NASA CR65425

RESEARCH AND DEVELOPMENT OF AN OPTIMIZED LABORATORY SPEECH-COMPRESSION SYSTEM FOR SPACECRAFT APPLICATION

FINAL REPORT

25 May 1966

Contract No. NAS 9-4523

[redacted]

Period Covered:

25 May 1965 to 25 May 1966

Prepared For

Manned Spacecraft Center
National Aeronautics and Space Administration
Houston, Texas

PHILCO
A SUBSIDIARY OF Ford Motor Company
COMMUNICATIONS AND ELECTRONICS DIVISION
BLUE BELL, PENNSYLVANIA 19422

RESEARCH AND DEVELOPMENT OF AN OPTIMIZED LABORATORY SPEECH-COMPRESSION SYSTEM FOR SPACECRAFT APPLICATION

FINAL REPORT

25 May 1966

Contract No. NAS 9-4523

Period Covered:

25 May 1965 to 25 May 1966

David M. Jurenko
James M. Loe
John P. Schultz

Prepared For:

Manned Spaceflight Center
National Aeronautics and Space Administration
Houston, Texas

PHILCO
A SUBSIDIARY OF Ford Motor Company
COMMUNICATIONS AND ELECTRONICS DIVISION
BLUE BELL, PENNSYLVANIA 19422

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF ILLUSTRATIONS (Continued)

SECTION 1

INTRODUCTION

This final report is submitted in accordance with the Requirements of Contract NAS 9-4523 with the Manned Spacecraft Center of the National Aeronautics and Space Administration. The program was monitored by Mr. Carrington Stewart, of NASA-MSC.

1.1 OBJECTIVE

The objective of the program was to develop a speech bandwidth-compaction technique, based on an electrical analog of the hearing mechanism, which extracts the slowly varying parameters that define the invariant descriptors necessary for incremental speech-sound recognition. The parameters shall be capable of transmission in a 160-Hz communication channel, and, at the receiver, should result in resynthesized speech having a phonetically balanced word intelligibility test score of 80 percent. The breadboard was aimed at minimizing circuit complexity, thereby yielding size, weight, and power consumption of a developed unit.

1.2 STUDIES

The following two studies concerning the Narrow-Band Speech-Compression (NBSC) system were also performed and published under separate cover.

1.2.1 Spacecraft Applications Study

A study was made of the NBSC in order to determine its advantages when used in the Apollo spacecraft.

1.2.2 Microminiature Study

Another study of the NBSC was made in order to delineate the ease with which microminiature circuit techniques could be applied; and also, the minimum volume, weight, size, and power consumption that would result from a microminiature NBSC system.

SECTION 2

PROGRAM SUMMARY

2.1 SINGLE-EQUIVALENT FORMANT (SEF) CONCEPT

This program centered around a new concept in speech compression, one that extracts a single-equivalent formant of normal speech, that the human hearing mechanism recognizes as intelligibile information.

The feasibility of this approach has already been proven in the laboratory, but this Section is intended to restate the theory and also familiarize the reader with the SEF concept.

The fact that a single formant could evoke vowel responses in the human was observed early in the 1800's by Willis and others.[1] Much more recently this same fact was reported by Delattre et.al.[2]

The presence of this phonetic equivalence between single-formant sounds and human vowels can best be shown by comparing the stimulus and perceptual response for the two types of sounds. This equivalence is shown in Figure 2-1. In this figure, 10 vowels are plotted on the horizontal axis, and their corresponding formant stimulus frequency is plotted on the vertical axis. For the human vowels, each of the first three formant frequencies, as taken from the Peterson-Barney data[3] is plotted against its corresponding perceived vowel response. The heavy line superimposed on this data represents the perceived vowel response to a single-formant stimulus. Note that the single-formant response does not correspond to any one formant for all the vowels. In the region of the back vowels it is closes to the first formant, while for the first vowels, it approaches the second formant. Note in these results that each three-formant human vowel can be replaced with an equivalent single formant without destroying

---

1. Lord Rayleigh, The Theory of Sound, Volume II, pp. 470, 478, Macmillan and Co.

2. P. Delattre et.al., "An Experimental Study of the Acoustic Determinants of Vowel Color," Word, Vol. 8, No. 3, pp. 195-210, December 1952.

3. G. Peterson and H. Barney, "Control Methods Used in a Study of the Vowels," JASA, Vol. 24, No. 2, pp. 175-184, March 1952.

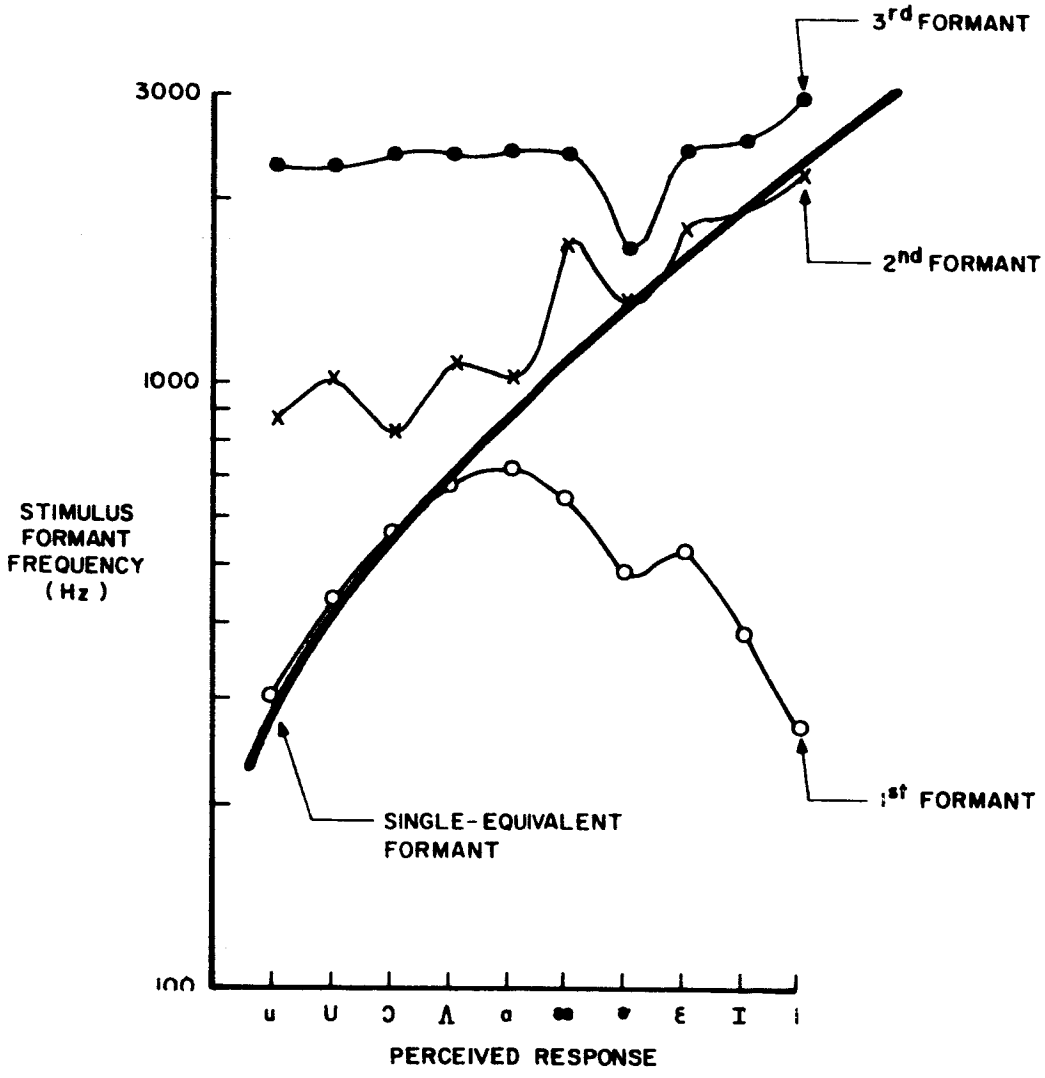


Figure 2-1. Correlation of the Single-Equivalent Formant and Human-Vowel Formant Locations

its phonetic value. Further experiments have shown that when equivalent-formant amplitude and voicing are added as parameters, a single-formant equivalent exists for all speech sounds; vowels, nasals, fricatives, voiced-fricatives, plosives. Thus, it is possible to replace three-formant human speech with its single-formant equivalent, thereby reducing by 2 to 1 the information needed to specify the phonetic content of speech.

To aid in understanding the latter part of this report, it is necessary to describe the factors involved in single-equivalent formant speech. It is postulated that when a human hears a multi-formant sound, as in human speech, his attention focuses upon only one formant, called the dominant formant. The presence of any other formants (called recessive formants) serves only to shift the perceived phonetic value slightly away from that of the dominant formant. This process is analogous to the hue shift that occurs in the human eye when two different unmixed colors are presented simultaneously. The color upon which the eye has affixed its attention (analogous to the dominant formant) is perceived as shifted slightly in hue because of the presence of the adjacent different color (analogous to the recessive formant).

This dominant and recessive nature of formants is revealed by a series of listening tests; the results of these tests are shown graphically in Figure 2-2. The figure shows the perceived phonetic value of two formant vowels as a function of their dominant and recessive formants. Areas of equal phonetic value are bounded by the contour lines. It is quite obvious that for any two formant frequencies, two phonetic values could exist, depending upon which formant is dominant. Thus, the mechanism by which dominance is assigned in the human is of great importance. Further experiments also show that the relative amplitude of the two formants is the principal factor in determining which formant is dominant. It was found that if a high-frequency pre-emphasis of 6-9 db/octave were applied on the speech spectrum, a simple measurement of the formant amplitudes determined dominance–the largest amplitude being dominant, the second largest being recessive.

It should further be noted in Figure 2-2 that the diagonal line represents the value of the single-equivalent formant and that it intersects all phonetic areas in the diagram. Thus, no vowels are excluded from this equivalence.

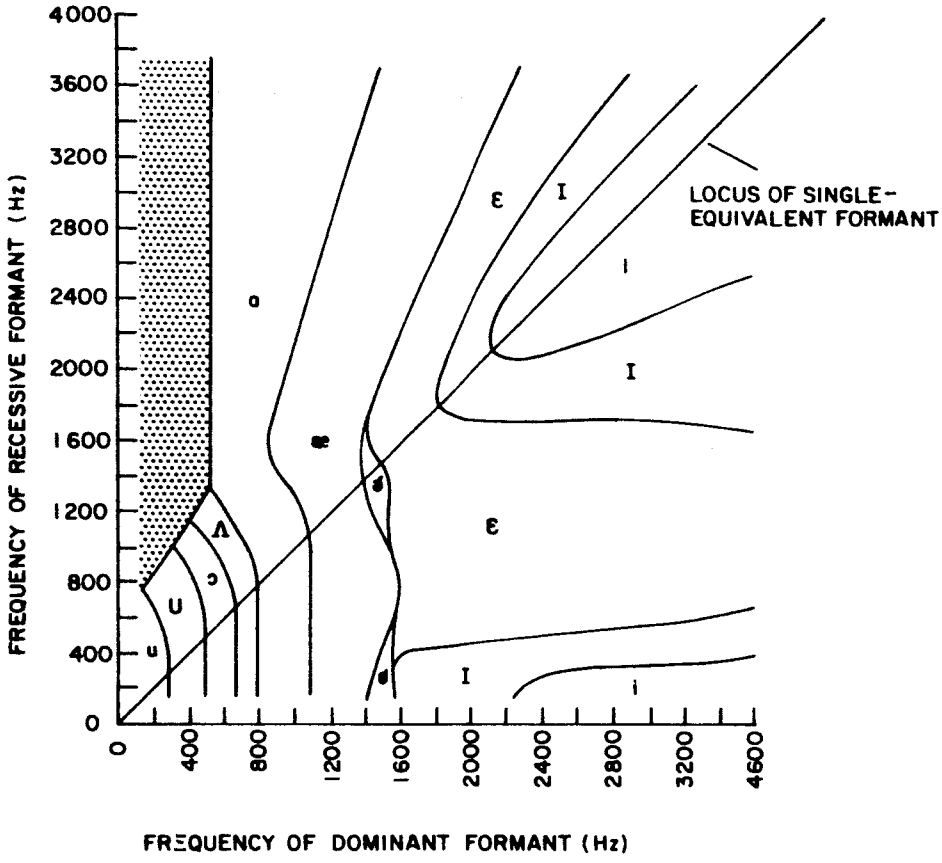


Figure 2-2. Dominant-Recessive Formant Determination of Vowel Color

2.1.1 Single-Equivalent Formant Frequency Extraction

The extraction of a parameter "approximately equal" to the single-equivalent formant is based upon the characteristics of the speech wave and the psychological factors just described. Figure 2-3A shows the formation of a three-formant speech sound. The shock of the glottal wave train excites the various vocal cavities, producing a series of damped sinusoids. The ringing frequencies of the damped sinusoids are the formant frequencies. The amplitude and formant frequencies are thus controlled by the movement of the articulators. The addition of these formant energies produces the complex wave known as speech. Figure 2-3B, C, and D show in detail how the addition of these waves is affected by the relative formant amplitude. When the first formant is larger in amplitude than the second formant, Figure 2-3B, the period of the first major oscillation of the complex wave (a result of the glottal excitation) is approximately equal to the period of the first major oscillation of the largest or first formant. The same holds true when the second formant is larger than the first, as shown in Figure 2-3C. The period of the major oscillation of the complex wave is again approximately equal to the period of the larger second formant. However, when both formants are of equal amplitude (Figure 2-3D), the resultant period of the major complex wave oscillation is different. It is the average value of the two formant periods.

Referring again to the nature of the amplitude of the dominant formant in human speech, recall that after a 9 db/octave pre-emphasis, the largest amplitude formant will be the dominant formant. Replotting Figure 2-1 to show only the largest amplitude formant for human speech after a 9-db pre-emphasis produces the curve of Figure 2-4. An examination of the curve shows that, for the back vowels, the first formant is dominant and nearly equals the frequency of the equivalent formant. With the front vowels, a similar situation holds true for the second formant. The central vowel, however, has first and second formants of nearly equal amplitude and, conveniently, the frequency of the single-equivalent formant is approximately an average of the first- and second-formant frequencies in this region. Thus, it can be seen that the period of the first excursion of the complex wave at each glottal pulse provides a measure of the single-equivalent formant frequency.

