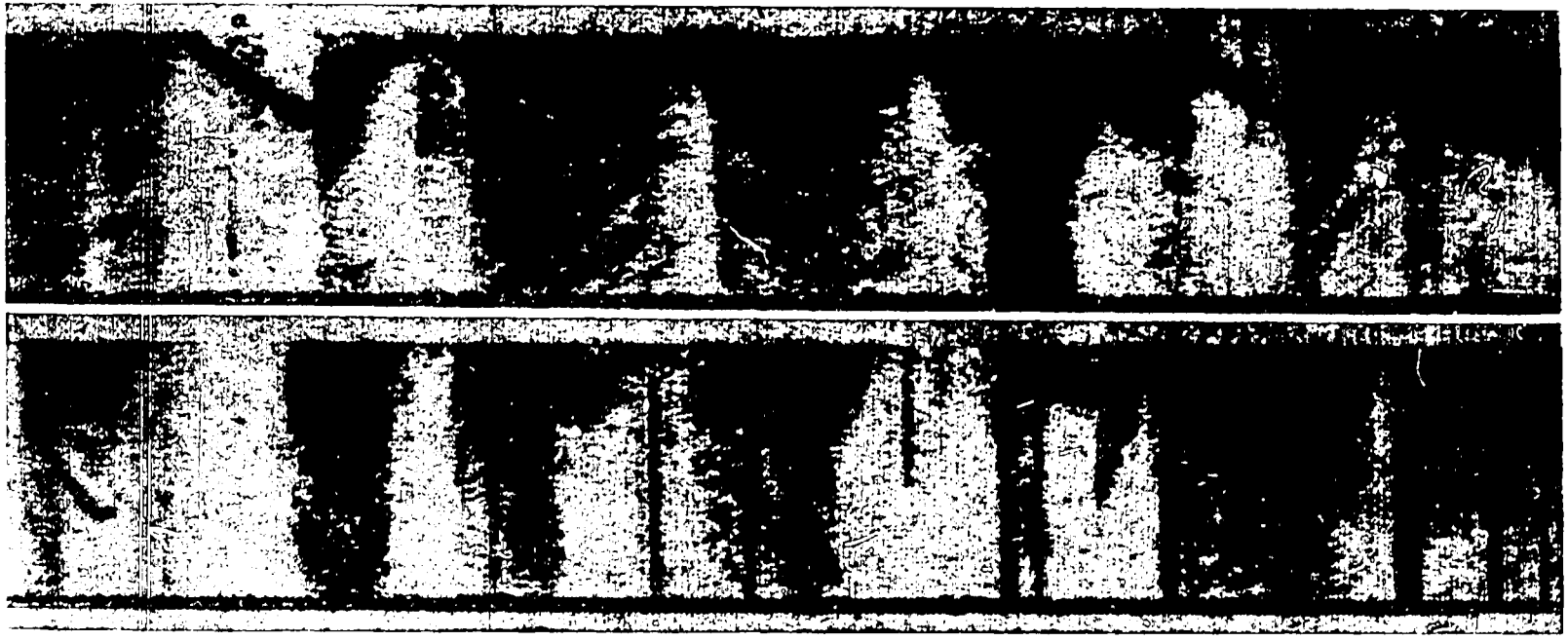
Recognition of this condition is aided by familiarity with the appearance of normal speech in spectrograms. All the vowels have characteristic resonant areas close to what would be the base line (zero frequency) in normal speech, and the glides such as occur at a and b tend to start from this region.

At c, the harmonics look as though they could meet at a point if extended to the right. This point would be the true base line.

A displacement of this type could be produced by modulating with a carrier of 1000 or 2000 cycles, and suppressing the lower (inverted) sideband with a high pass filter. In practice, however, a double modulation process is used, because the displacement may be changed at will without changing the filter cutoff.

In both spectrograms, a small amount of lower sideband can be seen. This incompletely suppressed sideband would look the same whether produced by single or double modulation.

Figure 50 - Fixed Displacement



The upper spectrogram shows an example of wobbled inversion. Note the occurrence of harmonics symmetrically disposed about a suppressed carrier frequency which is varying between 2500 and 3500 cycles. Note how the harmonics remain essentially parallel. At points a and b the sidebands appear to consist only of low frequency noise. The carrier wobble is irregular in shape but regularly repeated in time.

If there is a low pass filter in the system, its cutoff frequency is higher than 3500 cycles. Note that if it were lower, it would occasionally cut off some of the wanted (lower) sideband.

The lower spectrogram shows a wobble covering a much wider frequency range. The lower sideband dips below the 3500 range of the spectrogram only part of the time. This would certainly be diagnosed as a band displacement system involving double modulation. If it were encountered in practice, wide band spectrograms would be used to determine the exact displacement.

Both of these systems are laboratory simulations for illustrative purposes.

Figure 51 - Wobbled Displacement

The upper spectrogram was artificially produced by cutting up and rearranging a spectrogram of simple inversion.

There are vertical boundaries about 275 milliseconds apart. There are also horizontal boundaries, but these are not continuous.

Elements 3 and 6 show no horizontal boundaries, but they show the signs of simple inversion.

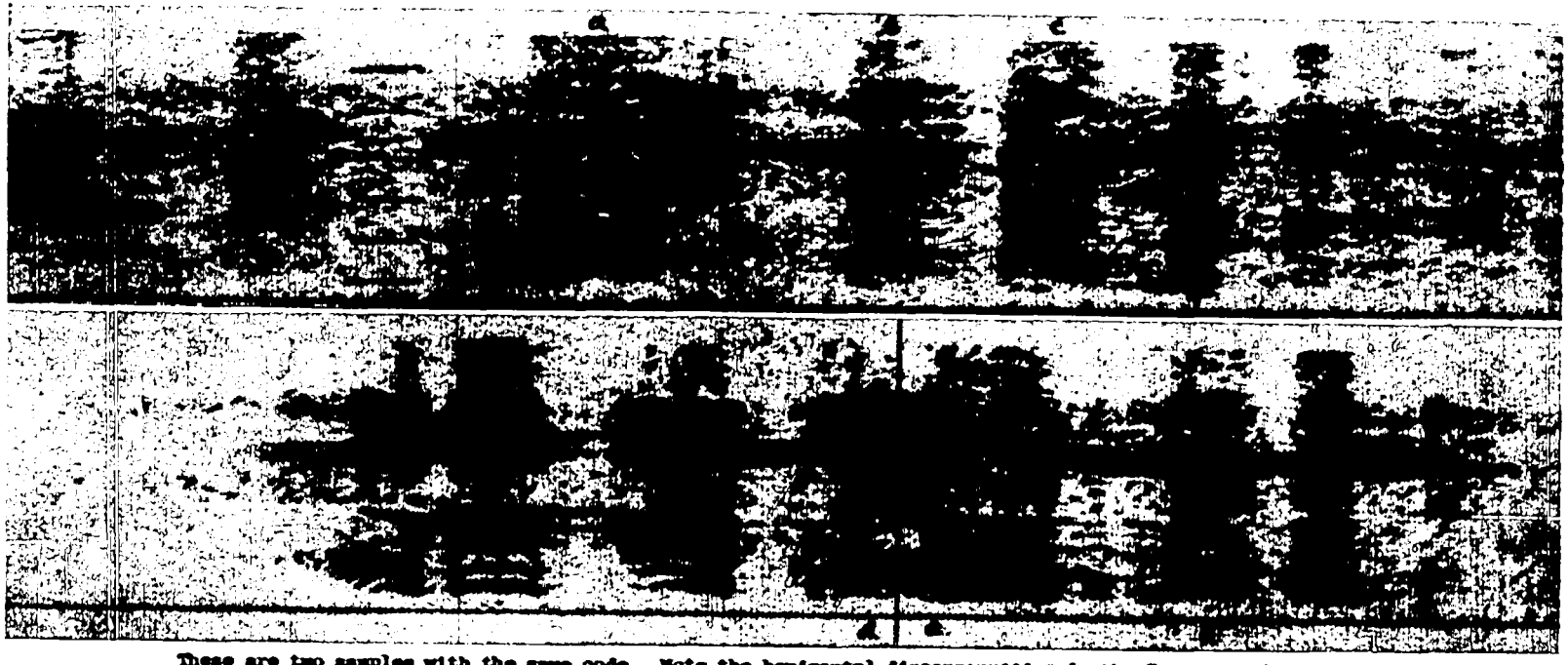
All the other elements show horizontal boundaries, with higher slopes above the boundary than below. If the two portions of each element were interchanged, the slopes would be in the correct order (for inverted speech).

