365 — YES

[redacted]

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

INTERNAL NOTE NO. MSC-EB-E-68-4031-U

DIGITAL TECHNIQUES FOR SPEECH-TO-NOISE RATIO MEASUREMENT

[redacted]

[redacted]

MANNED SPACECRAFT CENTER

HOUSTON, TEXAS

November 17, 1968

[redacted]

INTERNAL NOTE NO. MSC-EB-E-68-4031-U

DIGITAL TECHNIQUES FOR SPEECH-TO-NOISE RATIO MEASUREMENT

PREPARED BY

D. R. Cook

D. Cook
Project Engineer

APPROVED BY

Sam T. House

S. House
Head, Data Systems Analysis Section

J. Overton
Chief, Data Systems Development Branch

P. Vavra
Chief, Information Systems Division

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
MANNED SPACECRAFT CENTER
HOUSTON, TEXAS
November 1968

ACKNOWLEDGMENTS

The project engineer wishes to acknowledge the efforts of Larry M. Hirsch, data systems engineer for the Philco Ford Corporation, Oron L. Schmidt, SETB project engineer, and Charles W. Brewer, ITT / Federal Electric Corporation technical writer. Their cooperation and contributions were instrumental in creating this document.

ACKNOWLEDGMENTS

The project engineer wishes to acknowledge the efforts of Larry M. Hirsch, data systems engineer for the Philco Ford Corporation, Oron L. Schmidt, SETB project engineer, and Charles W. Brewer, ITT / Federal Electric Corporation technical writer. Their cooperation and contributions were instrumental in creating this document.

TABLE OF CONTENTS

	TITLE	PAGE
1	SUMMARY	1
2	INTRODUCTION	2
2.1	SCOPE OF THE MANUAL	2
2.2	BACKGROUND	3
3	THE PROCESSES AND PROGRAMS DEVELOPED BY DSDB FOR SPEECH-TO-NOISE RATIO MEASUREMENT	6
3.1	THE PROCESSES DEVELOPED BY DSDB FOR SPNR MEASUREMENT	6
3.2	PROGRAMS DEVELOPED BY DSDB FOR COMPUTATION OF SPEECH-TO-NOISE RATIOS	8
3.2.1	Mean Square Value Speech-to-Noise Ratio Derivation Program	9
3.2.2	Amplitude Value Speech-to-Noise Ratio Derivation Program	12
3.2.3	Autocorrelation Function Speech-to-Noise Ratio Derivation Program	15
3.3	CORRELATION OF SPEECH-TO-NOISE RATIOS AND WORD INTELLIGIBILITY SCORES	21

TABLE OF ILLUSTRATIONS

	TITLE	PAGE
1	Equipment used in SPNR measurement	7
2	MSQ SPNR computation program	11
3	AMP SPNR computation program	14
4	An autocorrelation function of speech data	16
5	An autocorrelation function of noise data	17
6	ACF SPNR computation program	20
7	An SPNR-WI correlation graph	22

SECTION 1
SUMMARY

Noise degradation of Apollo voice communication has heretofore been measured at communication link outputs by the use of techniques which derive a signal-to-noise ratio (SNR). These measurement techniques have produced results which have not agreed with word-intellegibility evaluations performed at communication link outputs because of the use of a 1000-Hz tone by the SNR measurement technique to simulate voice input to communication links. The Data Systems Development Branch (DSDB) of the Information Systems Division (ISD) has investigated the possibility of measuring actual speech waveforms transmitted through the links rather than the transmitted tone which simulates speech waveforms. As a result of its investigations, DSDB has developed three techniques for measuring speech waveforms and from them deriving a speech-to-noise ratio (SPNR) measurement. Initial tests indicate that SPNR measurements made by the use of these techniques agree more closely with word-intellegibility evaluations. The SPNR measurement techniques developed by DSDB are now being tested by the Systems Engineering and Test Branch (SETB) of ISD in order to determine exactly their limitations and applicability.

SECTION 2
INTRODUCTION

2.1 SCOPE OF THE MANUAL

This manual presents a summary of the investigations into digital techniques of speech-to-noise ratio derivation for Apollo communication voice tapes. It describes the processes used by the Data Systems Development Branch (DSDB) to digitize the analog information present on voice tapes, and it describes the method used to compute the speech-to-noise ratio for the communication system from which the tape has been recorded. It discusses three programs for speech-to-noise ratio derivation from digitized signal waveforms.

Further information about speech-to-noise ratio measurement techniques is available in the following publications on file in the Information Systems Division Library:

a. Apollo Voice Intelligibility Measurement Techniques and Procedures, B68-3701 (U), January 1968

b. Development of a Speech-to-Noise Ratio Measurement Utilizing Digital Techniques, PHO-TN228, 12 June 1968

c. Development of a Speech-to-Noise Ratio Measurement Utilizing Analog Techniques, PHO-TN248, 6 September 1968

d. Verification Test Report of a Speech-to-Noise Ratio Measurement Method Utilizing Digital Techniques, PHO-TN270, 1 October 1968

SECTION 2
INTRODUCTION

2.1 SCOPE OF THE MANUAL

This manual presents a summary of the investigations into digital techniques of speech-to-noise ratio derivation for Apollo communication voice tapes. It describes the processes used by the Data Systems Development Branch (DSDB) to digitize the analog information present on voice tapes, and it describes the method used to compute the speech-to-noise ratio for the communication system from which the tape has been recorded. It discusses three programs for speech-to-noise ratio derivation from digitized signal waveforms.

Further information about speech-to-noise ratio measurement techniques is available in the following publications on file in the Information Systems Division Library:

a. Apollo Voice Intelligibility Measurement Techniques and Procedures, B68-3701 (U), January 1968

b. Development of a Speech-to-Noise Ratio Measurement Utilizing Digital Techniques, PHO-TN228, 12 June 1968

c. Development of a Speech-to-Noise Ratio Measurement Utilizing Analog Techniques, PHO-TN248, 6 September 1968

d. Verification Test Report of a Speech-to-Noise Ratio Measurement Method Utilizing Digital Techniques, PHO-TN270, 1 October 1968

2.2 BACKGROUND

Over the past 3 years the Information Systems Division (ISD) and the Space and Electronic Systems Division (SESD) have been actively engaged in testing and evaluating the Apollo communication system. During this testing program, techniques and procedures have been developed to measure the percentage of word intelligibility of all voice communication links in the system.