The block diagram shown in Figure 2-5 shows how this process is implemented. The input speech is amplified and then fed into a high-frequency pre-emphasis network, and then to a Schmitt trigger set to clip about the zero axis. Thus, the negative excursion of the Schmitt trigger indicates when the oscillation of the input wave crosses the zero axis in a negative

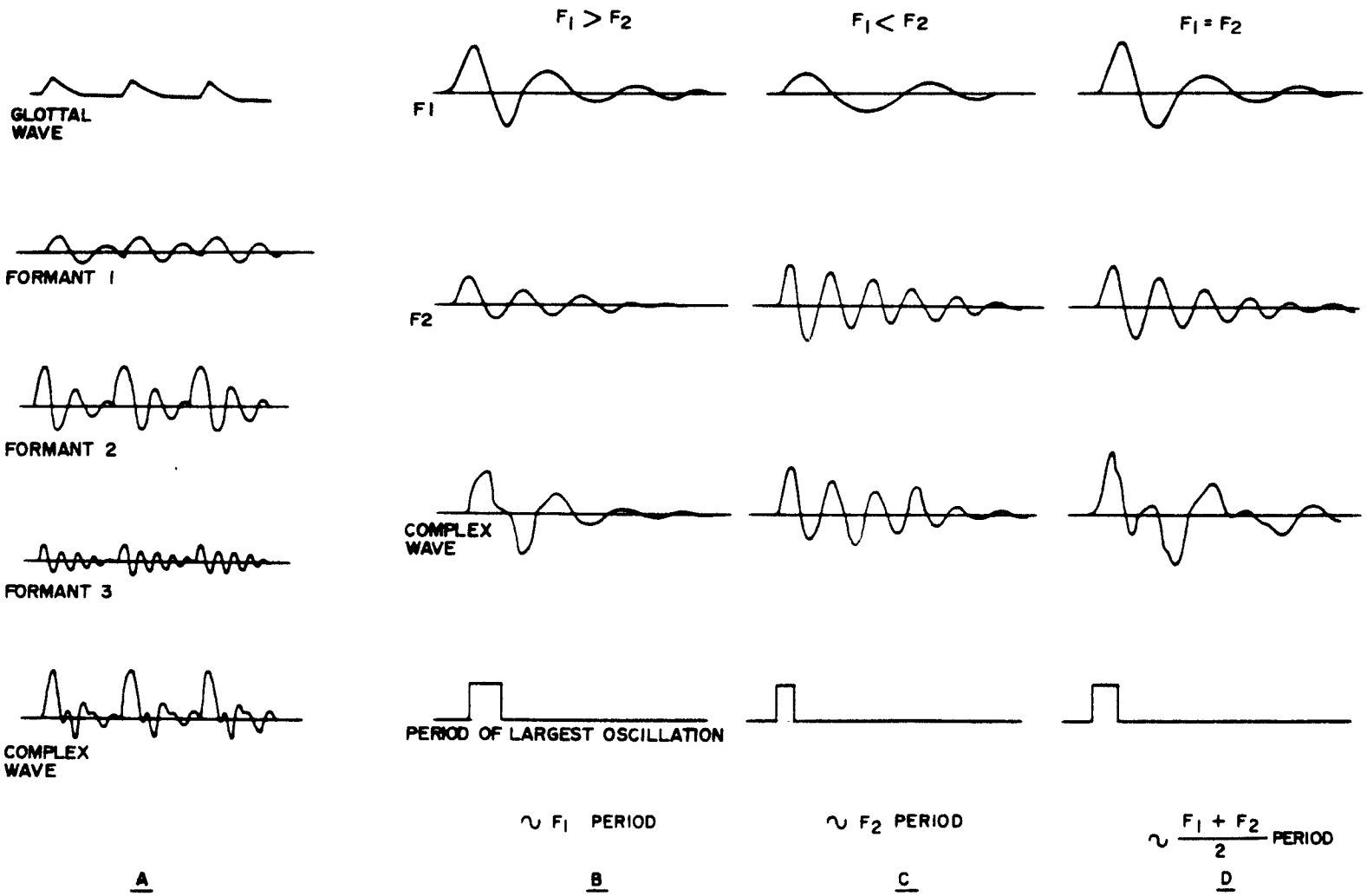


Figure 2-3. Formation of the Complex Speech Wave

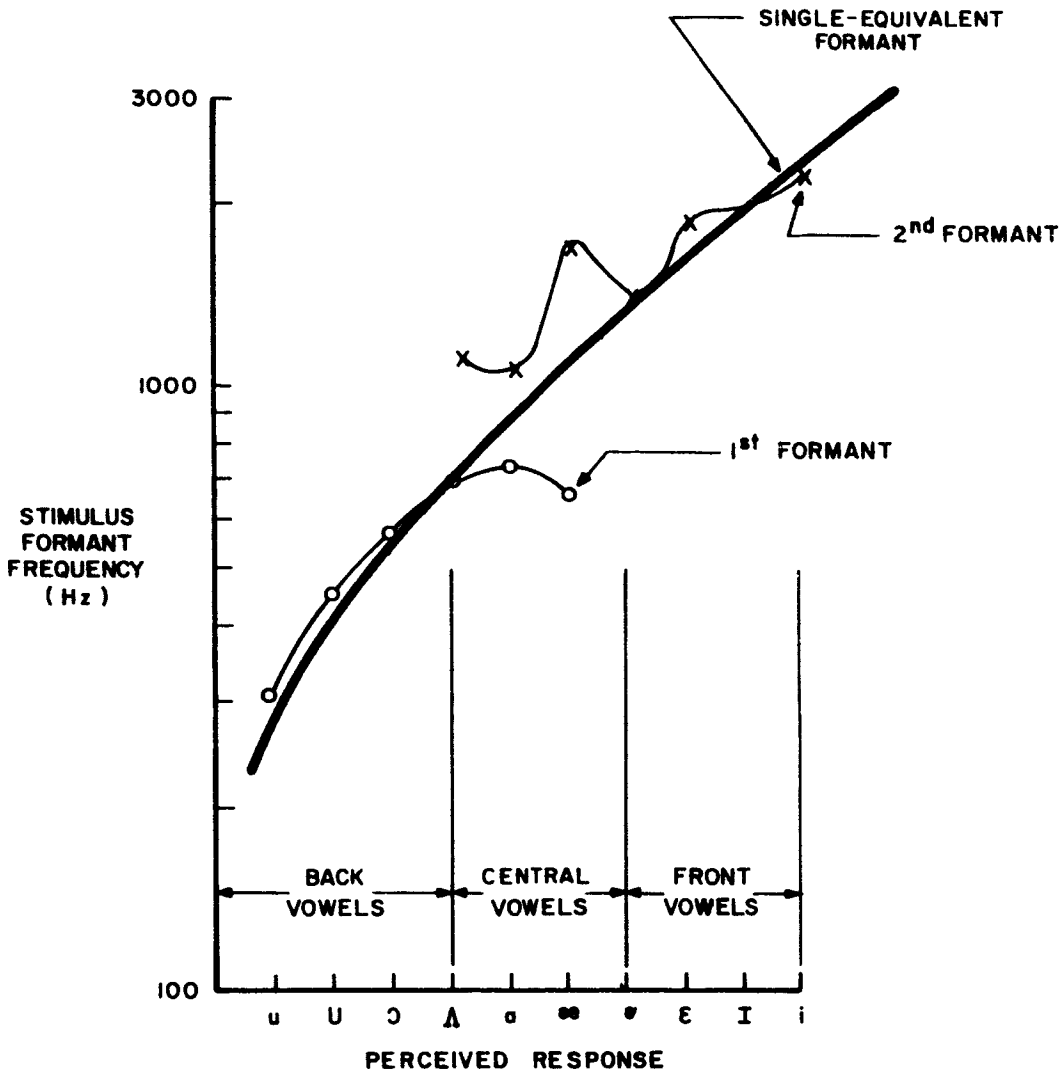


Figure 2-4. Comparison of Largest-Amplitude Human Formants and the Single-Equivalent Formant

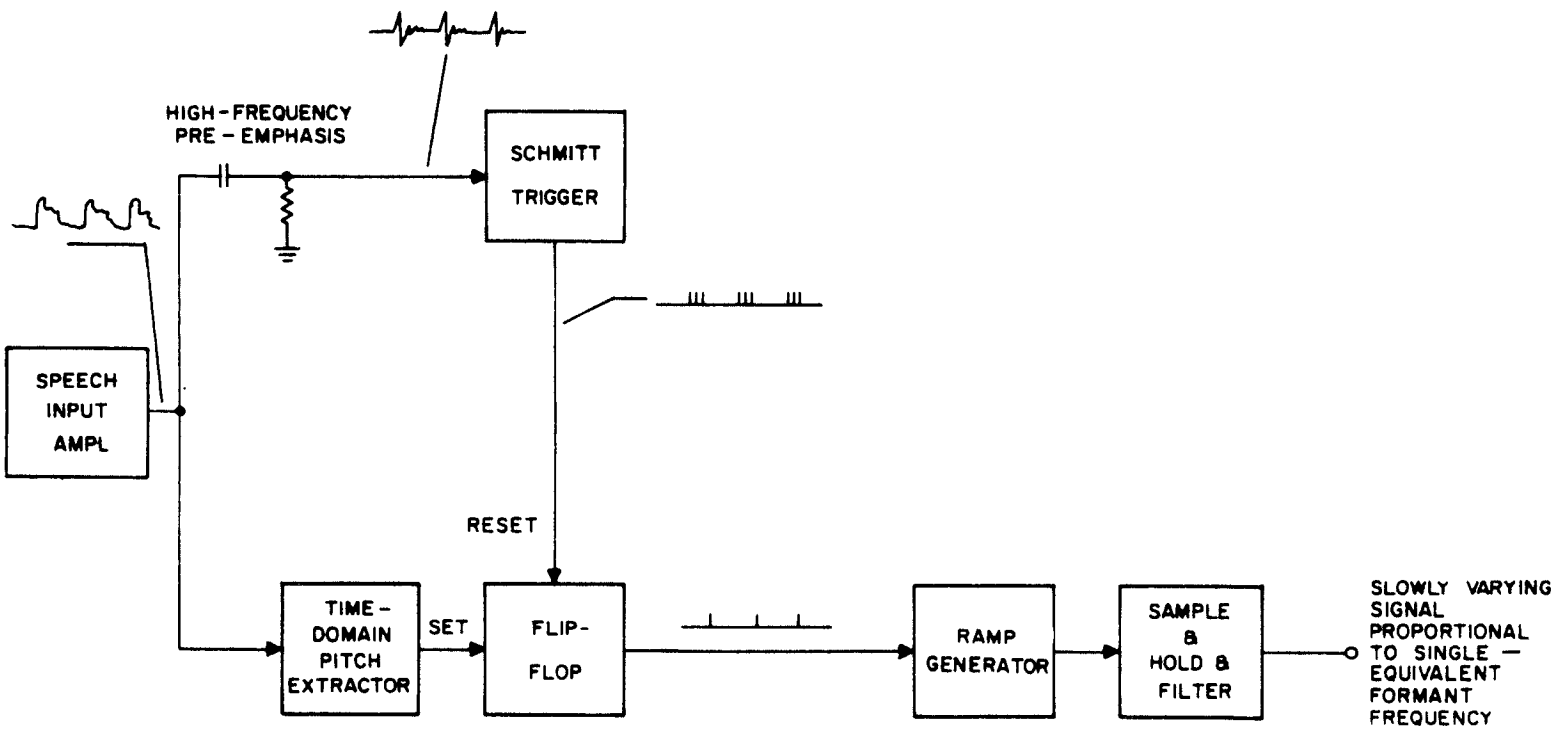


Figure 2-5. Single-Formant Extraction

direction. A time-domain pitch extractor indicates when the glottal pulse occurs. This sets a flip-flop which, in turn, is reset by the first following negative-going excursion at the Schmitt trigger. The output of the flip-flop is thus a pulse-length-modulated signal proportional to the period of the single-equivalent formant. A ramp generator and a sample-and-hold circuit then produce a continuous, slowly varying signal proportional to the frequency of the single-equivalent formant.

2.2 PROGRAM SUMMATION

This report discusses study and experimental progress aimed at developing an optimized laboratory speech-compression system for spacecraft application. Using the Single Equivalent Formant (SEF) concept as a starting point, research and design work was carried out in order to develop a breadboard system capable of being transmitted over a 160-Hz bandwidth, have an intelligibility of 80 percent at a S/N ratio of 20 db, and be easily adaptable to microminiaturization.

This work was accomplished through research and design in the following four areas.

2.2.1 Refinement and Simplification of Parameter Extractors

The purpose of this task was to breadboard and study the existing SEF, amplitude, pitch, and voicing extractors. This resulted in refinements in the amplitude, pitch, and voicing extractors, and a departure from the SEF concept in that, ultimately, two formant frequencies are extracted with frequency domain extractors rather than a single formant extraction by a pitch-synchronous extractor.

2.2.2 Refinement and Simplification of Synthesis Process

The purpose of this task was to breadboard and study the existing SEF synthesis process. The multivibrator synthesizer technique was replaced by a ringing-filter technique termed the Active Voltage-Tuned Filter (AVTF). Design of this type of filter utilizing matched field-effect transistors (FET) is reported in Appendix C.

2.2.3 Development of Three-Formant Synthesis

The purpose of this task was to study the feasibility of re-creating three-formant speech from the single-formant parameter. Study of this

technique has concluded that it is feasible, and could possibly lead to the very economic phoneme transmission system with added experimental effort.

2.2.4 Development of the Multiplexing Technique

The multiplexing technique chosen as the one needed for a minimum-bandwidth system was Pulse-Amplitude Modulation (PAM). This multiplexing technique had a stringent requirement in that approximately 120 Hz of information was to be transmitted in a 160 Hz bandwidth. Two techniques to cancel the crosstalk distortion were investigated: the Return to Complement (RTC) and the Finite Memory Filter (FMF). Ultimately, tests showed the FMF system as the more optimum and simpler of the two systems.

SECTION 3

FACTUAL DATA

This section delineates the research and design work accomplished during the program. Since research on the program was built around the basic SEF concept and existing circuitry, it is only logical that this Section should follow, chronologically and experimentally, the progress of the program. Therefore, the following factual data records the step-by-step process through which the experimenters ultimately arrived at the decision to build a two-channel analyzer-synthesizer system.

The final Narrow-Band Speech-Compression (NBSC) system operates at an information rate of 100 Hz. It is multiplexed by Pulse Amplitude Modulation (PAM), filtered so as to fit in a 160-Hz transmission channel, demultiplexed using a digital approximation of the transversal filter equalizer[1] termed the Finite Memory Filter (FMF), and synthesized using an Active Voltage Tuned Filter. Synchronization is achieved using threshold detection of the sync pulse whose integrity is maintained by immediately preceding it with a blank or zero information channel.

3.1 PRELIMINARY DATA

By the time design work was started on the analyzer-synthesizer portion of this program, another program with similar goals and approach was well under way. At the time, this program used an analyzer based upon an SEF extractor, pitch extractor, voicing extractor, and amplitude extractor. The synthesis technique involved the use of a pitch oscillator which cohered a square-wave SEF oscillator. The square-wave was first converted to a damped square-wave train in a switch modulator. A series of integrations and integration corrections converted this signal to a damped sine wave. The amplitude information was then applied in a linear amplitude modulator. This synthesis technique is quite complex, as shown in the block diagram of Figure 3-1.

---

1. Locky, R. W., Automatic Equalization for Digital Communication, Bell System Technical Journal, Vol. 44 No. 4, April 1965.