It can also be seen that if the elements were thus rearranged, the harmonics of each element would match those of the preceding and following elements. This, of course, should be confirmed in practice by trial.

These configurations would be produced by reentrant inversion. In terms of these spectrograms, this process results in inverting successive elements about frequencies of 1000, 2000 and 3000 cycles respectively, removing the upper sideband, and replacing it with that portion of the lower sideband extending below about 200 cycles.

The lower spectrogram is a duplicate of the one used above. The boundaries have been marked off to show that the elements rearranged as suggested above, would form continuous inverted speech.

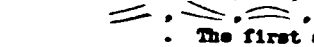
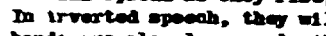
Figure 52 - Reentrant Inversion



These are two samples with the same code. Note the horizontal discontinuities in the frequency distribution of energy, which can best be seen by looking endwise along the patterns. These are the filter boundaries. This system shows five bands, covering the range from 250 to 3000 cycles. In split band systems, the subbands are generally equally wide, for practical reasons.

There are no vertical discontinuities, other than the normal sequence of words and spaces.

Note that the harmonics slope or curve in opposite directions within a word or syllable. This indicates that some of the subbands have been inverted.

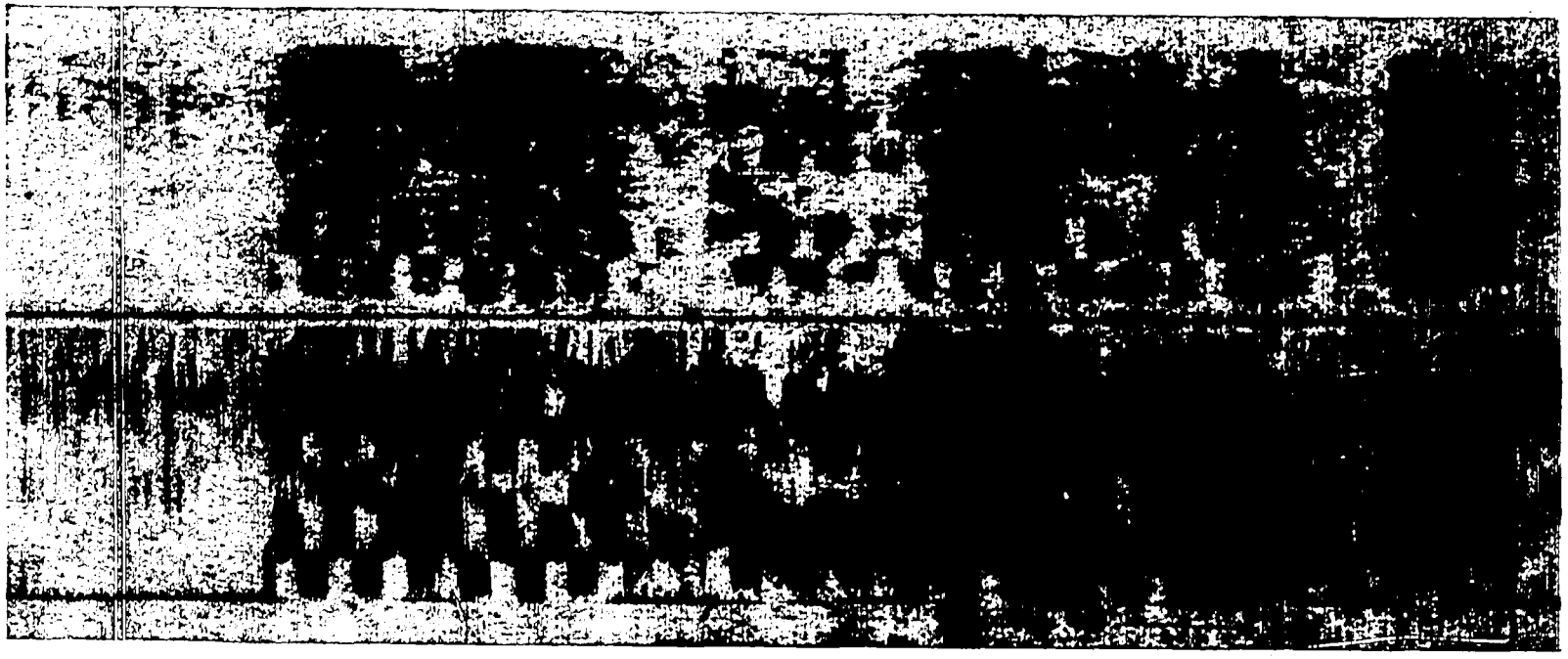
In normal speech, the harmonics will spread as they rise, showing the following configurations:  In inverted speech, they will show the opposite trends:  The first and fourth bands are clearly normal, the others are inverted. This shows up very clearly at point a.

Note that the fourth band shows the least slope or curvature; this must have been the lowest band originally. The middle band shows the most slope or curvature; this must have been the highest originally. The others can be similarly located, combining the indications from all the indicated points. Any one point is sometimes misleading due to the proximity of a harmonic to the filter boundary, as in the top band at a.

At points d, e there is a double inflection. This can be misleading, unless the slopes are estimated for simultaneous instants. The vertical line was drawn as a guide.

The code is 4, 2', 1, 5', 3', the primes denoting inversion.

Figure 53 - Fixed Split Band Scramble



These spectrograms show both horizontal and vertical boundaries. However, the level of the elements as a whole shows a rather smooth flow, as of words and spaces. Note the clear space ahead of the first word. These indications suggest that the elements have not been shuffled in time.

The presence of band shifts, however, is quite obvious. Harmonics can be seen sloping in both directions within the elements, particularly in the first word group.

The checkerboard effect in the lower spectorgram, however, suggests that only two codes were used alternately. This is corroborated by the fact that the middle band shows no vertical discontinuities, indicating that it was never switched.