Voice communication among astronauts inside spacecraft, astronauts engaged in extravehicular activity (EVA), and mission support personnel on the ground has always been recognized as a requirement for a successful mission. To insure the reliability of Apollo voice communication links, extensive development and testing programs are carried out prior to the use of the links on an actual mission.

To measure the performance of the voice communication link, a measurement program has been developed. Under this program, a monosyllabic word intelligibility test was used to obtain word intelligibility (WI) scores for all voice tapes made with the Apollo Communication System (ACS). Voice tapes of the ACS were made in the Systems Engineering and Test Branch (SETB) Laboratory of ISD and sent to Fort Huachuca, Arizona, for WI scoring.

In preparation for measurement of voice intelligibility at the output of any communication link or configuration of communication links, source tapes to be used as communication system inputs are prepared under rigidly controlled conditions. These tapes are recorded in a quiet room, using the same types of suits and microphones to be used on actual missions, and voice input is formatted in the special manner required for evaluation at Fort Huachuca. A 1000-Hz tone and a noiseless interval are recorded just prior to the voice format.

From the output of a communication system link being tested, another tape is made. This tape, in the past, has been scored in two ways. A WI score has been given to it by the evaluation center, and SETB has used the transmitted tone and noiseless intervals to arrive at a signal-to-noise ratio (SNR).

A lack of correspondence between WI scores and SNR measurements, however, has caused ISD to investigate more accurate methods of electronically measuring the performance of voice communication links. Studies into possible sources of measurement error during SNR determination indicate that the method used at the present time does not account for either changes in the speech input level in the source tape or the presence of suit noise in the source tape. To minimize error from these sources, it was decided that measurement of actual voice and noise amplitude recorded on the data tapes during the monosyllabic word intelligibility test format was preferable to measurement of the 1000-Hz tone and link noise measurement. Because measurements made by this refined method are more dependent on the audio input to the communication link during speech transmission than they are on electronic signal simulation of the audio input, the calculation in which these measurements are used is called the speech-to-noise ratio (SPNR) calculation.

Special characteristics of SPNR measurement indicated that digital measurement techniques, rather than conventional analog measurement techniques, would be desirable. Response times of analog measurement devices such as RMS meters are much too slow for real-time analysis, and analysis over extended periods of time would cause undesirable delay in communication link performance evaluation. In addition, the number of measurements it would be necessary to make to accurately represent the complex waveforms of speech and noise over the entire testing period virtually eliminates the possibility of unautomated computation.

SETB requested the support of DSDB in developing techniques for computerized SPNR derivation. The current state of development of these techniques and programs is reported in this manual. After SPNR measurement techniques are permanently established, SETB will process all voice data tapes now stored, and procedures will be established to process further voice data tapes soon after they are made. The more accurate correspondence between word intelligibility and electronically measured speech-to-noise ratios will enable more accurate evaluation and prediction of communication link performance. The SPNR measurement techniques which are described in this manual are now being evaluated by SETB in order to determine their limitations and applications.

SECTION 3
PROCESSES AND PROGRAMS DEVELOPED BY DSDB FOR SPEECH-TO-NOISE RATIO MEASUREMENT

3.1 THE PROCESSES DEVELOPED BY DSDB FOR SPNR MEASUREMENT

Apollo voice tapes are recordings of a communication link output. The electrical output of the link recorder is a complex wave which is a function of time, noise, and speech. For any particular time during the recording, the waveform output of the recorder has one amplitude value. This amplitude value, which is the sum of instantaneous amplitude value of noise and the instantaneous amplitude value of speech, is the basic datum used in speech-to-noise ratio computation. Samples of waveform amplitude from the recording are digitized by an analog-to-digital converter and stored on magnetic tape. These values are the data input for the SPNR computation program (section 3.2).

To implement amplitude sampling and conversion from analog to digital representation, equipment already in operation in the DSDB Laboratory has been used. Apollo voice tapes are played on a Magnecord model 1028 recorder. The recorder is interfaced with a Scientific Data System (SDS) model A/D A30A analog-to-digital converter. The digital output information from the A/D A30A is recorded onto magnetic tape by a Data Machines Inc. (DMI) model 620 computer at a rate of 8330 samples per second to accurately represent a waveform with the bandwidth required for the Apollo communication links.

The IBM 360 model 44 computer located in the DSDB Laboratory is used for SPNR calculation from digital data input. Because of the incompatibility of the A/D A30A output format with the input format of the 360/44, the DMI 620 computer is used to reformat the A/D A30A output onto another reel of magnetic tape which is then transferred to the 360/44 (fig. 1).