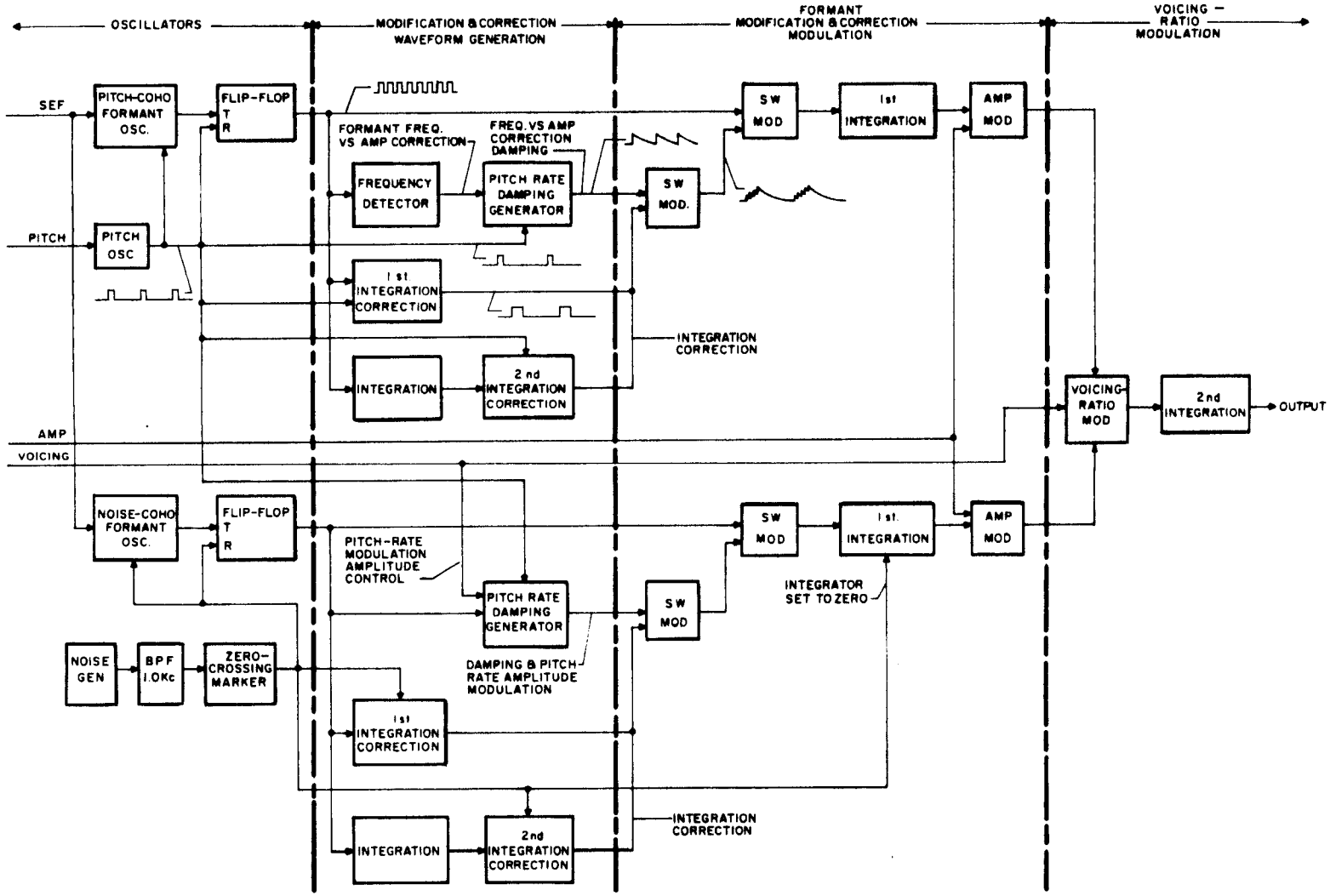


Figure 3-1. Early RADC Synthesizer, Block Diagram

3.2 RINGING-FILTER SYNTHESIZER

It was believed that another type of synthesizer could be built using voltage-tunable filters which would be excited with pitch-oscillator pulses in a manner analogous to the human vocal tract. Amplitude modulation would be obtained by switch modulation of the pitch-pulse amplitude. With this technique the output of the tunable filters would be a damped sine wave suitable for use directly as the synthesized speech. This ringing filter synthesizer is shown in block diagram form in Figure 3-2.

3.2.1 Three-Formant Synthesis

Ringing-filter synthesis is so much simpler than the cohered oscillator type that it was decided to build a three-formant synthesizer. It was expected that this would result in much more natural synthesized speech and greater intelligibility. The analyzer would be modified somewhat in that three-amplitude parameters extracted from the three-formant regions would be transmitted in place of a single-amplitude parameter. The modified analyzer is shown in block diagram form in Figure 3-3.

In the synthesizer it was planned to regenerate the signals necessary to control the synthesized formants from the SEF signal. This is possible because of the known relationship between the Single Equivalent Formant and the first, second, and third formants. This relationship, which was developed by Philco and is excerpted from Appendix B of this report, is shown in Figure 3-4.

It was decided that the circuits used to generate the formant frequency control signals would be simple analog diode shaping circuits. The reasoning behind this approach was that with three-formant sounds presented to the listener the location of each formant would be less critical than for a single formant. Figure 3-5 shows the first attempt at three-formant synthesis.

The three-formant synthesizer with analog deviation of the formant-frequency signals was an improvement over single-formant synthesis, and on a short RT word intelligibility test scored 72 percent. Although the 72 percent score was encouraging it was judged not adequate to produce 80 percent intelligibility on a PB word test. To increase intelligibility and quality a closer study of three-formant synthesis was begun. The three-formant synthesizer was used to generate 10 vowel phonemes, and by subjective listening tests the center and limit value phoneme parameters were obtained.

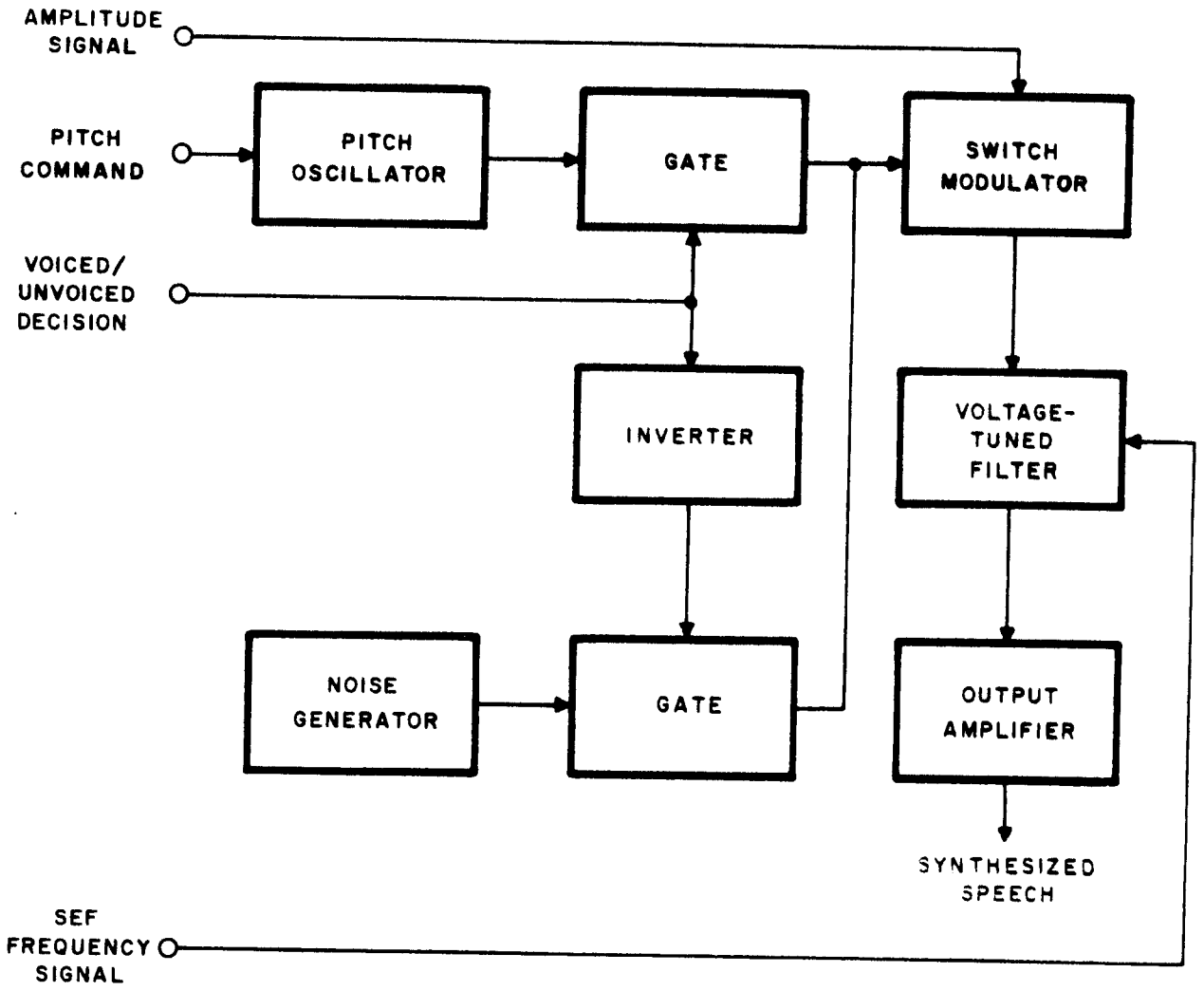


Figure 3-2. Single-Formant Ringing Filter Synthesizer

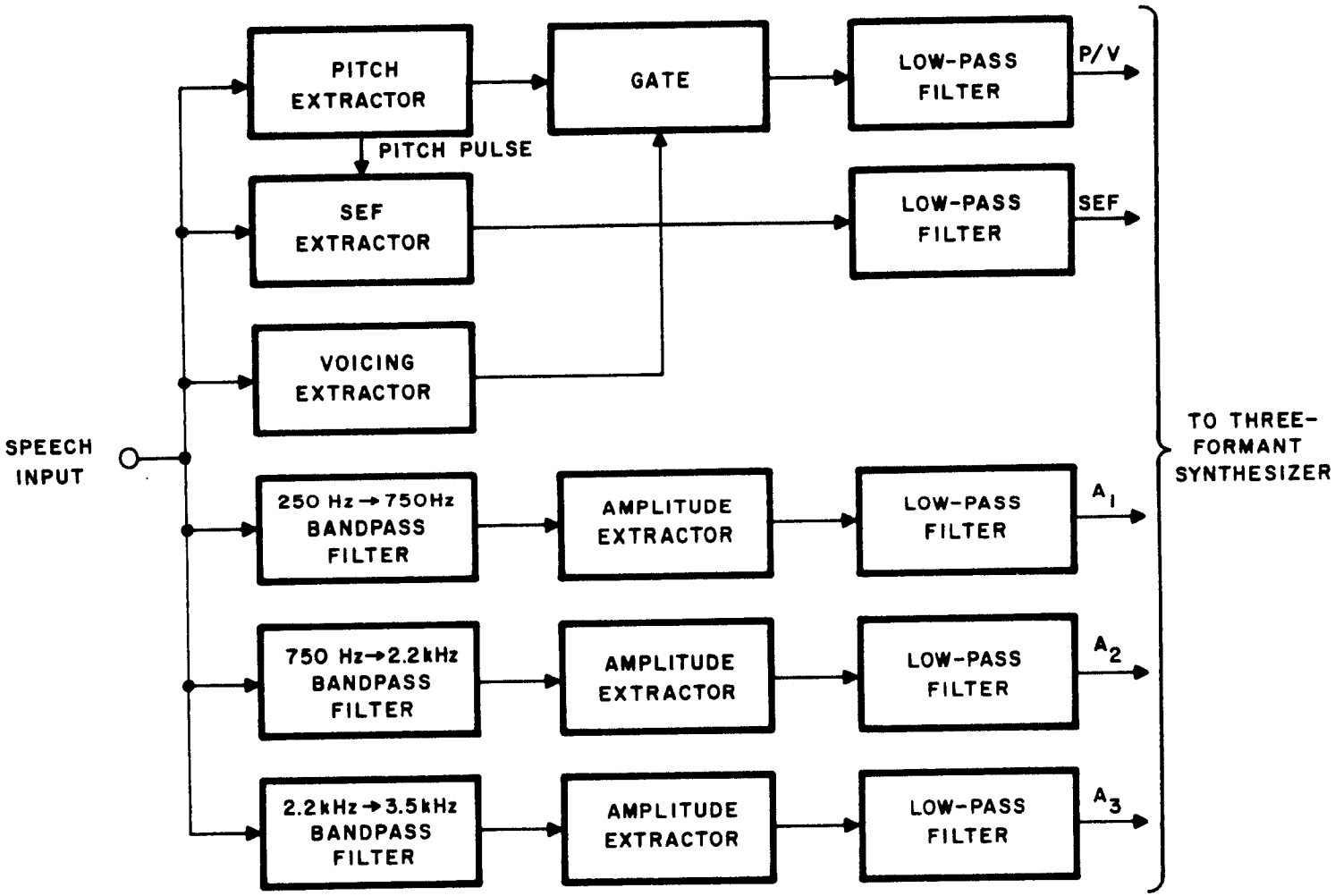


Figure 3-3. SEF Analyzer for Three-Formant Synthesis

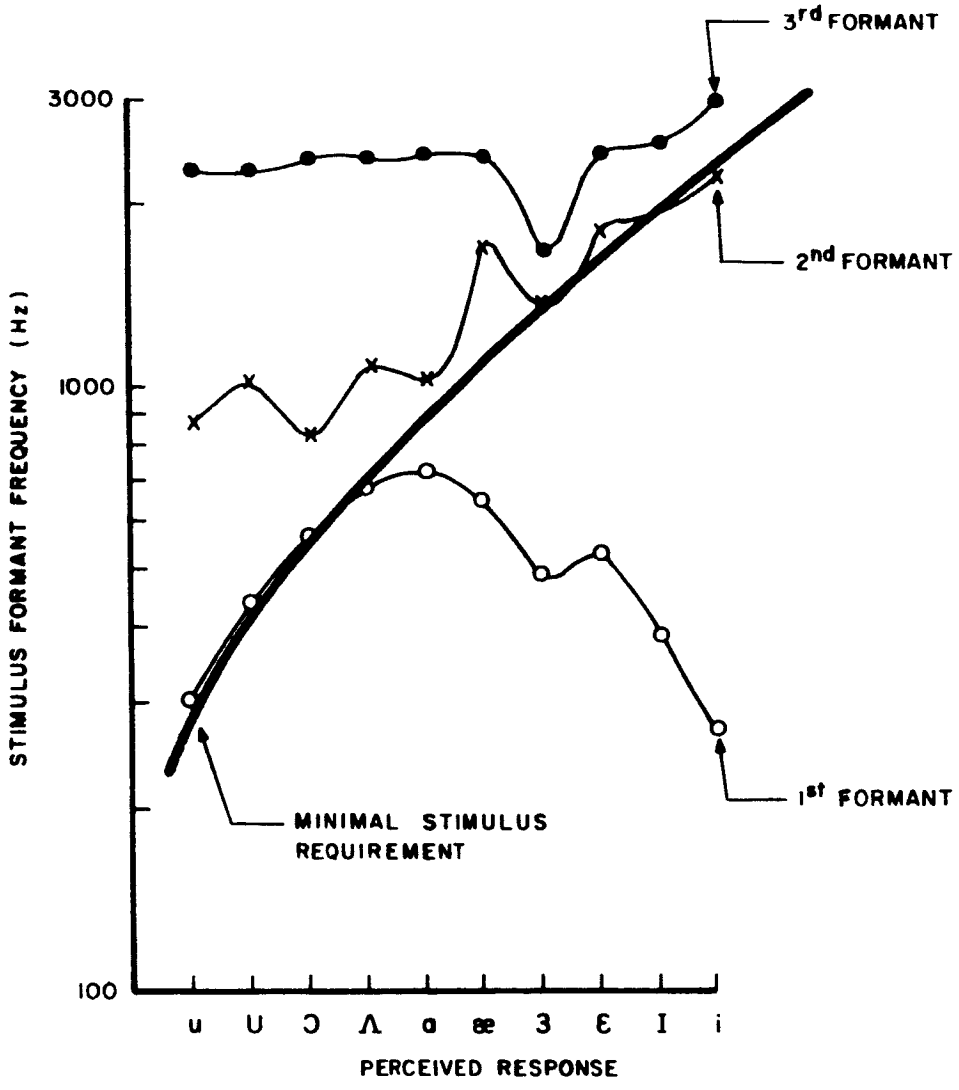


Figure 3-4. Correlation of Minimal Stimulus Requirement and Human Vowel Formant Locations