Figure 54 - Rapidly Switched Split Band Scramble - Example 1

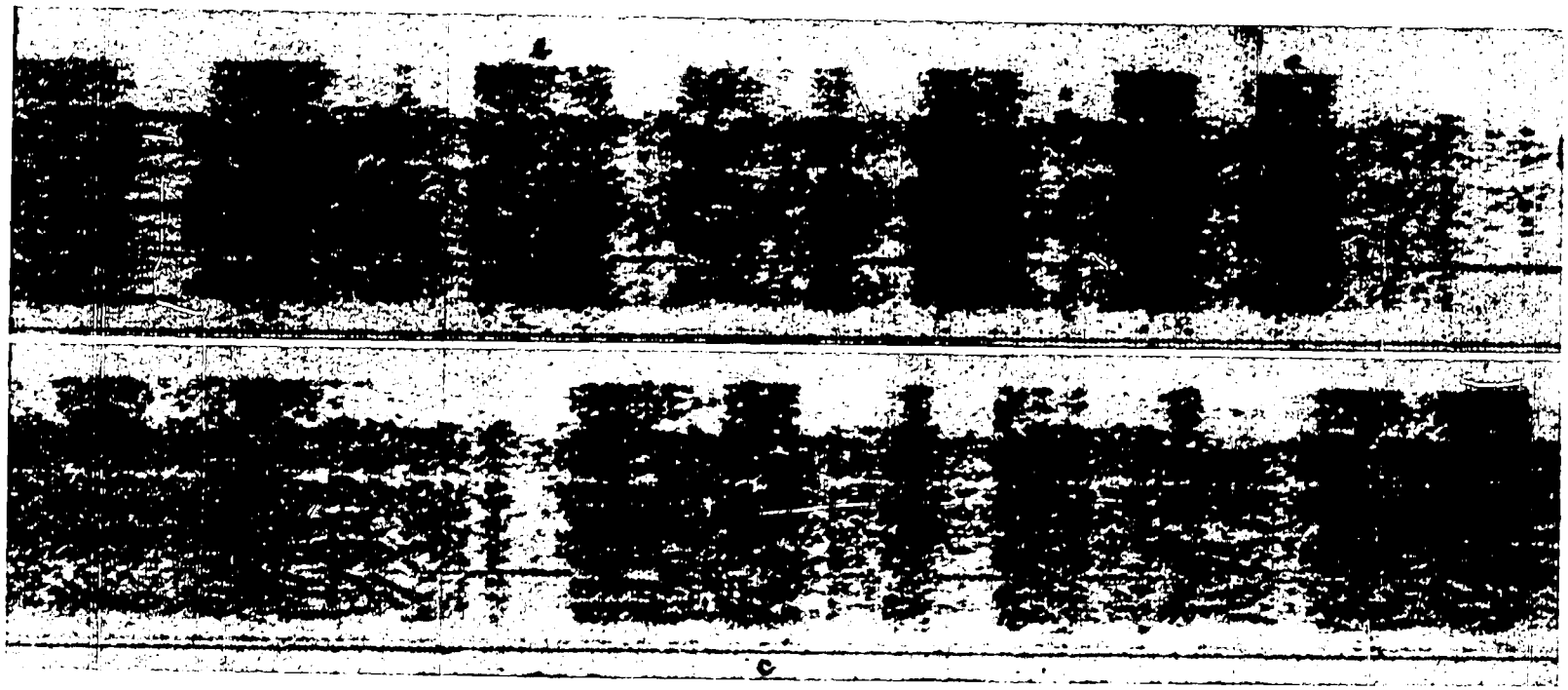
Here there are both horizontal and vertical boundaries. There is no evidence, however, that the elements have been time shifted. There are unbroken clear spaces as at a, the long word groups such as b show no abrupt changes in level or in pitch, and elements having different characteristic appearance, such as those at c and d, are not intermingled.

There is abundant evidence of band shifts: harmonics sloping in opposite directions, indications of inversion, and abrupt changes in the resonance areas.

Close scrutiny of long word groups such as b shows that several codes are being used, although it would take several samples to establish just how many.

This sample illustrates that if the codes are rapidly shifted, any one element seldom contains enough clear information to determine which code applies to it. However, if accumulated information about the system can be brought to bear, two clear bands may sometimes be sufficient to identify the code.

Figure 55 - Rapidly Switched Split Band Scramble - Example 2



These look at first glance like split band scrambles. The horizontal boundaries are quite evident. The bands are 600 cycles wide, beginning at 400 cycles.

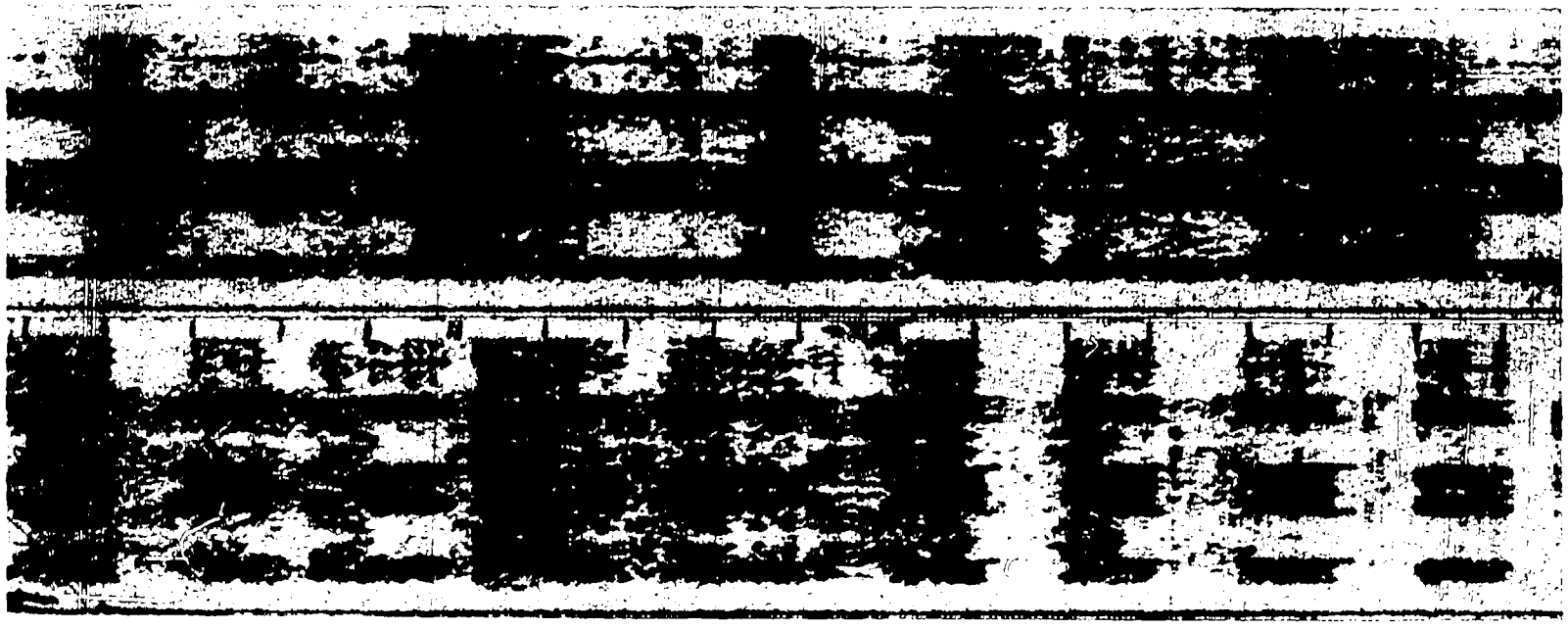
It will be noted, however, that all four bands look alike, except that alternate ones are inverted. Otherwise the slopes and curvatures are alike in all bands. An outstanding example is at point a. There is no gradation in slope.