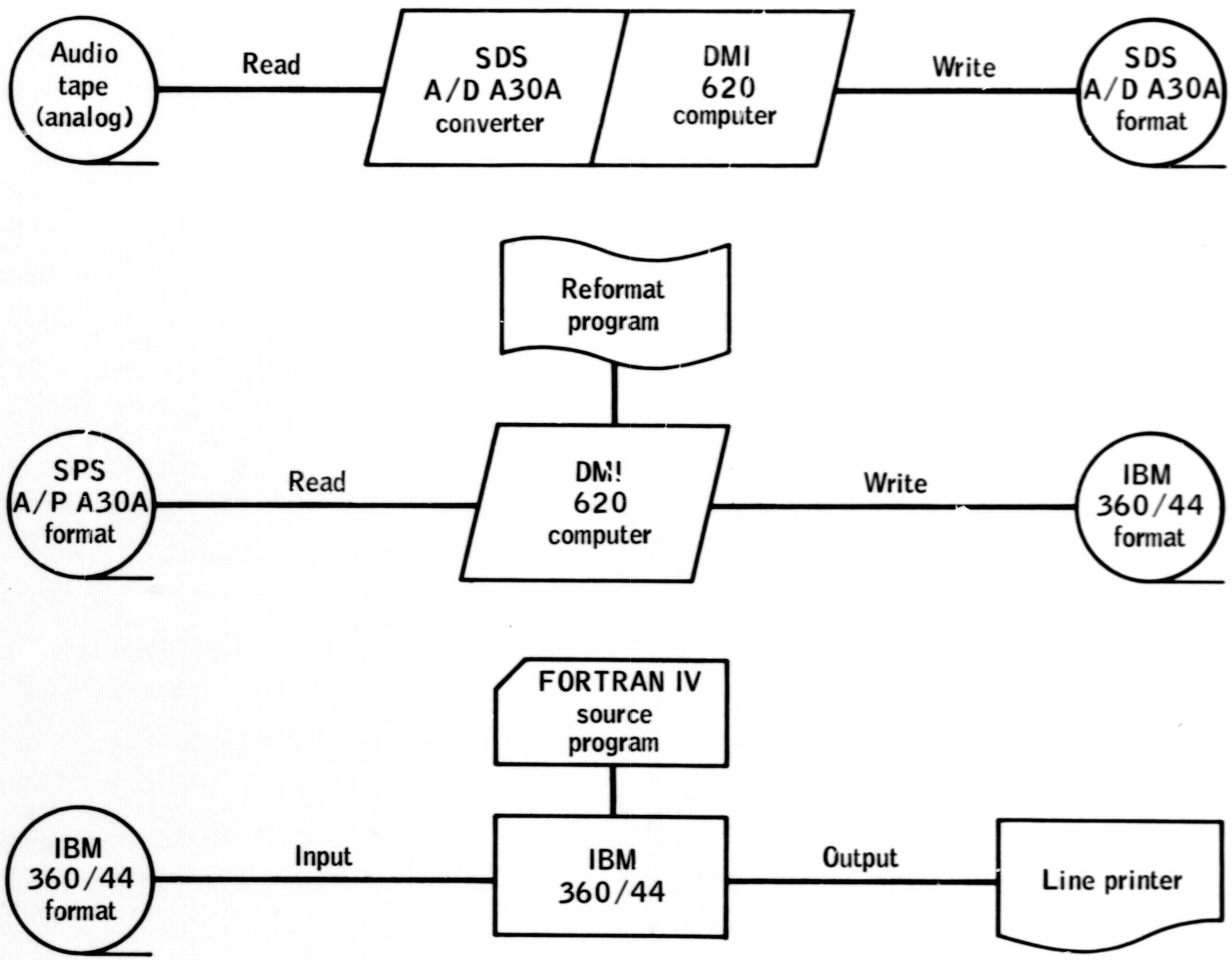

Figure 1.- Equipment used in SPNR measurement.

3.2 PROGRAMS DEVELOPED BY DSDB FOR COMPUTATION OF SPEECH-TO-NOISE RATIOS

By using the computational equipment described above, SPNR measurements have been made on Apollo voice data tapes by three different programs. Each program is written in FORTRAN IV, each uses the digitized waveform information as input, but each calculates the SPNR by a different method. Program 1 uses each waveform amplitude value in a mean square computation to arrive at numbers proportional to the waveform power content; it is therefore designated the mean square value (MSQ) program. Program 2 uses each amplitude value in a derivation of representative noise amplitude numbers and representative speech-plus-noise numbers before it calculates any waveform content values; therefore, it is called the amplitude value (AMP) SPNR computation program. Program 3 uses each waveform amplitude value in determining values of the waveform autocorrelation function and from these values it derives power values. Program 3 is therefore called the autocorrelation function value (ACF) computation program. Each of these programs is described in the following three sections.

3.2.1 Mean Square Value Speech-to-Noise Ratio Derivation Program

In SPNR derivation by the MSQ program, amplitude samples taken during a 3-second interval are used in a FORTRAN IV program to calculate values proportional to the power content of the waveform at the time of the sample. These relative power values are considered in 170-sample groups by the computer, and a relative power average is computed for each of the groups. Relative power average values are then classified by their amplitude into three categories: (a) those values representing only noise power, (b) those representing both noise power and speech power for speech of sufficient duration to be meaningful, and (c) those which, because of rapid change of value, cannot be classified in either (a) or (b). Those power values classified in the first two categories form the basis for calculating the SPNR over the 3-second interval to which they belong. SPNR values for periods of time longer than 3 seconds are computed by averaging the values obtained for the 3-second intervals contained in the period. The following list enumerates the steps taken in the program which derives speech-to-noise ratios by the mean square value calculation method:

(1) The squared values of 24 990 amplitude samples are averaged in 147 groups of 170; a power proportional value is thus obtained for each real-time communications interval of 20.4 milliseconds, a time determined to be one-tenth the duration of the shortest vowel sound in speech. Several average power proportional values of this duration would occur within the time taken for significant speech to occur.

(2) Each average power-proportional value is then compared to the two values consecutive to it. Each time three of these values agree within 1 dB, the average of the three is entered into the computer memory.

(3) The values now stored in the memory are arranged in ascending order.

(4) The least value in memory is defined to be representative of noise power alone, and all values in memory which differ from it by 1 dB or less are also considered noise. They are stored in a noise array.

(5) All values which differ from the least value by 3 dB or more are defined to be speech-plus-noise power-proportional values. They are stored in a speech-plus-noise array.

(6) After consideration of 24 990 sample values, 3 seconds of communication data, power-proportional noise values are averaged to derive a single value to represent noise power and power-proportional speech-plus-noise values averaged to derive a single value to represent speech-plus-noise power.

(7) The values obtained in step 6, which represent speech-plus-noise power and noise power are substituted in the equation which follows in order to derive a communication system SPNR for a 3-second period:

$$\text{SPNR} = 10 \log \frac{(\text{SPEECH} + \text{NOISE}) - \text{NOISE}}{\text{NOISE}}$$

(8) The 3-second SPNR is printed and stored in memory. A running average is computed from the values stored in memory, and when all amplitude samples are processed in the steps listed above, the average SPNR for the duration of the data tape is printed.