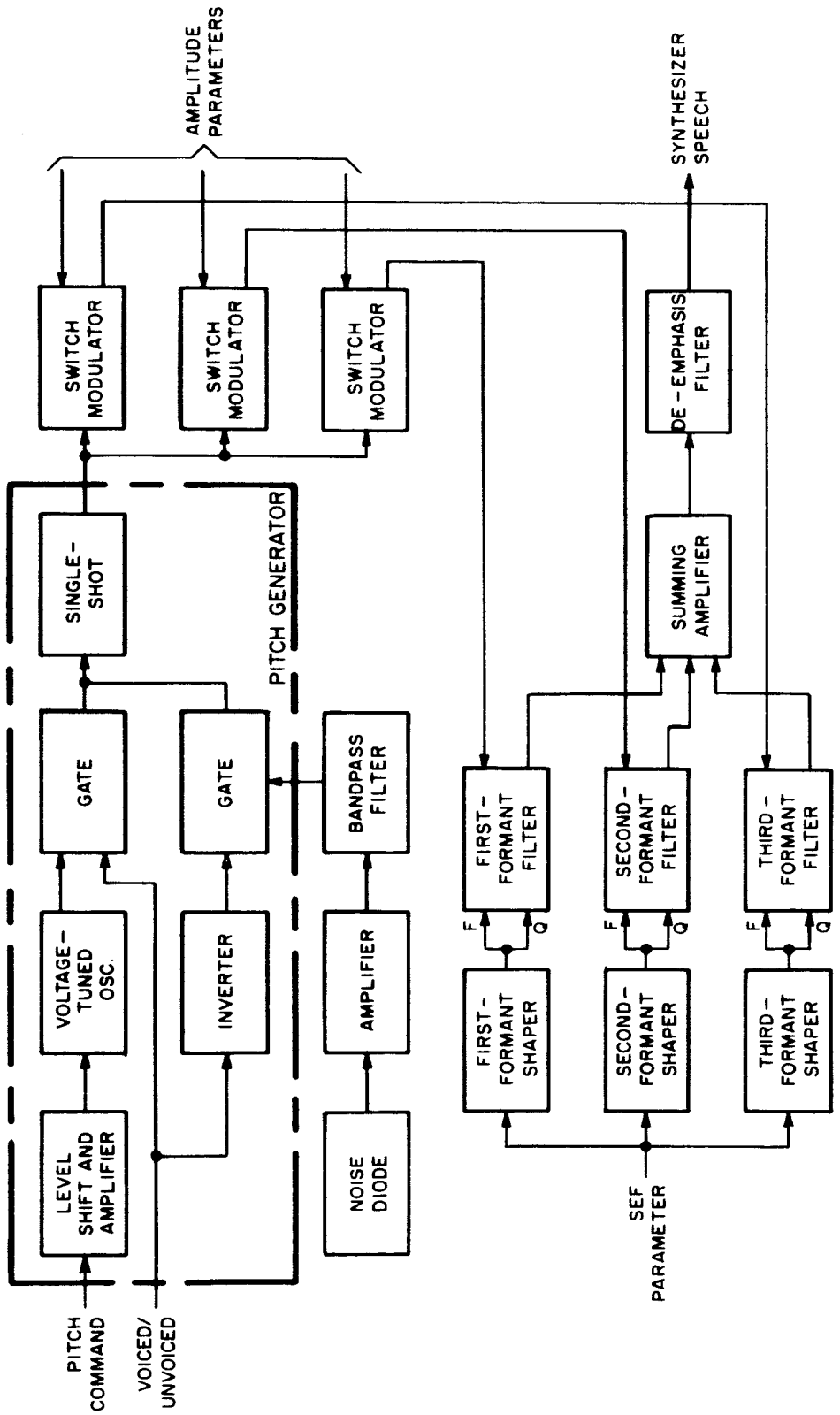


Figure 3-5. Three-Formant Synthesizer, Overall Block Diagram

Figure 3-6 shows the results of the study, with curves for the first, second, and third formants. The broken-line curves correspond to the average formant frequencies versus perceived phoneme, as presented by G. Peterson and H. Barney.[1] The solid curves represent the formant frequencies for the perceived phoneme, using the three-formant synthesizer. The vertical bars on the curves, aligned with the perceived phonemes at the bottom of the graph, represent the maximum change in formant frequency that can be tolerated and still allow perception of the phoneme. The results plotted in Figure 3-6 established the fact that the formant-frequency parameters must be more closely controlled than was previously believed.

3.2.2 Digital Formant Shaper

Because of the complex shape of the formant curves (Figure 3-6) it was decided to build a digital formant shaper in order to obtain the required accuracy. It was decided that the digital formant shaper should control the relative formant amplitudes as well as the frequencies. Basically the shaper consisted of 12 threshold detectors (one for each of the vowel phonemes plus S and SH) connected with diode logic so that only one threshold could be activated at any input voltage. The threshold detectors drove two diode-resistor matricies, the outputs of which were dc voltages proportional to the frequency and amplitude of the three formants for the detected phoneme. Figure 3-7 shows the digital formant shaper in simplified block diagram form, and its connection to the three-formant synthesizer.

Initially this setup was tested with synthetically produced analyzer waveforms, and the quality was quite promising. The vowels were particularly good. As expected, the consonants were not quite as good due to the thresholding by the digital detectors during transient sounds.

When tested with real analyzer inputs, however, the quality of the synthesized speech was poor. This problem was studied by means of sonographs and visicordings, and it was concluded that the major cause of poor quality was the unsteadiness of the SEF signal. For a sustained vowel phoneme the SEF analyzer output was not a perfectly steady dc signal, but contained a randomness. Furthermore, the level of the SEF signal was not very predictable. This resulted in rapid whole-vowel fluctuations in the synthesized speech, as shown in Figure 3-8.

---

1. G. Peterson and H. Barney, "Control Methods Used in the Study of the Vowels," JASA, Vol. 24, No. 2, pp. 175-184, March 1952.