The components are not uniformly spaced within a band, and they frequently go in both directions within a band, as at b and c.

These are the characteristics of TDM scrambling. In the particular system illustrated above, the frequency range was divided into four bands, and all were modulated down to the lowest frequency. The switching rate was 600 per second, but the entire band was then modulated up 400 cycles, to avoid having the lowest sideband extend to frequencies too low for transmission channels to handle.

It is characteristic of TDM to produce upper and lower sidebands around the switching frequency and around its odd multiples. The sidebands differ only in the phases of their components.

Figure 56 - Time Division Multiplex

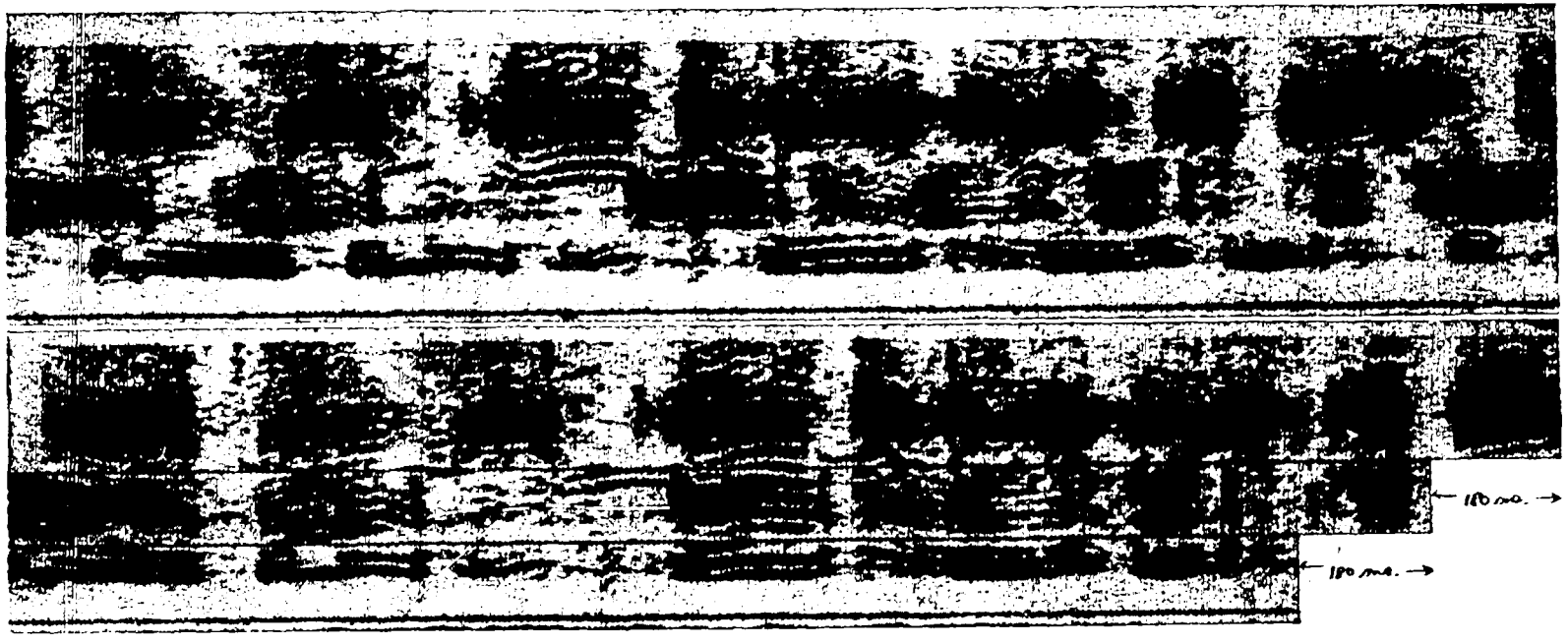


This is the same TDM system as the previous illustration, but a band of noise has been added to increase the privacy.

One half of the highest of the four frequency bands into which the speech channel was divided has been filled with thermal noise. In the upper spectrogram, this noise was steady, and in the lower one, it was turned on and off about 4 times per second.

Note that although the noise was introduced into only one subband, it appears in each of the four sidebands in the above patterns. This shows that in TDM, each sideband contains components from each subband.

Figure 57 - Time Division Multiplex with Noise Channel



Here there are horizontal boundaries, but the filters apparently do not cut off very sharply because the bands appear to overlap occasionally.

Note the staircase effect in the upper spectrogram, each syllable in the uppermost band appearing somewhat later in the middle band, and still later in the lowest band. This condition has been rectified in the lower spectrogram, by cutting the frequency bands apart and shifting them relative to each other, thereby restoring the normal appearance of words and spaces.

There has been no shifting or inversion of the subbands.

Note that the filter crossovers have been made very deep, as evidenced by the gaps between bands, and the lowest band has been severely curtailed in width, probably in an effort to reduce the amount of intelligence which may be gained by listening to any one band.

Figure 58 - Subbands Variously Delayed

The characteristics of TDS systems are vertical discontinuities, which must be regularly spaced, that is, the element lengths must all be equal, with the exception noted below.

The normal flow of words and spaces is destroyed by the time scramble. Adjacent elements show abrupt changes in total energy, in the location of resonance areas, and in pitch. All of these are readily discernible in the sample above.

This particular system uses a synchronizing pulse. Three of these can be seen in the spectrogram at about 2000 cycles, spaced 750 milliseconds apart. The length of individual elements is 1/20th of this, or 37.5 milliseconds.

Slightly ahead of each pulse can be seen two successive elements which appear to be twice as long as the rest. These are instances in which two speech elements which were adjacent in the original speech were left adjacent in the scramble. The fact that this occurs in each pulse cycle reveals (1) that the code cycle has the same length as the pulse cycle, and that the TDS code is repeated each cycle.