A flow chart illustrating the MSQ SPNR computation program is presented in figure 2.

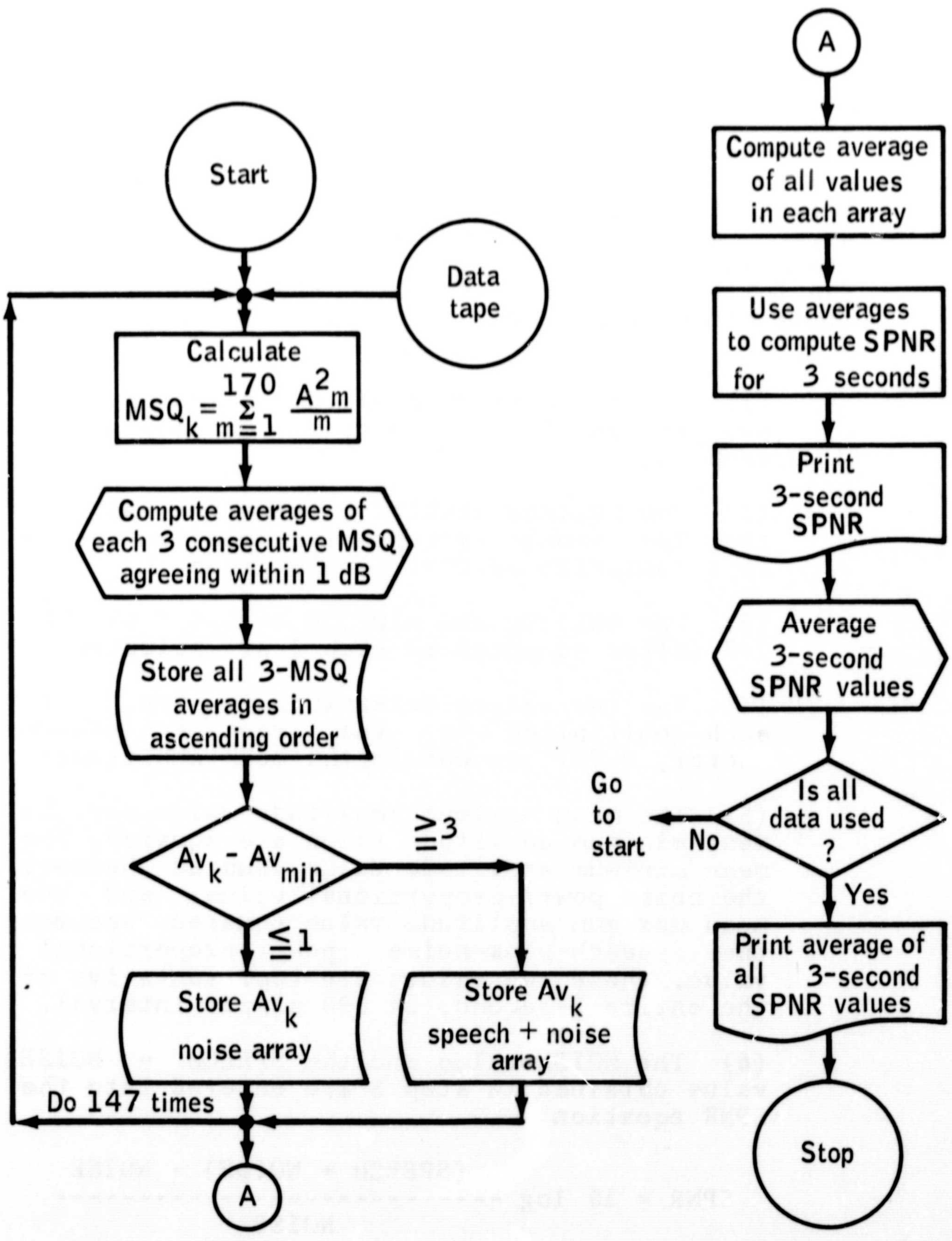

Figure 2.- MSQ SPNR computation program.

3.2.2 Amplitude Value Speech-to-Noise Ratio Derivation Program

Determination of SPNR value by the amplitude value (AMP) calculation program is similar to determination by the mean square value (MSQ) program. The AMP program is faster than the MSQ program because of assumptions made to simplify calculation. In the AMP program, sample values are obtained in the same manner as in the MSQ program, but the AMP program first discriminates between those amplitudes values representative of noise and those representative of speech-plus-noise. This saves computation time. The following list describes the computational steps taken in the AMP SPNR derivation program.

(1) 24 990 waveform amplitude value samples are grouped in 147 groups containing 170 samples.

(2) The maximum amplitude value in each of the 147 sample groups is selected as a representation of that sample group.

(3) The maximum and minimum values from the 147 values computed in step 2 are selected.

(4) The two values determined in step 3 are each multiplied by the root-mean-square factor, 0.707, to obtain the mean amplitude.

(5) The mean maximum amplitude value and the mean minimum amplitude value are squared. The mean minimum amplitude value squared becomes the noise power-proportional value, and the mean maximum amplitude value squared becomes the speech-plus-noise power-proportional value. These two values are representative of the entire 3-second, 24 990 sample interval.

(6) The NOISE value and the SPEECH + NOISE value obtained in step 5 are entered into the SPNR equation:

$$SPNR = 10 \log \frac{(SPEECH + NOISE) - NOISE}{NOISE}$$

(7) The APNR for the 3-second, 24 990 sample interval is printed and entered into memory; the values that accumulate in the memory are used in a computation of a running average. The value of this running average, when all information has been processed through the steps listed above, is printed as the SPNR value for the communication data tape.

Figure 3 is a flow chart illustrating the amplitude value speech-to-noise ratio derivation program.