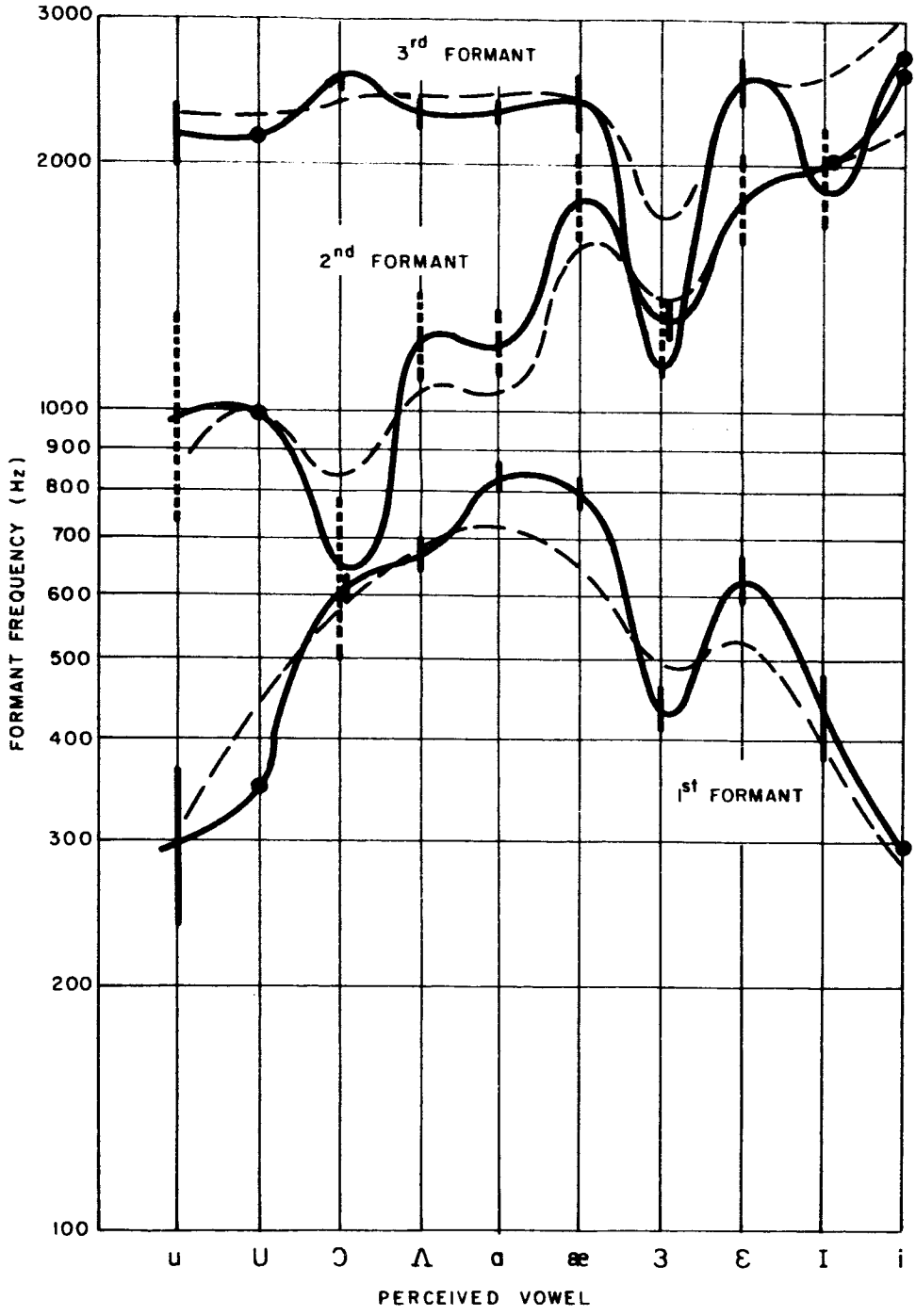


Figure 3-6. Formant Locations for 10 Vowel Phonemes

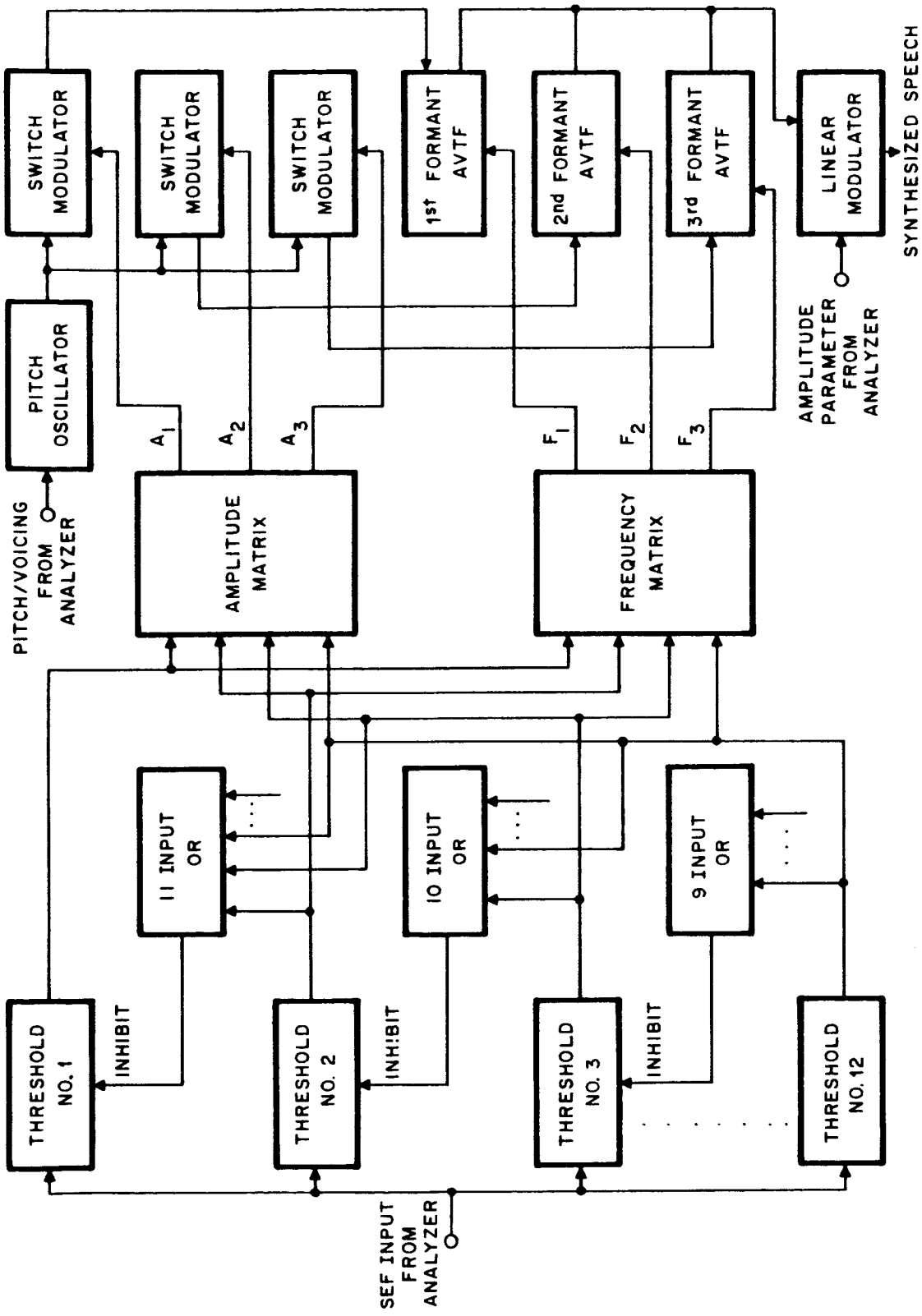


Figure 3-7. Three-Formant Synthesizer With Digital Formant Shaper

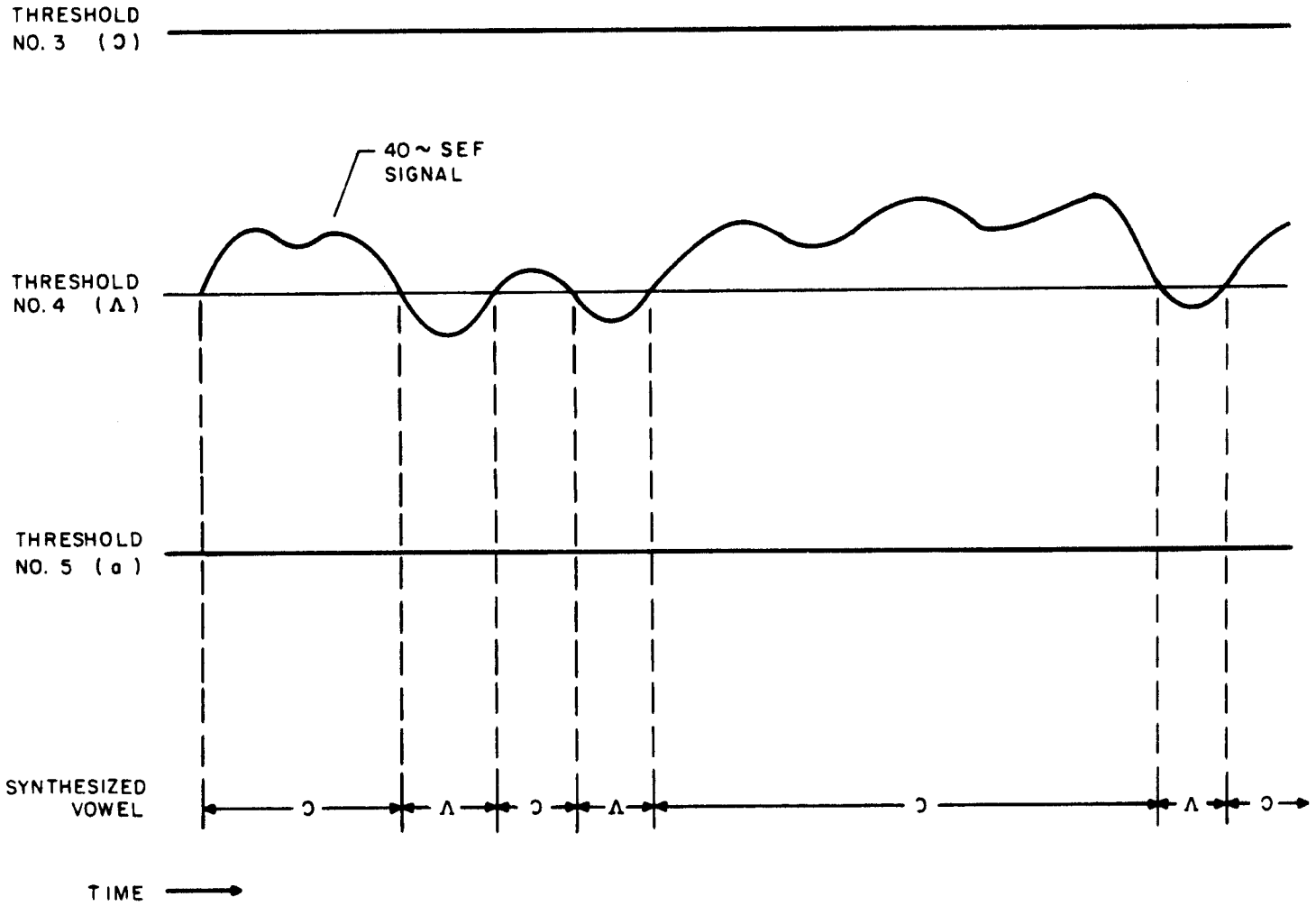


Figure 3-8. Digital Formant Shaper Synthesizer Output With Real-speech Input Parameters

Perceptually, this was quite poor. An attempt was made to smooth the synthesized output by RC filtering of matrix outputs (Figure 3-7) but this resulted with either slurring or clicking of the synthesizer output with no acceptable middle ground. At this point there were two alternatives: (1) increase the number of thresholds so that the matrix output signals would contain smaller transitions and therefore be smoother, or (2) attempt to obtain a smoother and better defined SEF signal from the analyzer. To increase the number of thresholds from 12 to 48 (4 per phoneme) would require a geometric increase in the number and complexity of the diode logic as well as in the output matricies. In addition, the threshold circuits would require extensive redesign since their accuracy would need a four-fold improvement. In view of these problems the second alternative was chosen: that of improving the quality of signal from the analyzer.

3.3 SINGLE EQUIVALENT FORMANT TRACKING FILTER ANALYZER (SEFTA)

A new method of SEF extraction, which was not pitch synchronous, was investigated. This new method utilized a tunable preselector filter, period detection, and period-to-dc conversion for the SEF output. Figure 3-9 shows a block diagram of the Single Equivalent Formant Tracking Analyzer (SEFTA). The speech input is first passed through an AVTF, then through another identical AVTF connected with a differential amplifier to form a high-pass filter. Both AVTF's are tuned to the same frequency, so that when both are applied to the differential amplifier, the overall response at the output of the differential amplifier is that of a bandpass filter. The bandpassed speech is then analyzed by a detector which converts period to voltage linearly. Since the AVTF's have a linear frequency versus control voltage, a shaping circuit is required to convert from period to frequency. This output is now of the proper shape and level to drive the AVTF's to the frequency detected by the zero-crossing detector and enhance its accuracy.

A SEFTA analyzer was constructed using the previously constructed pitch, voicing, and amplitude extractors, and tested with a single-formant ringing-filter synthesizer. The result was smooth and pleasing synthesis as long as the analyzed phonemes stayed in the Λ to i region, but when the analyzed phoneme moved into the w, U, ↄ region the analyzer tended to get trapped in this region and hesitate before following a glide into the Λ to i region. The perceptual result of this analyzer error can only be described as peculiar, and not of high intelligibility. The speech input to the SEFTA analyzer was then high-pass filtered, and the corner frequency gradually raised. When the high-pass setting reached 750 Hz the trapping tendency

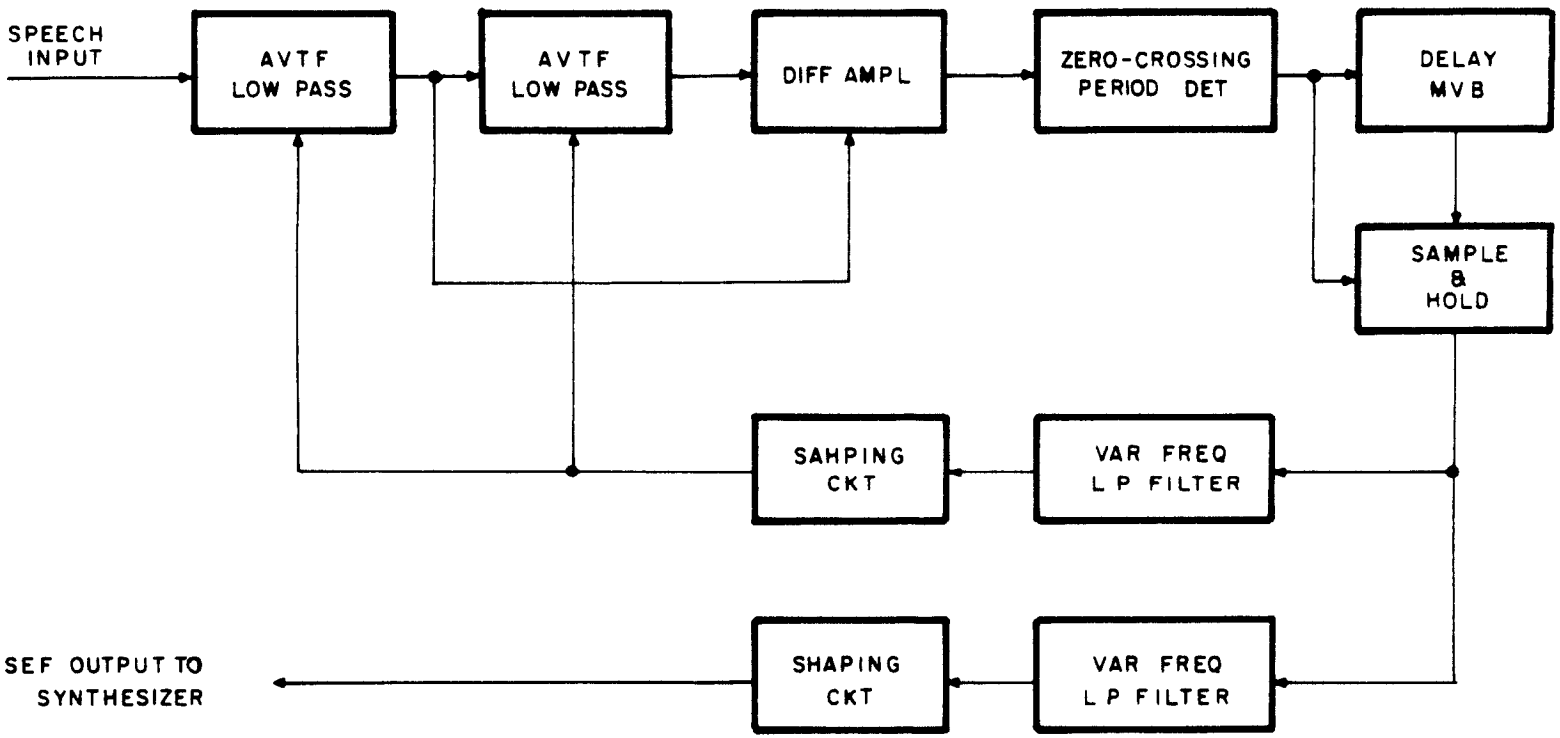


Figure 3-9. Single Equivalent Formant Tracking Analyzer (SEFTA)

was effectively eliminated, but the synthesized output now lacked accuracy in the back-vowel region. While in this configuration it was determined that the tracking filter could be replaced with a fixed-tuned bandpass filter designed to pass the second formant frequency region. The shaping circuit was retained since it produced the correct transfer function to control the synthesizer filters.

3.3.1 Two-Formant Analyzer

It was then decided to split the audio spectrum into two sections with a period analyzer and amplitude extractor for each section. Two fixed-filter analyzers with two amplitude extractors were built. The first analyzer was bandpass limited to the 200 Hz to 750 Hz region; the second analyzer was bandpass limited to the 750 Hz to 3.5 kHz region. This analyzer is shown in block diagram form in Figure 3-10.

This analyzer represents a departure from the Single Equivalent Formant (SEF) idea in that the extractor which operates in the 200 Hz to 750 Hz range is tracking or extracting the first speech formant. The second extractor retains some of the principle of SEF in that it extracts the second formant or third formant, whichever is dominant. Neither of these extractors operate on the first half-cycle of speech that is synchronous with the extracted pitch pulse, as did the original SEF extractor. They are not pitch synchronous, so that the system no longer requires the complex real-time pitch extractor which it now uses. Although time did not permit on this program, the pitch extractor could be considerably reduced in complexity by redesign to a frequency-domain extractor.

3.4 TWO-FORMANT SYNTHESIZER

The synthesizer used in the two-formant system uses the ringing-filter synthesis technique. It is quite similar to the human vocal tract in that pitch pulses (or noise pulses for an unvoiced sound) are amplitude modulated according to the formant amplitude signals A_1 and A_2 (Figure 3-11) and used to excite two tunable bandpass filters. The resonant frequencies of the filters are controlled by the extracted formant frequency signals from the analyzer. The tracking low-pass filters that follow the AVTF's in Figure 3-11 were found to be necessary to reduce high-frequency pitch harmonics that were passed by the AVTF's because of the method of excitation. This problem will be discussed in greater detail in the subsection on Active Voltage-Tuned Filters.