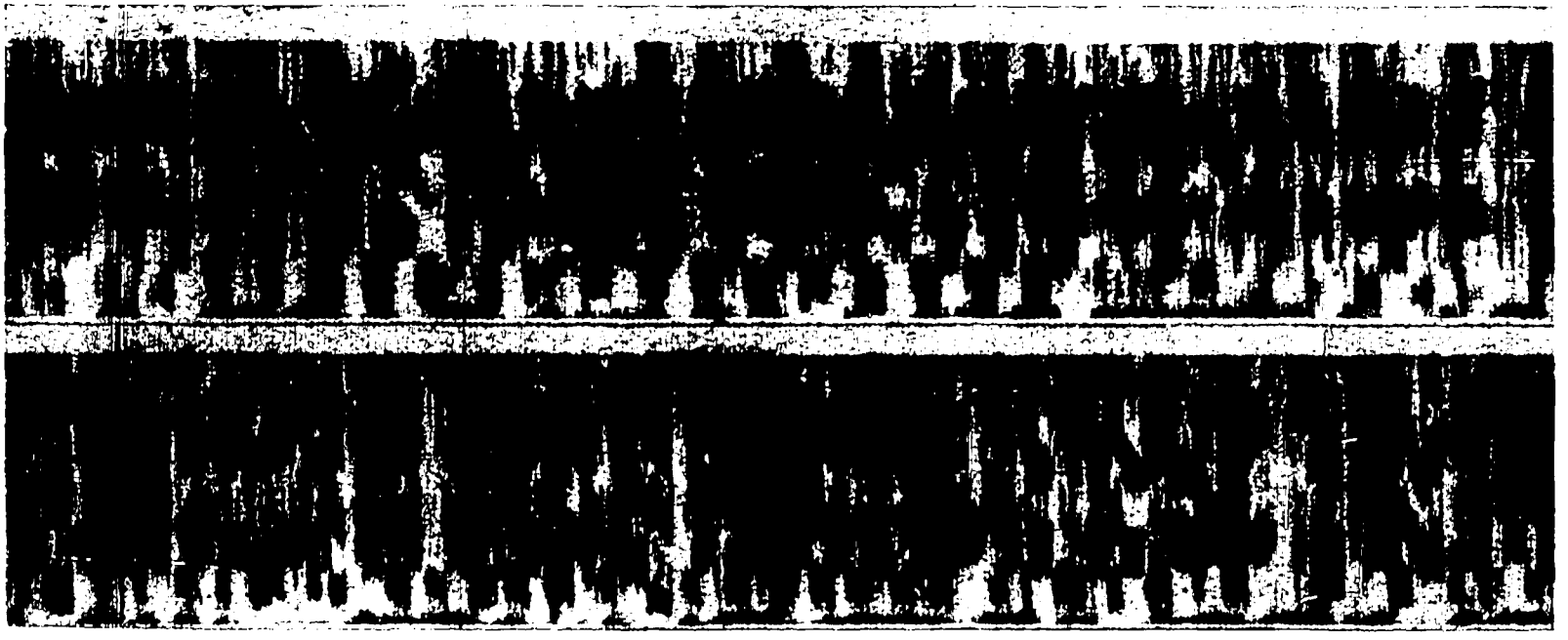
Figure 59 - Time Division Scrambling (TDS)

In this sample the evidences of band shifting are clear enough, as explained in previous examples. The fact that several different codes are used is suggested by the irregular distribution of the resonant areas over the frequency range. It is more conclusive, however, to examine the slopes of the harmonics in the upper spectrogram. At a, for instance, the lowest band shows more slope than either the second or fourth; at c, the opposite is true. At a, the harmonics in the second and third band slope in the same direction, at b in opposite directions. Just how many codes are used, of course, would require additional samples.

The evidences of time shifting are also clear. Elements a and b, for instance, are both strong, but are surrounded by gaps. At d, an element with energy distributed over the whole frequency range is surrounded by elements with entirely different distribution. There is also a marked difference in pitch between d and the surrounding elements. An even more marked change occurs at e.

This scramble, therefore, is the result of both band shifting and time shifting. It differs from a complete two-dimensional scramble in only one respect, which is described in a separate illustration.

Figure 60 - Combination of TDS and Rapidly Switched Split Band Scramble



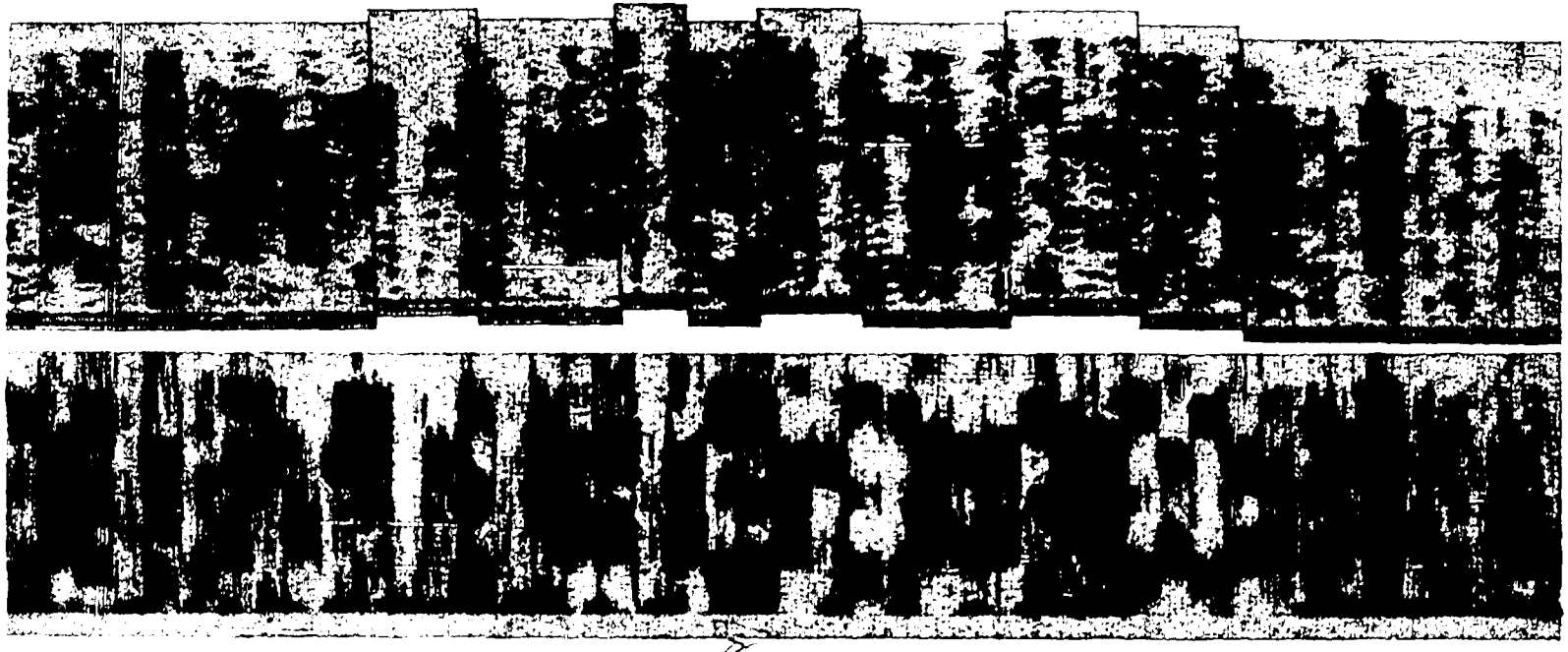
These are two combinations of TDS and rapidly switched split band scrambles. They differ from the previous illustration in that the two switching systems are independent. The split band code is changed at intervals of about 40 milliseconds, whereas the length of the TDS elements is about 34 milliseconds.

Each of these switching systems produces its own set of vertical boundaries. The distance between successive boundaries in the scramble, therefore, varies irregularly from a value corresponding to 34 milliseconds down to zero.