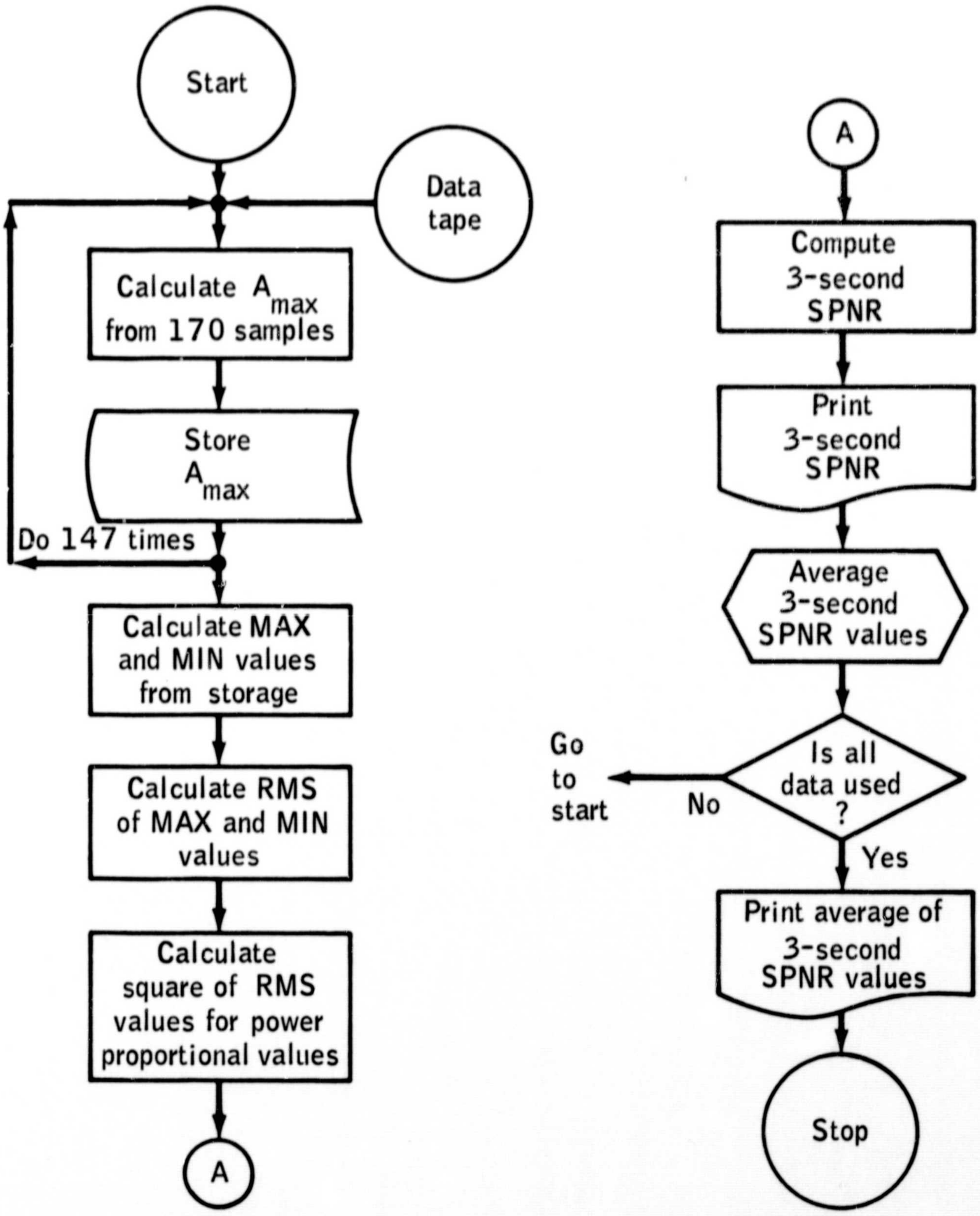

Figure 3.- AMP SPNR computation program.

3.2.3 Autocorrelation Function Speech-to-Noise Ratio Derivation Program

Amplitude values entered into the IBM 360/44 program are representative of either noise waveform amplitude or of speech-plus-noise waveform amplitude. Noise waveform amplitude values are randomly distributed through time, whereas speech amplitude values are distributed with a certain regularity or periodicity. The autocorrelation value (ACF) SPNR derivation program uses this periodicity to positively differentiate between amplitude values representative of speech and those representative of noise.

The autocorrelation function of random data describes the general dependence of the values of data at one time on the values of data at other times. Periodic waveforms, such as the sine wave function found in the vowels, nasals, and liquids of speech, have an autocorrelation function which persists over the duration of the waveform, whereas random-valued waveforms have an autocorrelation function which quickly diminishes to zero. The autocorrelation function is therefore a powerful tool for detecting deterministic data masked by a random background.

In determining SPNR values by the ACF program, amplitude values entered into the IBM 360/44 are considered in groups, just as they are in the MSQ and AMP determination programs. Three terms of the autocorrelation function for each group of amplitude values are calculated; experience shows the first three values alone to be sufficient. Figure 4 shows a plot of 50 values of an autocorrelation function for a 595-sample communications interval. The values are normalized; that is, they are represented proportional to the first value, which is plotted as unity. The first three values of the function are positive, and a plot of the entire function appears as a slowly damped periodic wave; therefore, these amplitude values are considered to be from a speech-plus-noise segment of the communication waveform. Figure 5 shows another autocorrelation function plotted over 50 values for another communications interval. The curve in this autocorrelation function shows the amplitude