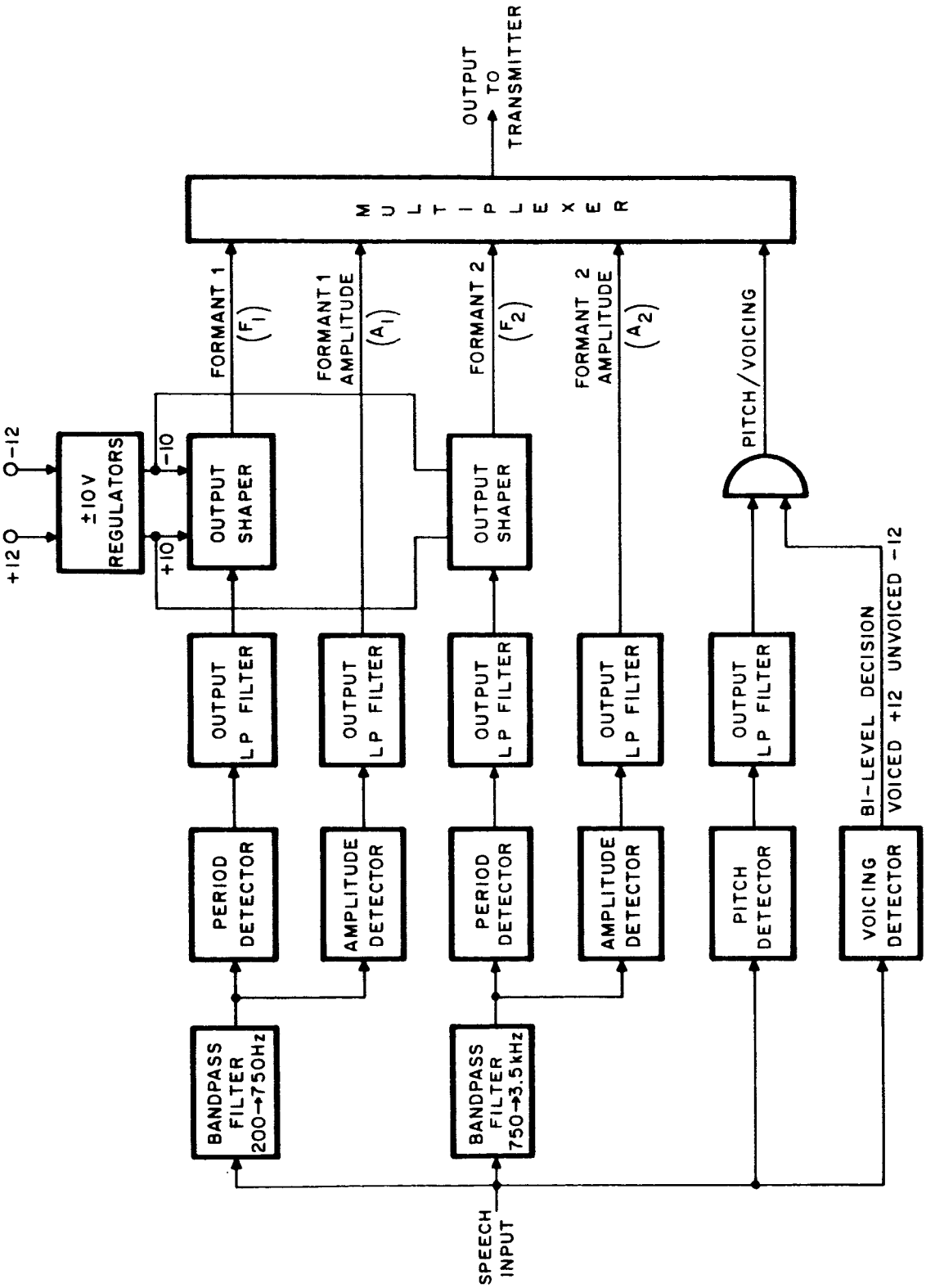


Figure 3-10. Two-Formant Analyzer, Block Diagram

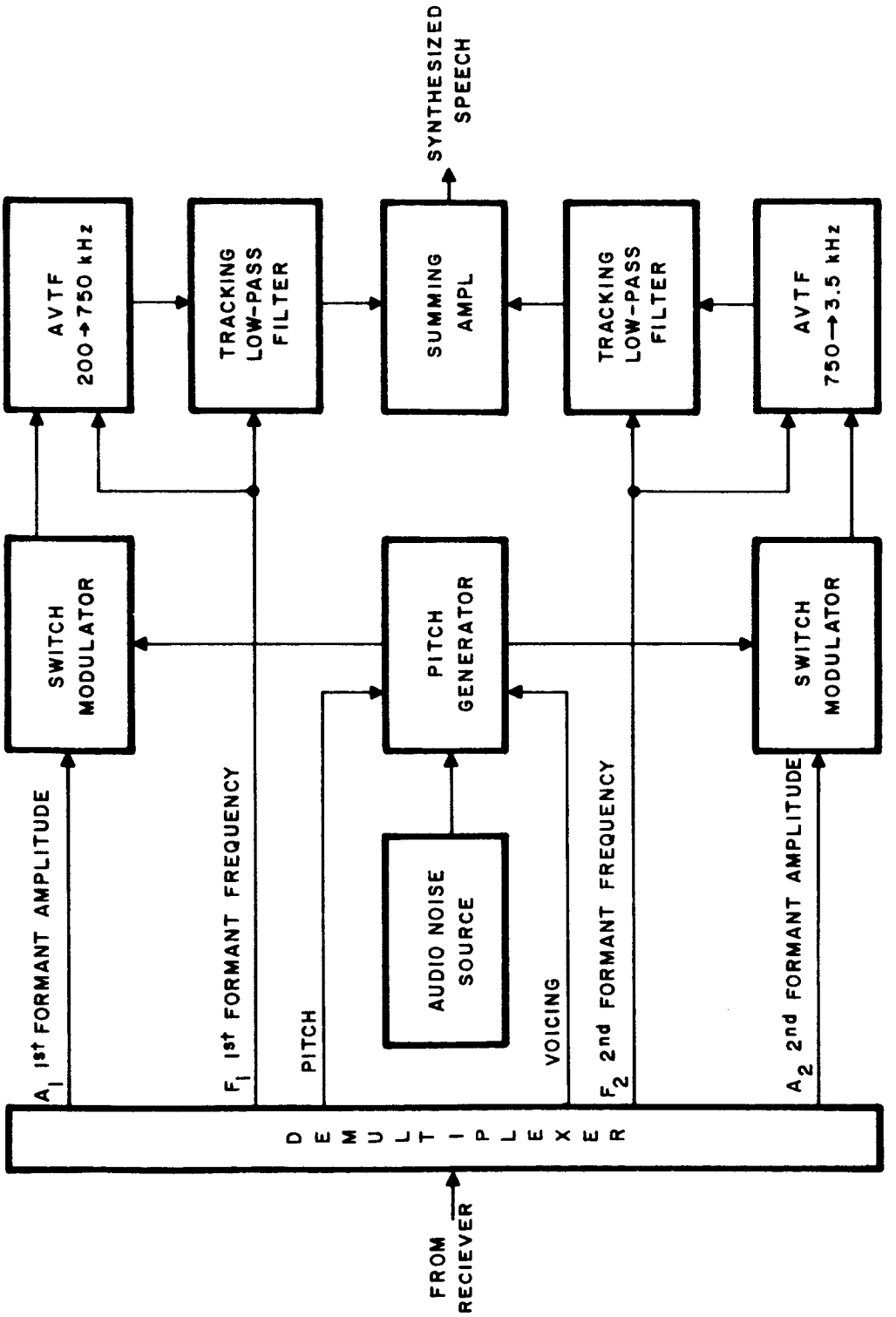


Figure 3-11. Two-Formant Synthesizer, Block Diagram

The first tests on the two-formant analyzer-synthesizer system made it apparent that this approach was far superior to anything we had previously tried. The quality of the synthesized speech was not unpleasant, and the intelligibility as measured by Rhyme Test (RT) method was 85 to 90 percent.

Now that the basic configuration of the system was chosen, a number of ideas were tried in an effort to improve system quality and intelligibility. These attempts, along with a short description and test results, are contained in a later subsection.

3.5 BANDWIDTH EXPERIMENTS

One of the more important experiments conducted concerned the optimum bandwidth of the parameter signals. The total analog bandwidth available was 160 Hz; allowing one 20 Hz channel for synchronization, this left 140 Hz for signal information. Four different arrangements of the signal bandwidths were tried, as follows:

Parameter	Parameter Bandwidth			
	System 1	System 2	System 3	System 4
First Formant Frequency	40 Hz	20 Hz	20 Hz	40 Hz
First Formant Amplitude	20	40	20	40
Second Formant Frequency	40	20	20	40
Second Formant Amplitude	20	40	20	40
Pitch/Voicing	20	20	20	40
	140	140	100	200

Systems 1, 3, and 4 had essentially equal scores on RT tests, with Systems 3 and 4 having the best quality. System 2 had the poorest quality and intelligibility due to the fact that the amplitude came on before the frequency got to the correct value. This is perceptually worse than the opposite situation, where the amplitude turn-on lags the frequency signal as in System 1. From the comparison between Systems 3 and 4 we learned

that the performance is not materially improved when the parameter bandwidth is increased beyond 100 Hz. We then have the happy situation of having 100 Hz of information (plus 20 Hz of sync) to transmit over a 160-Hz bandwidth link. Two multiplexing options are then possible: (1) sample at a 160-Hz rate and get better description of the 20-Hz parameters, or (2) sample at 120 Hz and get less channel crosstalk. It has been our experience that background noise or output from the synthesizer with no speech input to the analyzer is very annoying and detrimental to the intelligibility of the fricatives. This type of error would be a greater detriment to intelligibility than would be gained by better description at 160 Hz, so it was decided to sample at 120 Hz.

3.5.1 Detailed Description of Two-Formant Analyzer

In this subsection, each block of the two-formant analyzer (Figure 3-10) and the two-formant synthesizer (Figure 3-11) will be discussed in detail.

3.5.1.1 First-Formant Input Bandpass Filter

The input bandpass filters consist of a series of low-pass and high-pass active filters. In the first formant the first low-pass section is underdamped ($\rho = 0.25$) and the second low-pass section is overdamped ($\rho = 1$), and they are synchronous tuned, to give a high attenuation rate near the corner frequency. This is necessary to get good separation between the first and second formants where they are close together (see Figure 3-6). The high-pass sections are critically damped and synchronous tuned, which gives sufficient attenuation to reject the pitch fundamental frequency. The bandpass shape of the first formant filter is shown in Figure 3-12. The schematic diagram for the first- and second-formant filters is shown in Figure D-1 of Appendix D.

3.5.1.2 Second-Formant Input Bandpass Filter

The second-formant input filter uses an underdamped and overdamped pair of high-pass filters plus a critically damped high-pass section, all synchronous tuned (same corner frequency). This results in 35 db attenuation in the first octave (900 Hz to 450 Hz) which was found necessary to prohibit first-formant energy from causing errors in the second-formant extractor. The high-frequency corner in the second-formant bandpass is provided at 6.5 kHz by two synchronous-tuned, critically damped low-pass sections. It is necessary to provide the analyzer with information up to 6.5 kHz in order to get good differentiation between the fricative sounds.

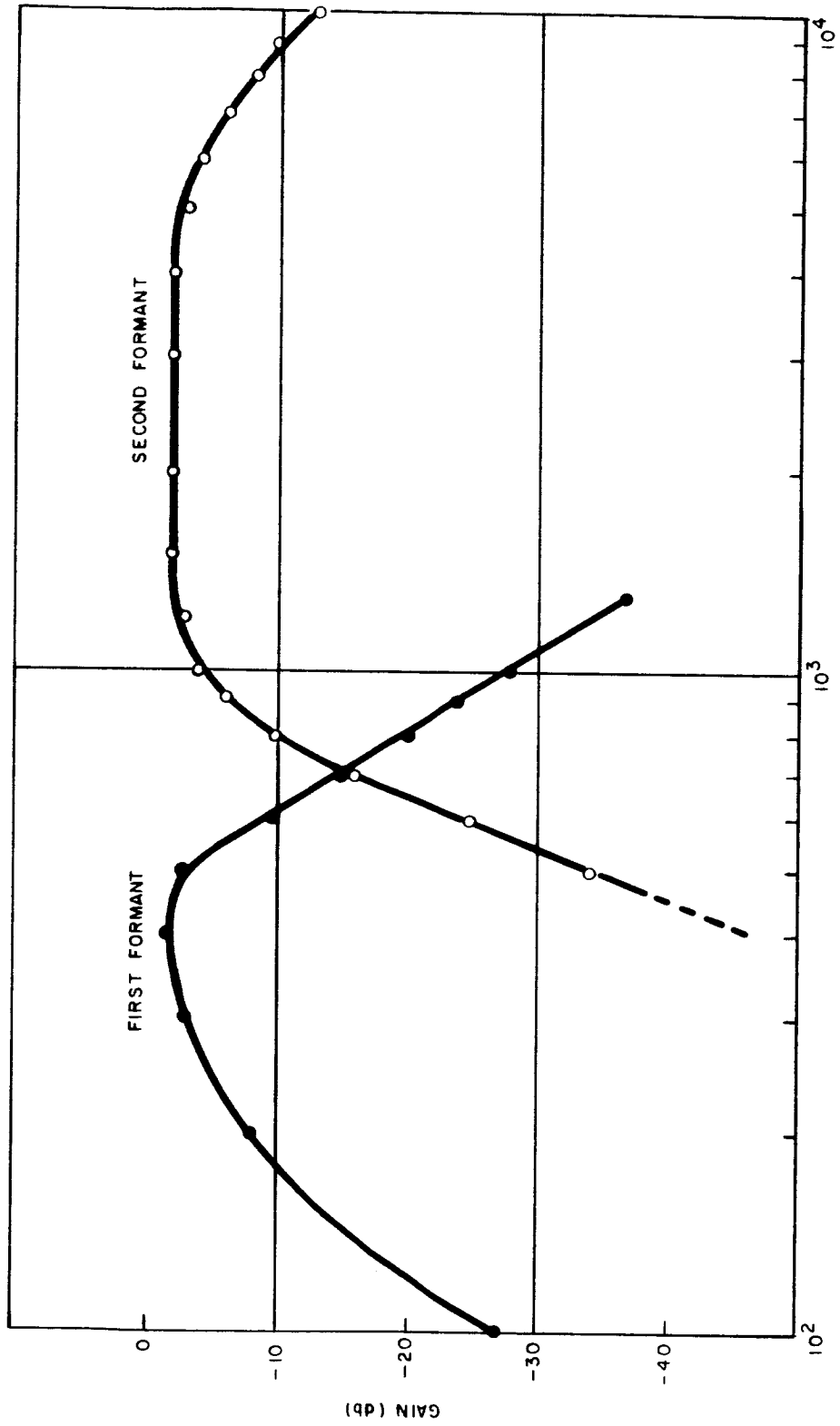


Figure 3-12. Formant Input Bandpass Filters

This is particularly true for the S and SH ($\int$) sounds, as can be seen from the sonograms of Figure 3-13. Figure 3-12 shows the second formant filter shape.

3.5.1.3 Period Detectors

The period detectors convert the zero-axis crossing rate to dc voltages, thereby producing a signal proportional to the period of the formant. Both first- and second-formant detectors are identical in concept, although they differ in actual circuitry by the transistor types and collector load values etc. A period detector is shown in block diagram form in Figure 3-14, and the timing diagram in Figure 3-15.

Referring to Figures 3-14 and 3-15, operation is as follows: at T_o, the formant waveform crosses the zero axis in the positive direction. The limiter output switches from negative to positive, and this output is differentiated and used to set MV_1 to the positive output state. The positive output from MV_1 acts through the NAND to open the charge switch and stop the increase in negative voltage on capacitor C. At the same time MV_1 closes the sample switch in the sample-and-hold circuit, thereby storing the voltage on C at T_o to T_1 in the sample-and-hold. At T_1 MV_1 returns to ground, opening the sample switch and triggering MV_2, which acts through the reset switch to discharge capacitor C. During the discharge time T_1 to T_2 the charge switch is held open by MV_2 acting through the NAND gate. At T_2 MV_2 returns to zero, opening the reset switch and closing the charge switch, thereby beginning the voltage run-up on capacitor C. The negative transition T_3 is prevented from reaching the logic, and the cycle does not repeat until T_4. In this way the period detector measures the full period. A switch is provided at the input to the limiter to select the phase (positive or negative zero-crossing) on which the system will operate, but it seems to have very little effect on the output.

3.5.1.4 Amplitude Detector

The amplitude detectors are simply formant amplitude envelope detectors. The output of the amplitude detectors consists of full-wave rectified ac. Filtering is done in the 20-Hz low-pass filters.