In the upper spectrogram, the speech was first subjected to TDS, and then to the split band scramble; in the lower spectrogram, the two scrambles were applied in the reverse order.

Only two split band codes were used, alternately. Since the second scramble tends to hide the first, the upper spectrogram shows the characteristic checkerboard effect noted in a previous illustration. In the lower spectrogram, the checkerboard effect is broken up by the TDS.

Figure 61 - Nonsynchronous Combinations of TDS and Split Band Scramble



An example of a two-dimensional scramble is not available, but the above spectrograms will serve to illustrate how such a scramble might be recognized.

The scramble above contains both time shifts (TDS) and band shifts (rapidly switched split band). Note, however, that in either of these scrambling systems, and in the combination of both, elements which are simultaneous in the scramble (that is, subbands within any vertical section) were also simultaneous in the original speech.

In the above example, therefore, a decided tendency may be seen for the character of vertical section to remain constant over the frequency range. That is, low level elements are low all over the frequency range, high level elements tend to be high all over.

Furthermore, subbands from voiced sounds do not occur in the same vertical sections with subbands from unvoiced sounds. Voiced sound may be recognized by the presence of harmonics in the 45° spectrograms, and regular vertical striations in the 300° spectrograms.

The most conclusive test for a two-dimensional scramble, however, is based on the fact that there will be differences in pitch within a vertical section. This can easily be tested as illustrated above. The spectrogram is cut down the middle of a vertical section, and the pieces shifted by one harmonic in either direction. If there is no change in pitch, the harmonics will still match all over, as above. If there is a change in pitch, the shift which is correct for one subband will be wrong for another. Two-dimensional scrambles, therefore, will not pass the above test.

Figure 62 - Test for Two-dimensional Scramble

The curvatures of the harmonics in the upper spectrogram look like voice inflections, except that they are abnormally frequent and rapid.

The resonance areas, as shown best in the lower spectrogram, also show a marked degree of curvature. Moreover, it will be noted that there is a marked correlation between the resonance areas and the pitch, that is, they reach their high and low points simultaneously.

In normal speech, the frequency and trend of the resonance areas are independent of the pitch trend. Wobbling the speed of a phonograph record or magnetic tape, however, multiplies the frequency of resonances as well as multiplying the pitch. These spectrograms were produced in this manner for illustrative purposes.

Figure 63 - Speed Wobble

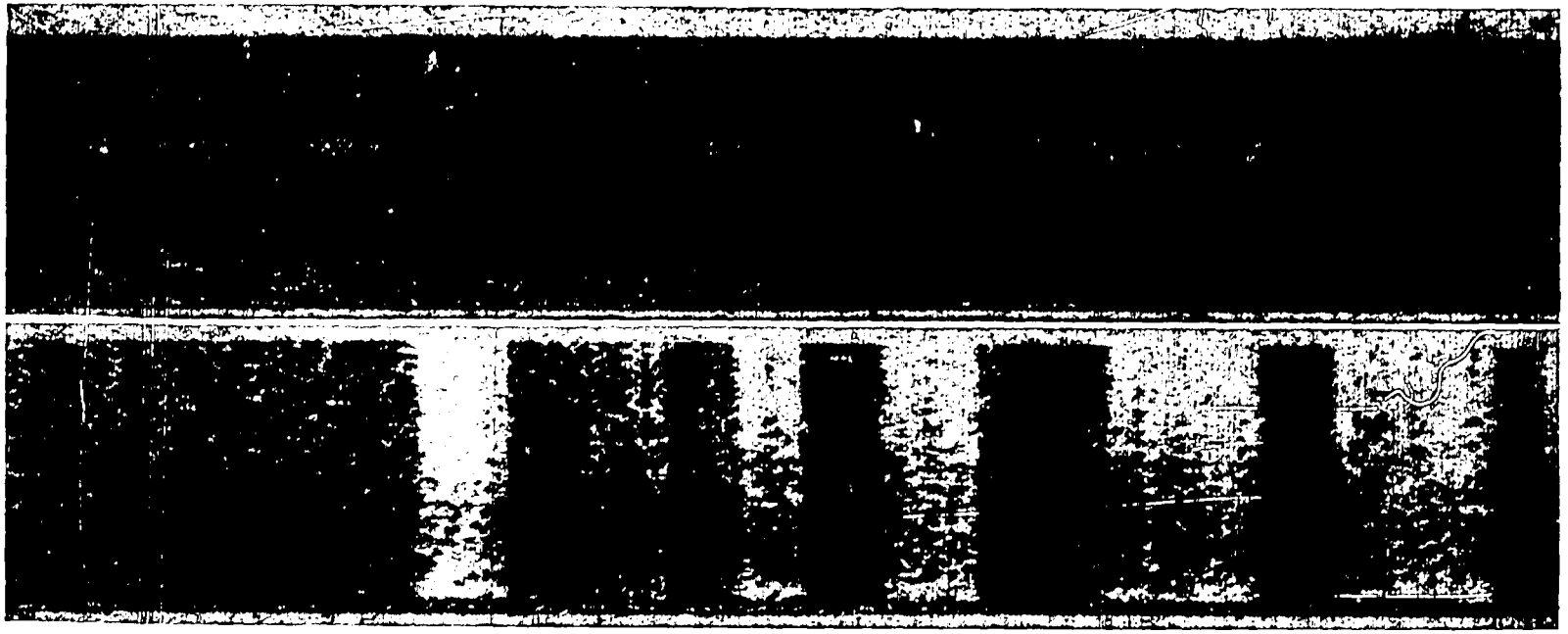
The above spectrograms were produced artificially by cutting up and rearranging spectrograms of normal speech.

The upper spectrogram shows the speech transmitted in sections about 160 milliseconds long, each section transmitted backwards. If the sections are as long as this, the condition can be recognised by familiarity with the normal speech formations, that is, by the way words normally start and end, and by the trend of the resonant areas. The slopes and curvatures of the harmonics, however, look perfectly normal.

If the elements are cut apart and matched, it will be found that the right hand edge of each element matches the left hand edge of the preceding element. This can be seen by inspection in the above example. The order of the pieces will be completely reversed after matching. If all the pieces are inverted, however, they will be found to match in their present order.

The lower spectrogram shows the same material, but in this case alternate elements are transmitted forwards and backwards. It will be found that none of the elements can be matched together at all. In order to match, alternate elements must be taken from a mechanically inverted spectrogram, as described in another illustration.

Figure 64 - Backwards



The upper spectrogram shows a continuous noise, with several words or syllables showing through. Counting the harmonics shows that the fundamental of the noise is about 100 cycles.

Examination of the signal with an oscilloscope shows that the noise consists of short pulses about 10 milliseconds apart. These can be removed by a blanking circuit.