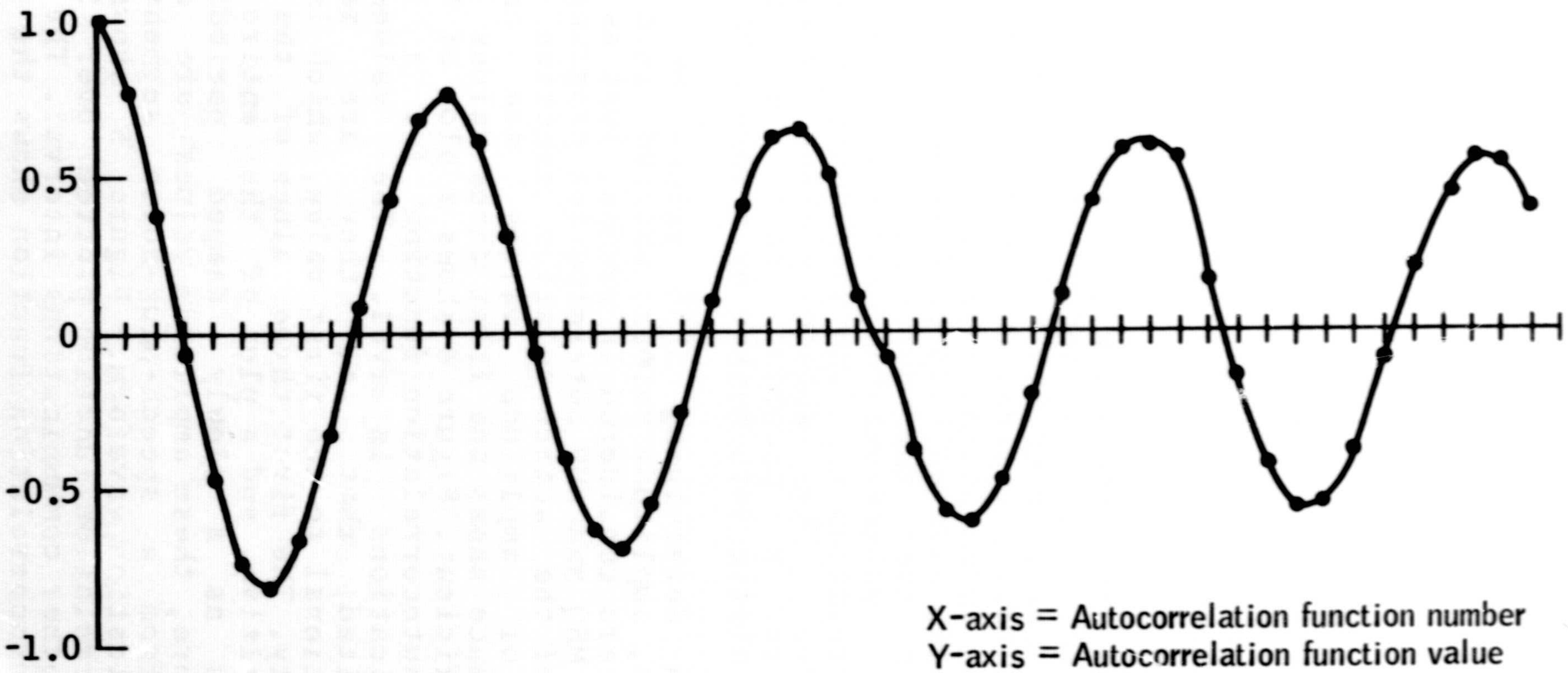

Figure 4.- An autocorrelation function of speech data.

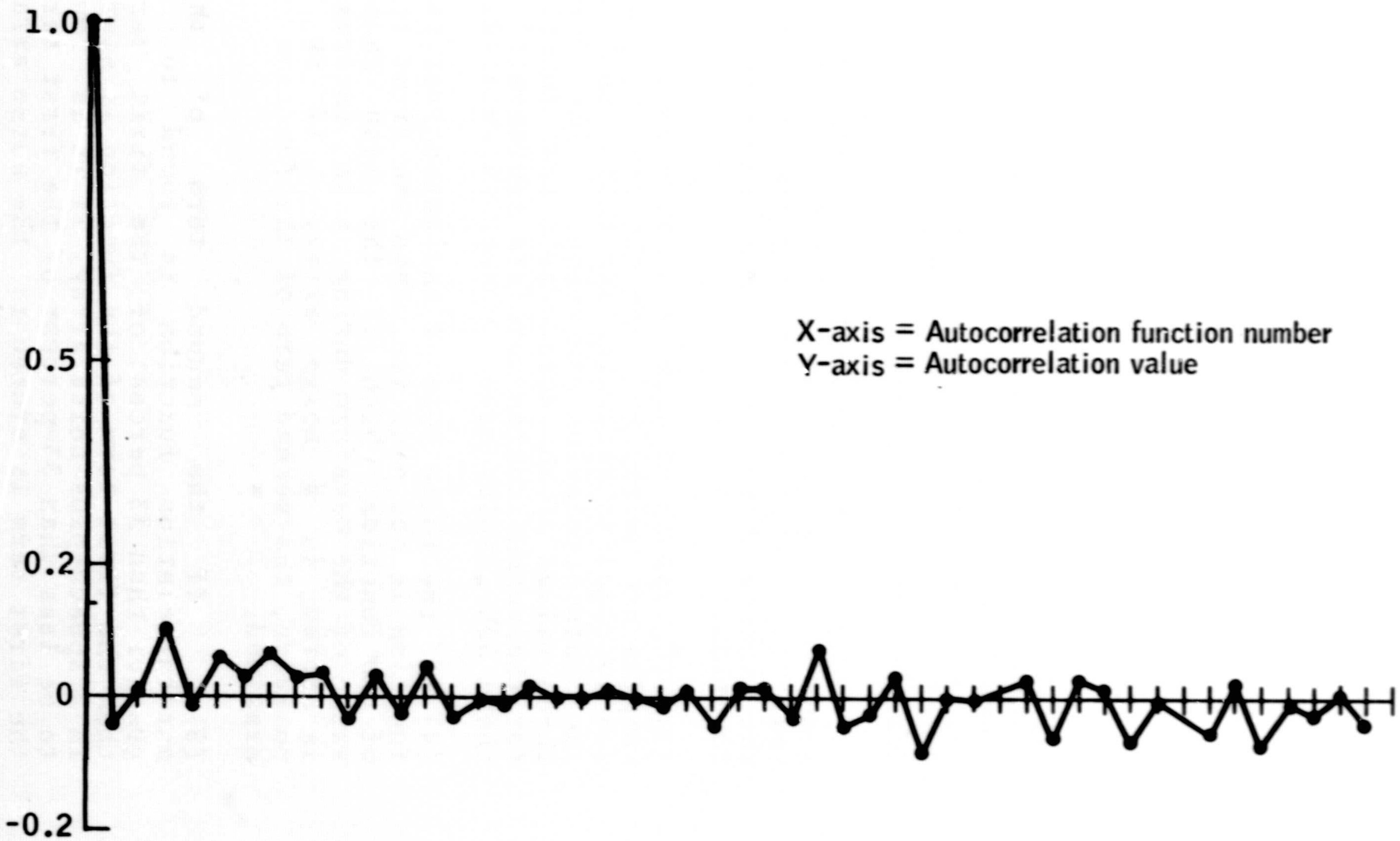

Figure 5.- An autocorrelation function of noise data.

values which it represents to be values of noise alone.