3.5.1.5 Output Low-Pass Filters

The output low-pass filters are active filters with 12 db/octave attenuation rate. They are designed for critical damping and are switch selectable at either 20 Hz or 40 Hz. All five filters are very similar; the

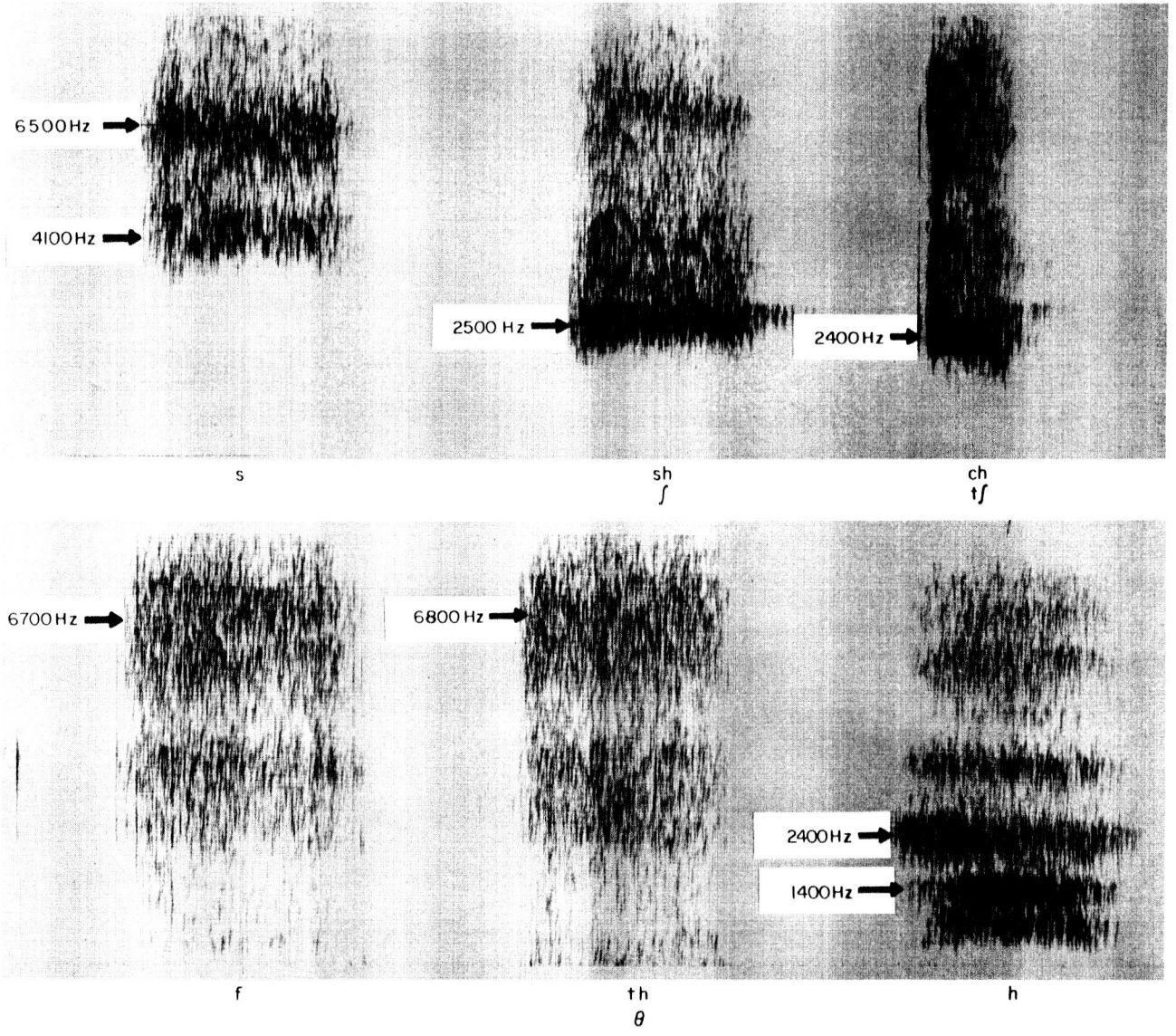


Figure 3-13. Sonographs of Unvoiced Fricatives

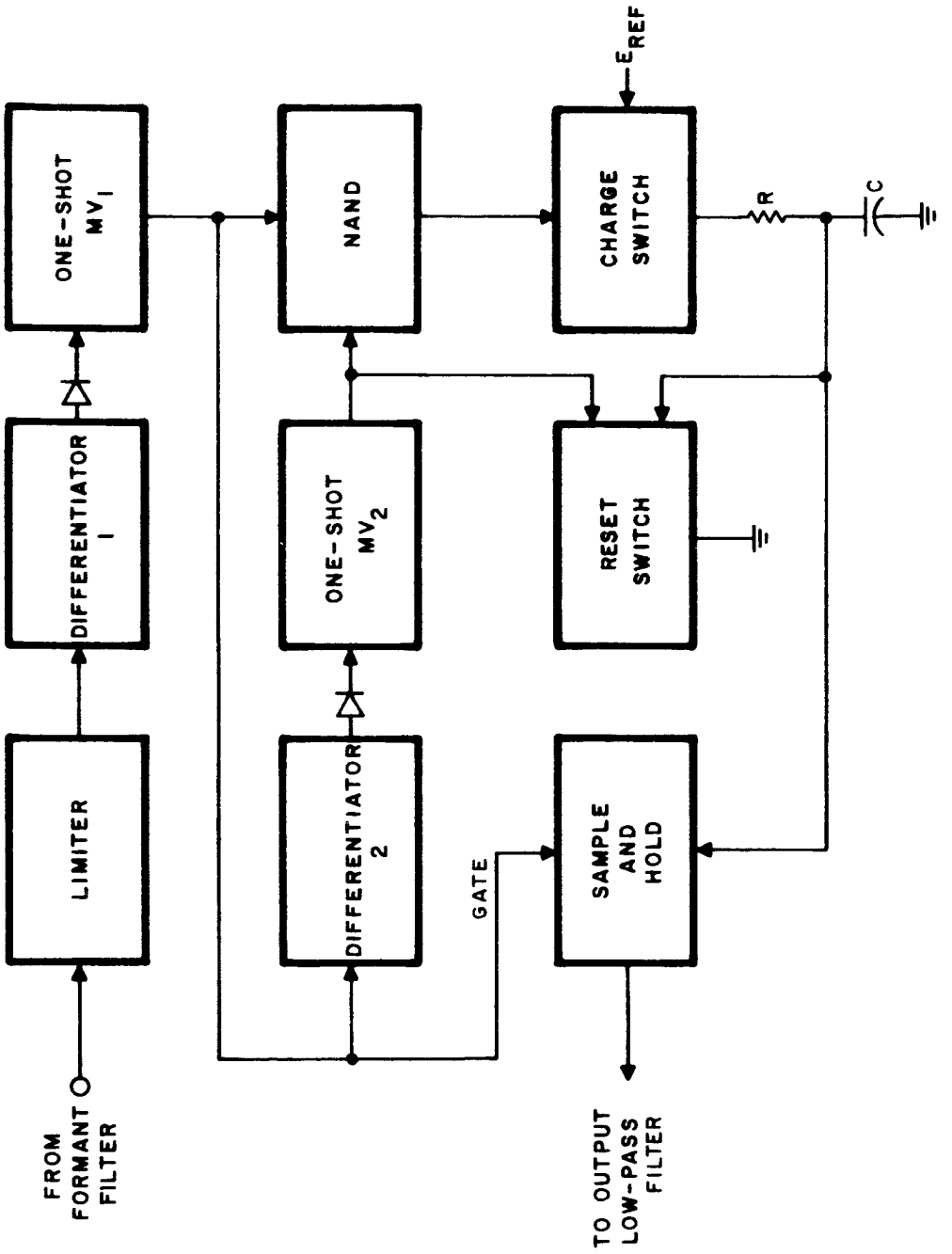


Figure 3-14. Period Detector Block Diagram

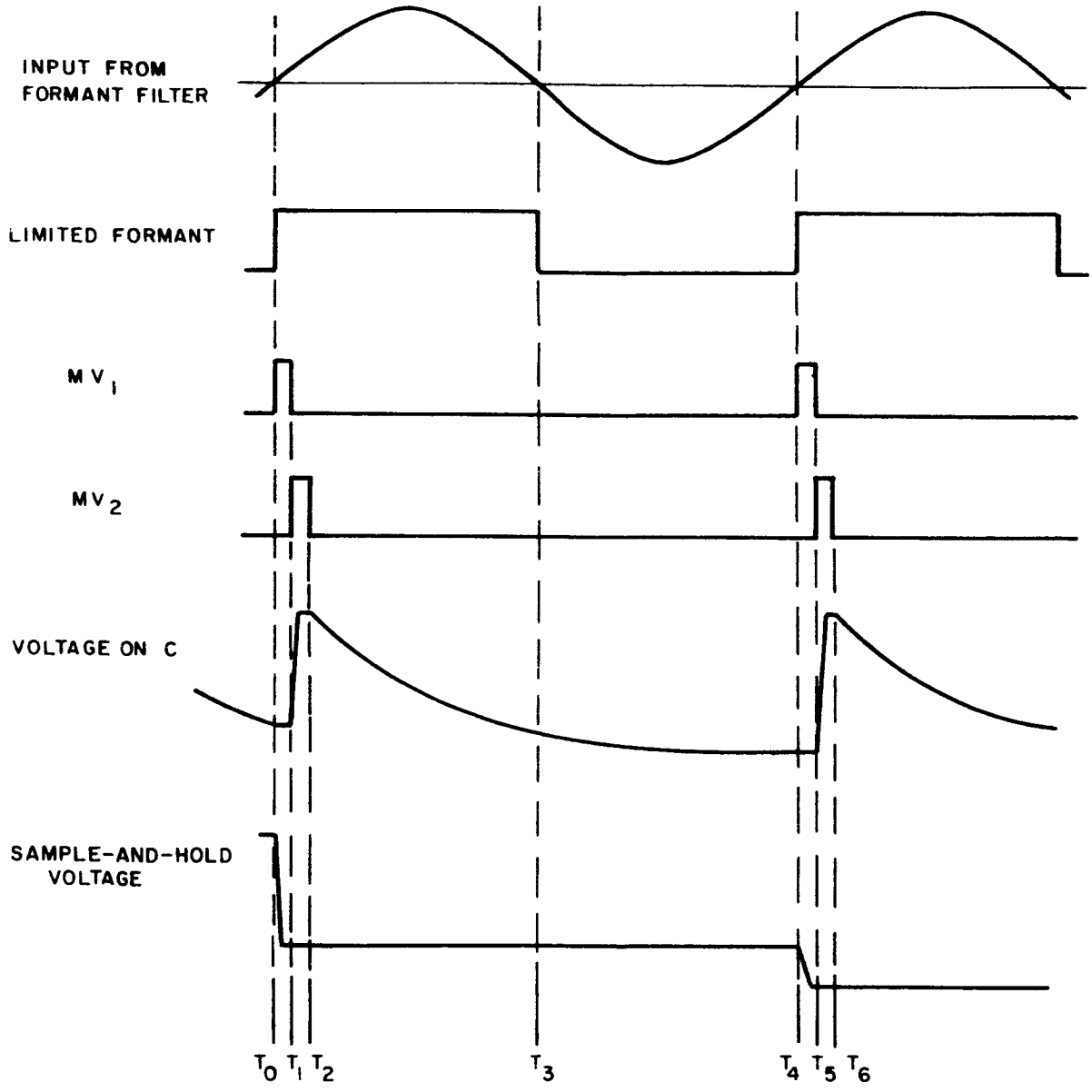


Figure 3-15. Period Detector Timing Diagram

only difference being the addition of another transistor in the amplitude-parameter filters, to provide higher input resistance in these filters.

3.5.1.6 Output Shapers

The output shapers convert the signal from the period detectors to a signal of the form $V_c = 1 - K_1 F_f$, where F_f is the formant frequency and K_1 is a positive number. This conversion is necessary because the synthesizer frequency control characteristic is of the form

$F_f = \frac{1 - V_c}{K_1}$, where V_c lies between zero and +1.

To arrive at the design of the output shapers the transfer function of the period detectors were calculated and plotted for various values of RC product (refer to Figure 3-14). A curve was selected that, when coupled with a simple diode break, would closely approximate a linear frequency-to-voltage conversion as required by the synthesizer.

Schematically, both output shapers are identical with different parts values to account for different frequency ranges. (See Figure D-9.)

The shaping network consists of R_1, R_2, R_3, D_1, and D_2. Phase inversion and isolation is provided by the differential amplifier for reasons of multiplexer compatibility. Diodes D_3 and D_4 provide limits for the formant frequency signal. Although these limits would never be exceeded during a speech passage, during a pause in speech, circuit noise can produce formant frequency signals which must not be applied to the synthesizer. Without these limit diodes the synthesizer generates disturbing spurious signals.

3.5.1.7 Pitch Detector

The pitch detector is a modification of a time-domain pitch extractor first described by Dolansky.[1] This pitch extractor has received considerable development at Philco and its characteristics are well understood. Its operation is based upon the principle given in the following paragraphs.

---

1. L. Dolansky, "An Instantaneous Pitch Period Indicator," JASA, Vol. 27, No. 1, pp. 67-72, January 1955.

Speech waveforms of voiced sounds exhibit a periodicity, as shown by Figure 3-16. A close examination of many such speech waveforms makes evident a point in time at which the vocal cavities are excited. The point (see Figure 3-16a) where the discontinuity occurs in the speech wave is an indication of the initiation of the vocal chord excitation function. The time-domain pitch extraction process is aimed at finding the point of discontinuity. The series of waveforms shown in Figure 3-16 depicts the basic process employed in the extracting of a time domain pitch pulse. Waveform a of Figure 3-16 shows a few cycles of a representative speech waveform. The points marked by arrows indicate the points of discontinuity where pitch pulses should be detected. Waveform b shows the resulting waveform after peak detection of waveform a. By differentiating waveform b, the initial transition of the original waveform is enhanced, as shown by waveform c. Waveforms d and e show the results of repeating the peak detection and differentiation process to further enhance the initial transition.

The preceding explanation of the pitch extractor is an over-simplication of the actual pitch extractor circuitry because nonlinear amplifiers, dual-time-constant peak detectors, and pre-emphasis networks are used to minimize pitch errors. The basic circuit elements of these operations are shown in Figure 3-17. The nonlinear or logging amplifier operates on the speech signal so that it occupies a relatively consistent dynamic range. As shown in Figure 3-18, the input speech waveform varies by 20 db between pitch pulses, whereas the signal output of the nonlinear amplifier is essentially unchanged from one pitch period to the next. The purpose of the logging amplifier is to preferentially amplify the low-level signal above the threshold of the peak detectors. The high-level signal is not amplified to the same degree, thus compressing the dynamic range of the signal. This method has been found effective in eliminating the problem of dynamic range of peak detectors.

The dual-time-constant peak detector eliminates errors occurring under the conditions of rapidly falling amplitude. This is shown in the block diagram by the two peak detectors, the emitter-follower, and the zener diode. The upper peak detector has a sufficiently long time constant to eliminate higher harmonics of the pitch from charging the peak-detector capacitor. Such a long time constant has an obvious disadvantage of decaying so slowly that rapid amplitude changes produce a loss in pitch pulses immediately after a fast drop in the amplitude of the speech wave.