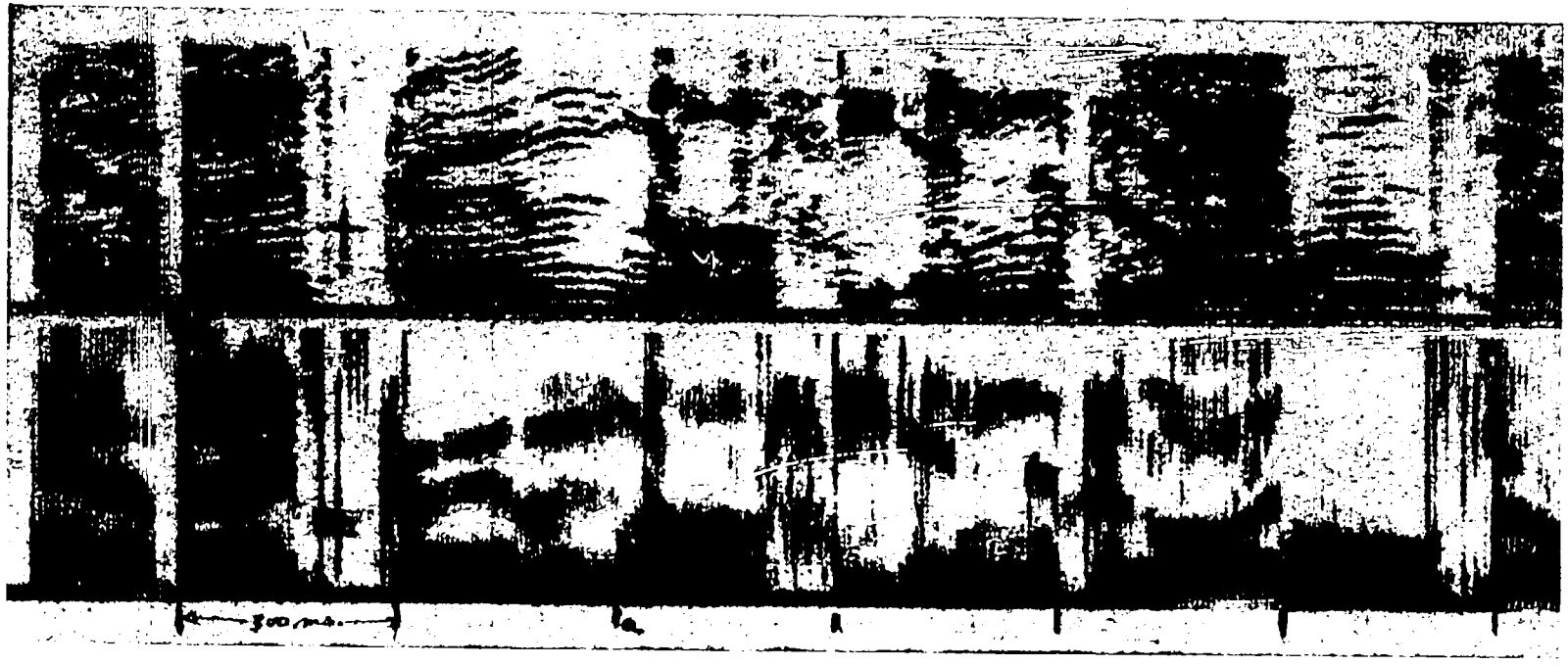
The lower spectrogram shows a sample (not the same as the one above) without the pulses. The outstanding characteristic, as in the sample above, is an almost complete lack of the harmonic structure of normal speech. Also, the energy is distributed more or less evenly over the frequency range for each word or syllable. There are no characteristic resonance areas.

There are no regular boundaries, either vertical or horizontal. The sequence of words and spaces looks normal in the spectrogram, and has the normal cadence of speech to the ear.

These characteristics are to be expected when the scrambling system operates on the wave form directly. In this particular system, the speech wave was multiplied by a coding wave. The latter was repeated 100 times per second, with a pulse between each cycle. It is obvious that a high degree of synchronism is required to remove the coding wave at the receiving end, which accounts for the high frequency of the synchronizing pulses.

It may be noted that phase reversal (at a sufficiently high and irregular rate to achieve privacy) is also essentially a multiplication process, except that the coding wave has no values other than plus and minus unity. Spectrograms of such a system would be expected to look like the above.

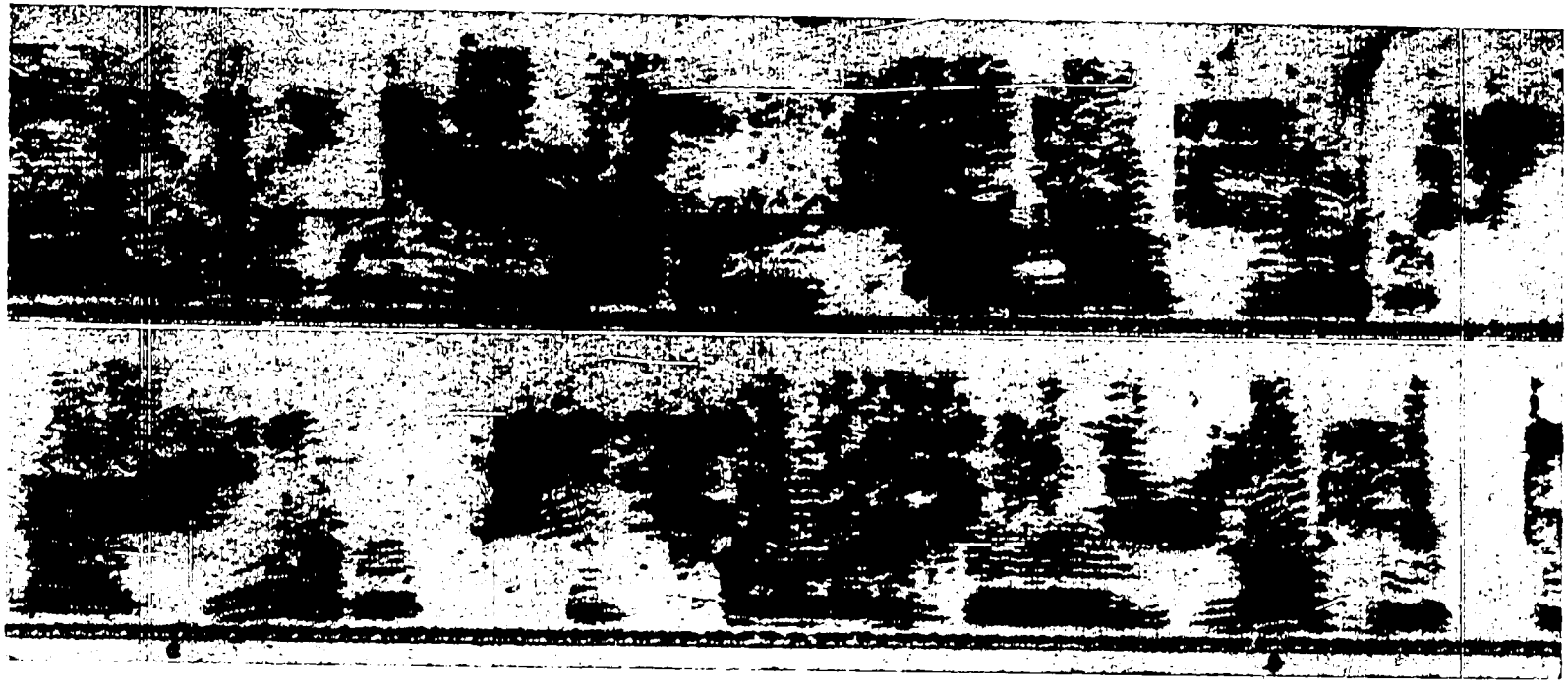
Figure 65 - Multiplication



Here two talking circuits have been switched between two transmission channels on a time division basis, switching at 500 millisecond intervals. The vertical discontinuities can usually be seen, but point a is an outstanding example of apparent continuity in pitch, inflection and resonance.