In determining the SPNR of an Apollo data tape by the ACF program, SPNR calculations are made over a communication interval of 3 seconds. During this interval, 24 990 samples of waveform amplitude are recorded. The program analyzes this data at the rate of 595 samples per computation in order to include autocorrelation function representation of the lowest frequency waveforms present in the communication waveform. The program operates according to the following steps:

(1) The values of the 595 consecutive samples are squared and averaged to concomitantly derive the mean square value of the waveform at the time of sampling and the first term of the autocorrelation function for that group of samples.

(2) The values of every two neighboring data samples in the group are multiplied, and the resulting 594 products are averaged to derive the second term of the autocorrelation function.

(3) The values of each two alternate data samples are multiplied, and the resulting 594 are averaged to determine the third term of the autocorrelation function. The first three values of the autocorrelation function for the communication waveform represented by these 595 samples have now been calculated.

(4) If the third term of the autocorrelation function is not positive, then the first term of the function, which is the mean square value of the waveform during this interval, is stored in a noise array. If it is positive, the second term of the function is examined.

(5) If the second term of the autocorrelation function is found to be greater than 33 percent of the first term, then the first term of the function is stored in a speech-plus-noise array. If it is found to be less than 33 percent of the first term, the first term is stored in the noise array.

It was determined by examination of many autocorrelation function graphs of speech-containing waveforms that the second term was always greater than 33 percent of the first.

(6) The program then repeats the steps listed above for the next 595 data points, storing values in the noise and speech-plus-noise arrays.

(7) The minimum mean square value stored in the noise array is determined. All values in the array which are greater than three times the minimum are assumed to be representative of at least some speech, and are therefore excluded from the array.

(8) After all 24 990 samples have been used in the computation steps listed above, values stored in each of the arrays are averaged to obtain one NOISE value and one SPEECH + NOISE value to represent the 3-second interval.

(9) The two values arrived at in step 8 are entered into the SPNR equation, and an SPNR for the 3-second interval is calculated, stored for future calculation, and printed.

(10) The first nine steps of the program are repeated for the next 3-second interval in the communication, and a running average for elapsed communication time is calculated. When all data has been processed, this average is printed and the program is terminated.

A flow chart illustrating the ACF program for SPNR derivation appears in figure 6.

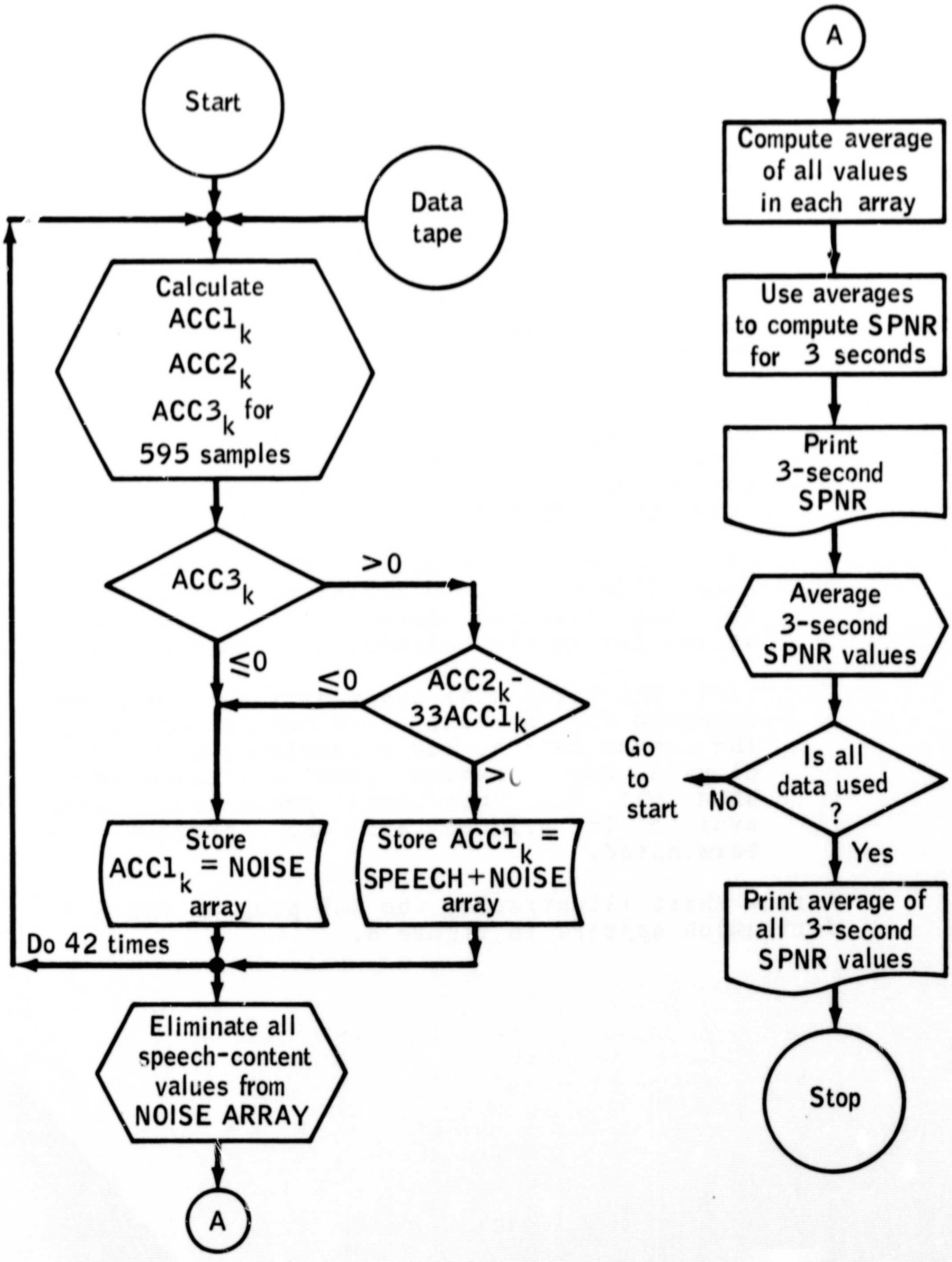

Figure 6.- ACF SPNR computation program.

3.3 CORRELATION OF SPEECH-TO-NOISE RATIOS AND WORD INTELLIGIBILITY SCORES

The computation processes and programs described in this document represent the current state of development of digital SRNR techniques in the DSDB Laboratory. The SPNR values which they provide correlate closely with WI scores. Figure 7 is a typical SPNR - WI correlation graph in which SPNR values for several data tapes are plotted against WI scores for the same tapes. The lack of exact correlation shown by this graph is accounted for by factors other than noise which have an effect on word intelligibility, such as distortion generated in the link. Many correlation graphs such as this one have been examined in the DSDB Laboratory to refine digital SPNR derivation techniques and programs to their present level of accuracy and sophistication. The computation programs described in this document have been found empirically to be practical and efficient programs for deriving speech-to-noise ratios for communication tapes. Further evaluation of these techniques, undertaken to determine more exactly their limitations and their applicability, is now being conducted by SETB.

Investigation has already begun to develop other processes and programs which combine the accuracy of the ACF program with the simplicity and economy of the AMP program by using a digital spectrum analyzer recently installed in the DSDB Laboratory.

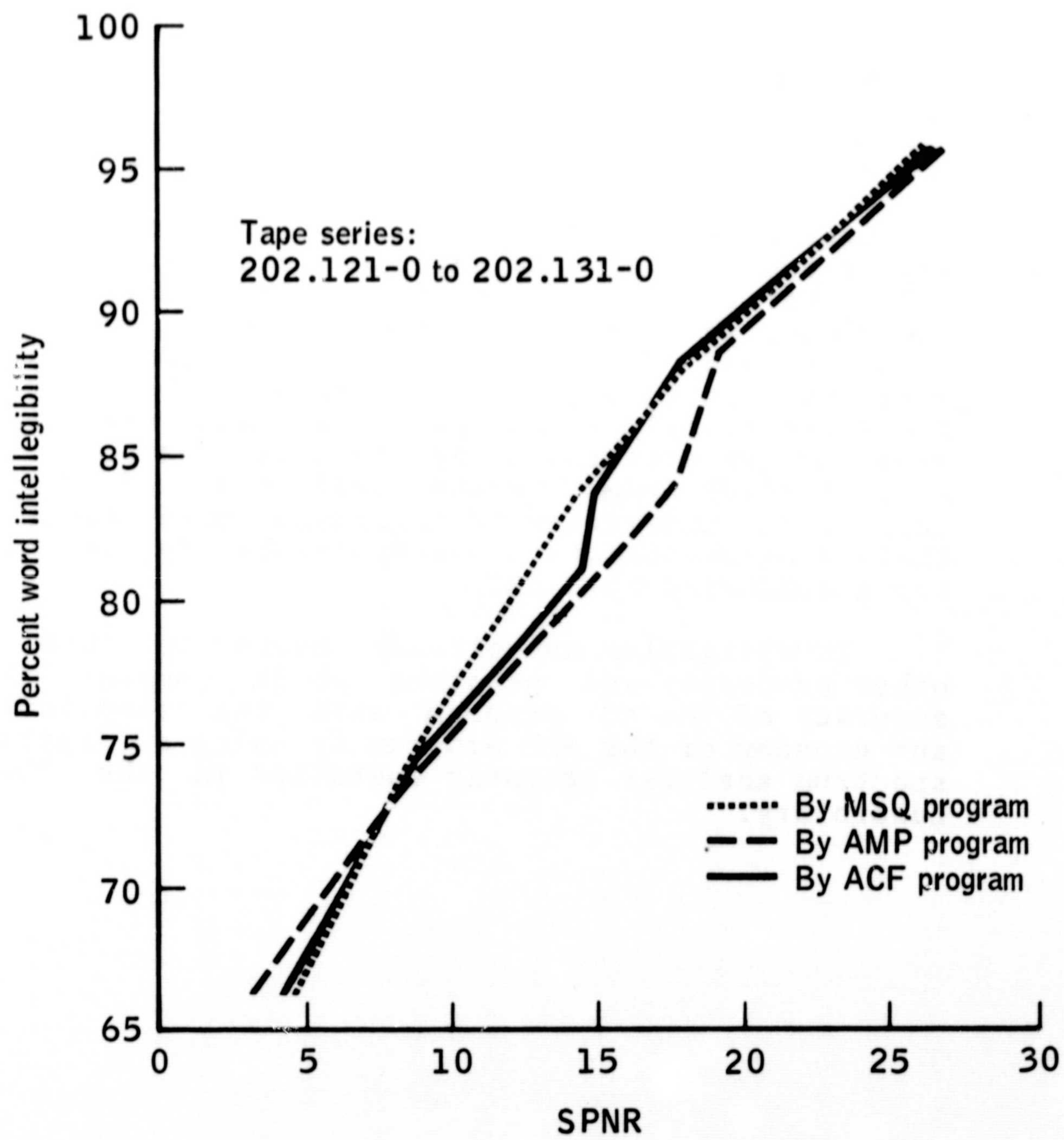

Figure 7.- An SPNR-WI correlation graph.