The ear can usually recognise the fact that two voices are present, at least at this switching rate. If the voices are nearly alike, or if recorded samples of the same voice are used, the nature of the scramble can be determined by cutting the pieces apart and attempting to rearrange them into continuous speech. This, of course, will be found impossible if it is a case of channel mixing. Another transmission channel should be found with the complementary elements.

Figure 66 - Time Division Channel Mixing



Here the harmonics occasionally curve in different directions, as at point a. A horizontal discontinuity is quite apparent at the indicated frequency, above which there are changes in pitch. These are not always readily apparent to the eye, but can be established by measurements.

In general, the syllables seem to begin and end at different times in the two bands. Formations such as c and d do not occur in normal speech.

No vertical discontinuities are apparent in either band, which indicates that if any time delays are involved, they apply to the whole bands. Yet (by trial) the speech in the upper and lower bands can not be matched by shifting the bands relative to each other.

It is apparent that in this case two talking circuits are being switched between two transmission channels. Another channel should be found which will contain the complementary subbands.

This was produced by a laboratory setup, for illustrative purposes. In practice, to obtain sufficient privacy, it would probably be necessary to combine this subband channel mixing with time division channel mixing illustrated elsewhere.

Point b, as a matter of interest, marks an outstanding example of apparent continuity in both pitch and slope.

Figure 67 - Subband Channel Mixing

Here the harmonics occasionally curve in different directions, as at point a. A horizontal discontinuity is quite apparent at the indicated frequency, above which there are changes in pitch. These are not always readily apparent to the eye, but can be established by measurements.

In general, the syllables seem to begin and end at different times in the two bands. Formations such as c and d do not occur in normal speech.

No vertical discontinuities are apparent in either band, which indicates that if any time delays are involved, they apply to the whole bands. Yet (by trial) the speech in the upper and lower bands can not be matched by shifting the bands relative to each other.

It is apparent that in this case two talking circuits are being switched between two transmission channels. Another channel should be found which will contain the complementary subbands.

This was produced by a laboratory setup, for illustrative purposes. In practice, to obtain sufficient privacy, it would probably be necessary to combine this subband channel mixing with time division channel mixing illustrated elsewhere.

Point b, as a matter of interest, marks an outstanding example of apparent continuity in both pitch and slope.

Figure 67 - Subband Channel Mixing

TABLE I

SPEECH SCRAMBLING DEVICES

	Illustrative Diagram Figure No.	Illustrative Spectrogram Figure No.	Type of System Studied by C-43	Noncryptographic Attacks Table II	Cryptographic Attacks Page
A. Single Modulation					
1. Inversion	7	48	Working	1a	
2. Variable Frequency Inversion	7	51	Simulated	1b, 2a, 2f	
3. Alternate Inversion		49	Simulated	1b, 2f	
4. Phase Reversal	7		Working	1b, 2e, 2g, 3e	52
5. Split Phase	8		Working	1a	
B. Double Modulation					
1. Fixed Displacement	9	50	Working	1a	
2. Stepped Displacement	9		Working	1b, 2a, 2f, 3b, 3f	
3. Wobbled Displacement, Regular	9	51	Working	1b, 2a	
4. Wobbled Displacement, Irregular	9		Simulated	3d	52
C. Triple Modulation					
1. Reentrant Inversion, Steps	10	52	Hypothetical	1b, 2e, 2f	
2. Reentrant Inversion, Continuous	10		Hypothetical	1b, 2a	52
D. Band Splitting					
1. Slowly Switched	11	53	Working	1b, 2b, 2f, 2g, 2h	52
2. Rapidly Switched	11	54 and 55	Working	2e, 2f, 2h,	52
E. Time Division Multiplex					
1. 4-band System	12	56	Working	1b, 2b	
2. With Noise Channel		57	Working	1b, 2b	
F. Magnetic Tape					
1. Delayed Subbands	13	58	Working	2b	
2. TDS, Repeated Code	14	59	Working	2e, 2g, 2h	52
3. TDS, Nonrepeated Code			Working		52
4. Speed Variations	15	63	Simulated	1b, 2a, 3e	52
5. Backwards Transmission	16	64	Hypothetical	1b	
6. Alternate Backwards and Forwards		64	Hypothetical	1b, 2e	

TABLE II

NONCRYPTOGRAPHIC DECODING METHODS

	Diagram Fig.	Used by C-43	Discussion Page
1. Captured Set or Functional Equivalent			
a. Fixed Condition - find by trial		Yes	33
b. Simple Program - get into step	27	Yes	33
2. Compromise Decoding Methods			
a. Intermediate Condition		Yes	34
b. Listen to Portion of Frequency Band	28 and 29	Yes	34
c. Listen Part Time			34
d. Limiter, Peak Chopper, Compressor	30 A,B,C		35
e. Rectifier	31 A,B, and 32	Yes	35
f. Superposition	33	Yes	36
g. Approximate Code by Trial		Yes	36
h. "Spoil" Good Code by Recoding		Yes	36
i. Directional Discriminator	34	No	37
3. Automatic Decoding			
a. Total Energy	35	Yes	37
b. Energy Frequency Distribution	36 and 37	No	37
c. Pitch Change Corrector		No	38
d. Wobble Corrector		No	38
e. Code Wave Generator		No	39
f. Parallel Automatic		No	39
g. Inharmonic Detector		No	39

~~SECRET~~ 054 450

ATI- 29345

TITLE: Final Report - Part I - Speech Privacy Systems - Interception, Diagnosis, Decoding, Evaluation

AUTHOR(S): Koenig, W.

ORIGINATING AGENCY: Bell Telephone Labs., Inc., New York, N. Y.

PUBLISHED BY: Office of Scientific Research and Development, NDRC, Div. 13

REVISION: (None)

ORIG. AGENCY NO.: (None)

PUBLISHING AGENCY NO.: 4573A

DATE	DOC. CLASS.	COUNTRY	LANGUAGE	PAGES	ILLUSTRATIONS
Oct '44	Secr.	U.S.	Eng.	111	photos, tables, diagrs

ABSTRACT:

The results of three years' experience in diagnosing, decoding, and evaluating speech privacy systems are summarized. Speech privacy systems may be used in connection with radio telephone systems or wire systems, but radio interception problems only are discussed. The decoding techniques described apply to wire as well as to radio communications. The sound spectrograph is described including its history, method of operation, and capabilities. It analyzes speech in terms of its three basic dimensions, frequency, amplitude, and time; and portrays the analysis in the form of spectrograms. Basic speech scrambling methods are also explained in which the original speech is transmitted with its parts modified, displaced, or interchanged. Cryptanalysis and cryptography, which apply to telegraph types of communication, are also described.

DIVISION: Electronics (8)

SECTION: Communications (11)

SUBJECT HEADINGS: Communication systems, Secret (23992.87); Decoders (28877)