The individual effect of the long time constant is shown in waveform a of Figure 3-19. The long time constant missed the second pitch pulse because of the previously mentioned rapid drop in amplitude of speech waveform. The short time constant peak detector, shown in b, intercepted the

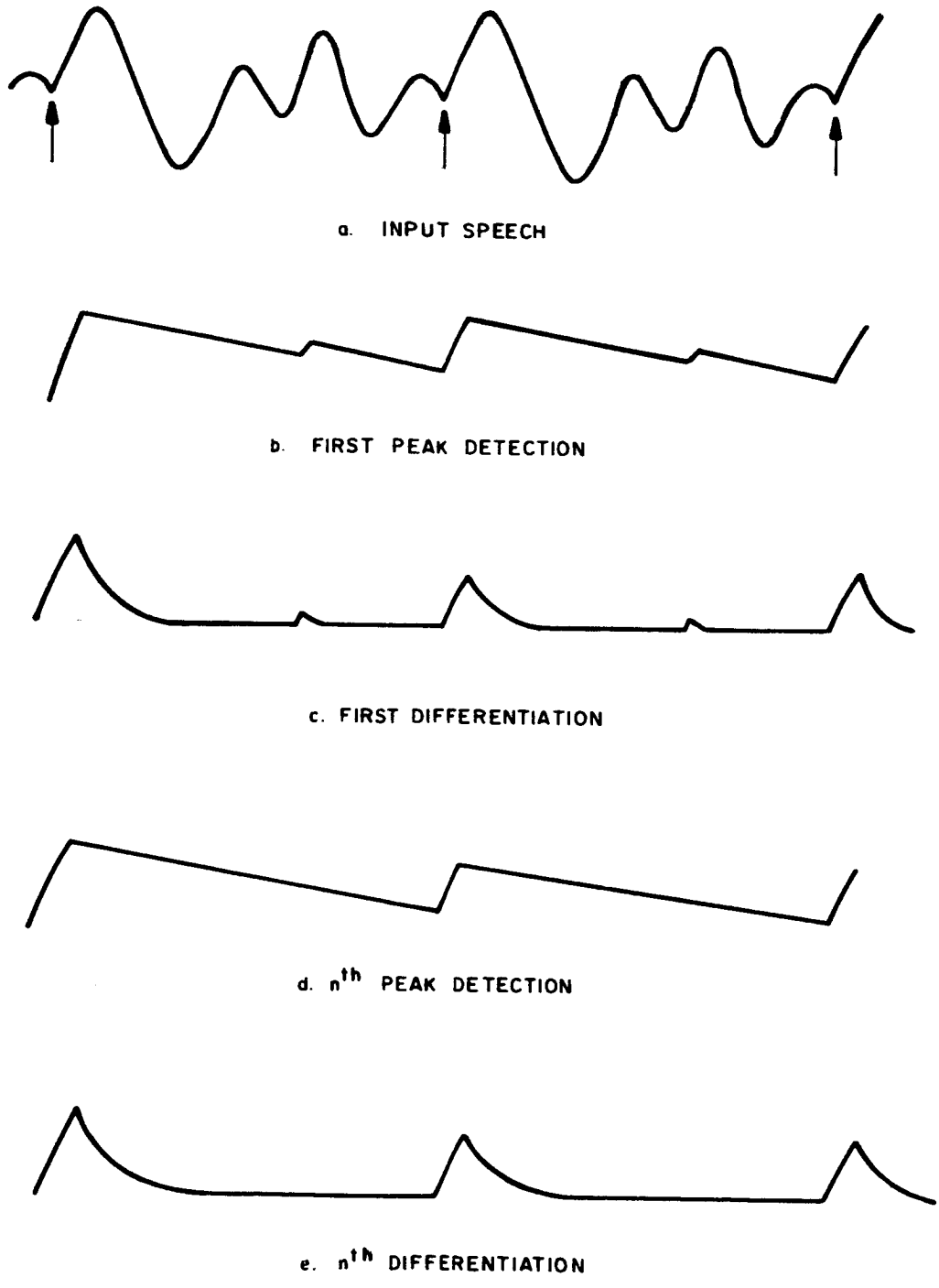


Figure 3-16. Basic Extraction-Process Waveforms

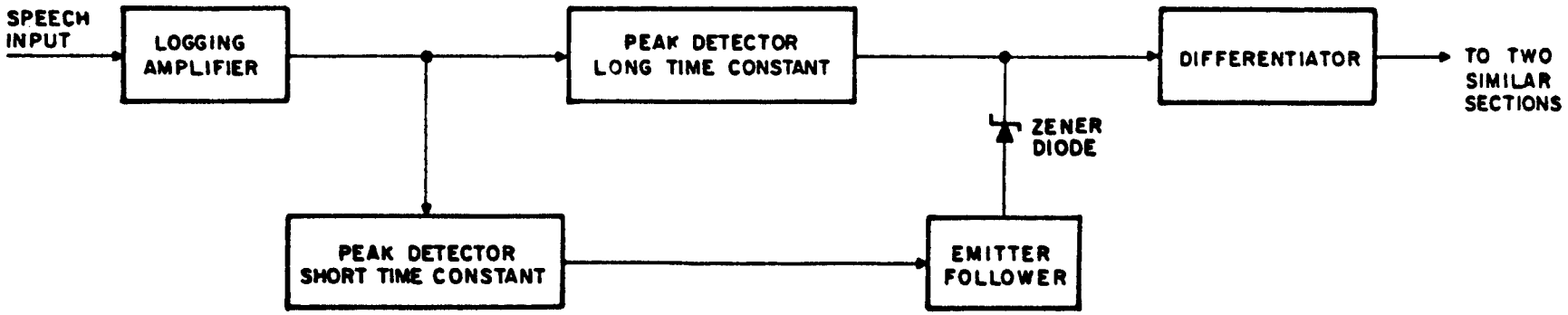


Figure 3-17. Pitch Pulse Extractor

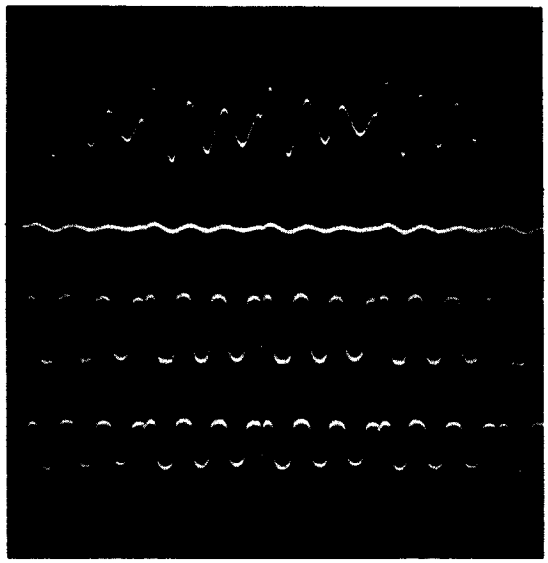


Figure 3-18. Nonlinear Amplifier Processing of Input Speech

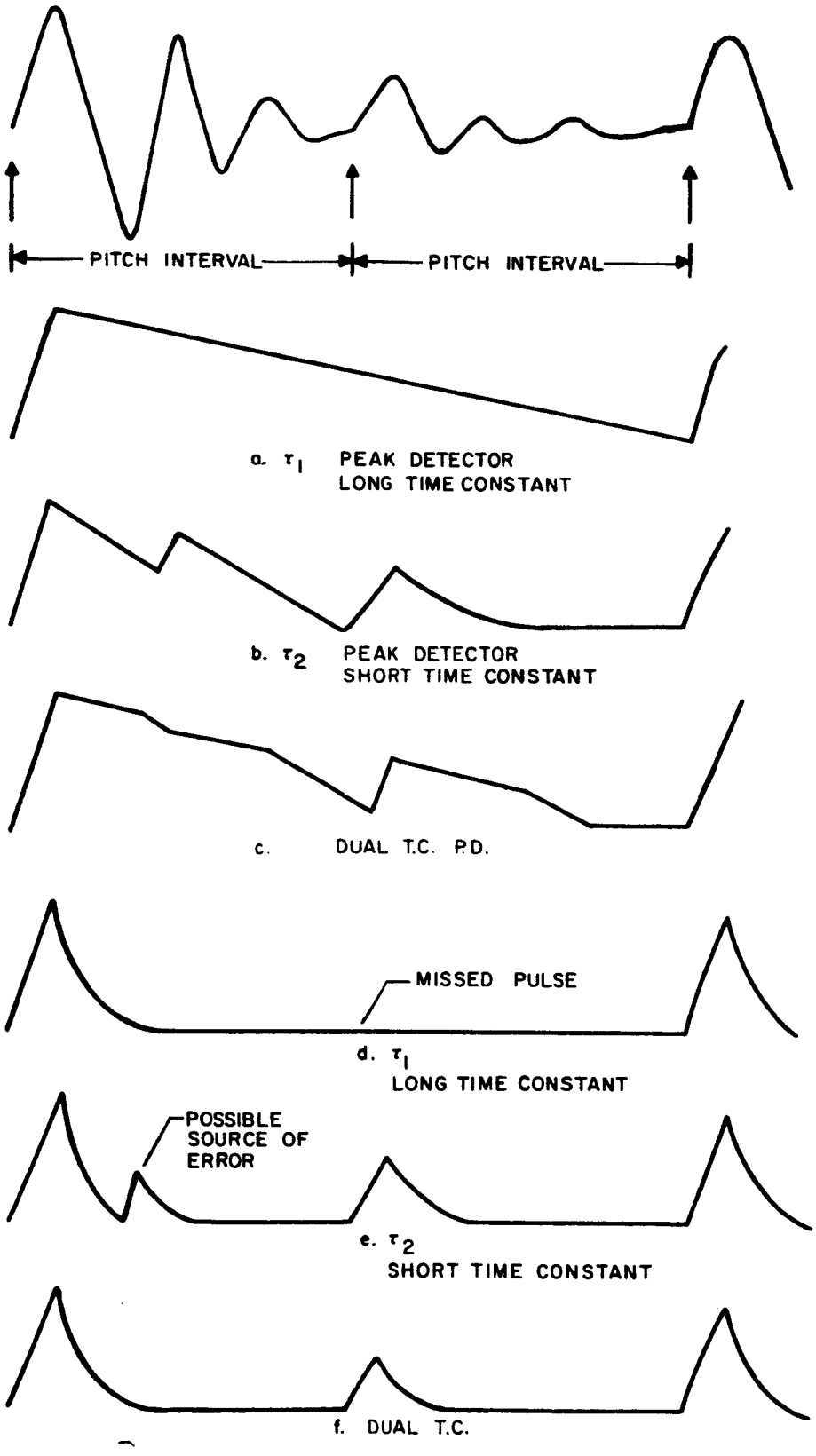


Figure 3-19. Comparison of Peak-Detector Waveforms

second pitch pulse. However, it also intercepted another undesired pulse in between the first and second pitch pulse.

The deficiencies of these two peak detectors are overcome by combining the two peak detectors with an emitter follower and zener diode. The zener diode conducts whenever the difference between the long- and short-time-constant waveforms falls below the zener breakdown voltage. Because of the emitter follower, the long time constant will now follow the short time constant waveform during the time when the zener diode is conducting.

The combined action of the dual-time-constant peak detection is shown in waveform c. Thus, when the amplitude changes rapidly, the combination will follow the amplitude with the required speed, but without a sacrifice in pitch second-harmonic rejection.

A block diagram of the complete pitch extractor is shown in Figure 3-20. As can be seen, the complete pitch extractor is a repetition of the nonlinear amplifier, peak detector, differentiator, and amplifier. The final pulse, after the last differentiation and amplification process, is digitalized and lengthened by a one-shot multibrator.

3.5.1.8 Voicing Detector

The first attempt to extract an accurate bilevel voicing decision was implemented using the existing multilevel voicing extractors. Its operation is based upon the determination of the average zero-crossing period for the speech wave. In this process, speech was infinitely peak clipped and the zero-crossings used to drive a ramp generator. The ramp generator was then peak detected and low-pass filtered. Thus, an average of the peak values of the ramp generator was obtained. Therefore, the output was a measure of the frequency differences of the zero-crossing rates of the voiced and unvoiced excitation. The filtered output was then clipped at the point where the highest voiced sound—and therefore the highest voiced zero-crossing rate— was found. This break occurs at the phoneme i at a zero-crossing rate of about 2200 hertz. Above this zero-crossing rate was designated unvoiced, and below was voiced.

Two problems with this approach made it necessary to modify the extractor slightly. The circuit failed when low frequencies were generated by wind blast into the microphone. To prevent this, speaking directly into the microphone was avoided. The major problem was the discrimination of i as a voiced vowel. Since i lies near the zero-crossing cut-off point,

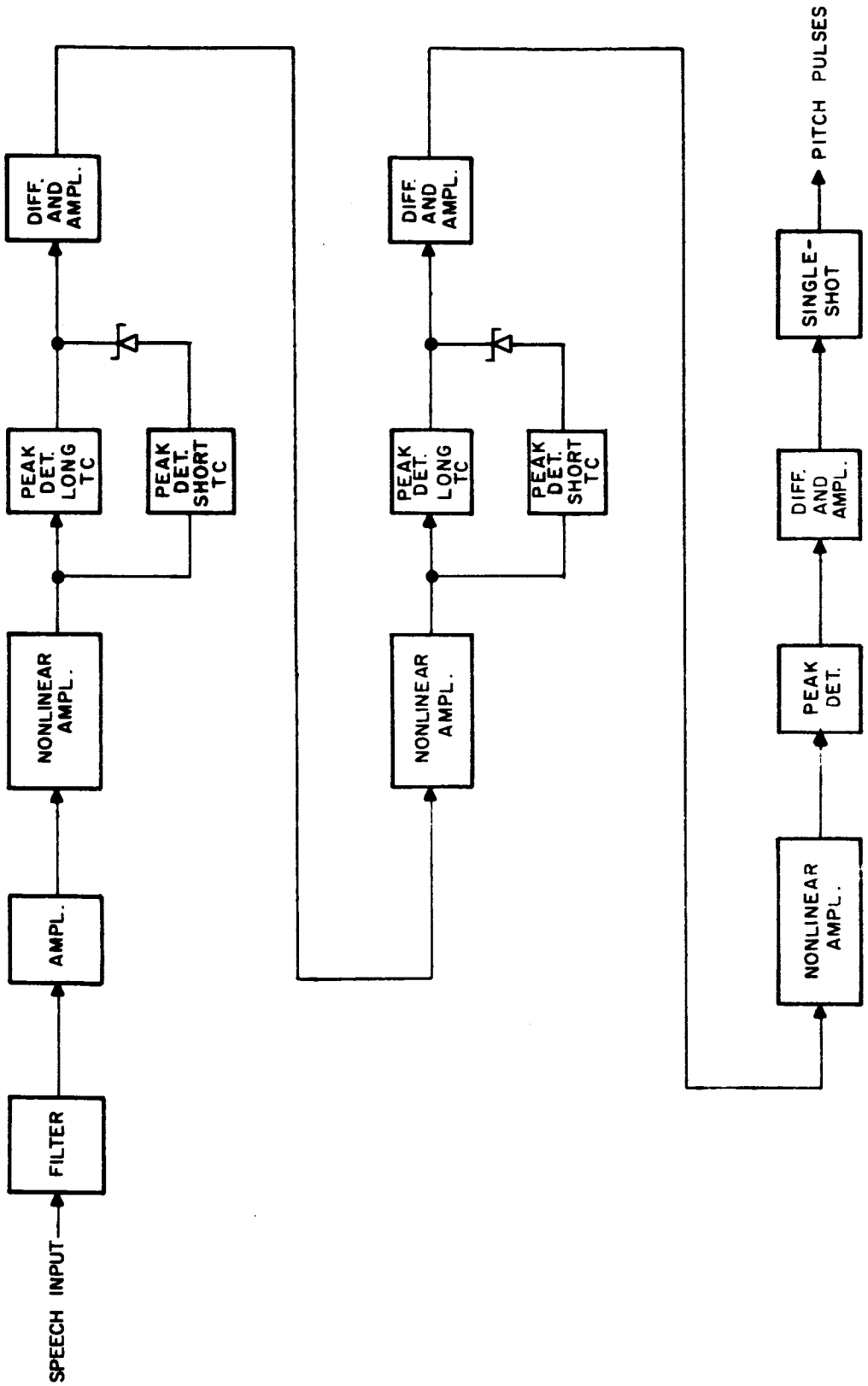


Figure 3-20. Pitch Extractor, Block Diagram

separating voiced and unvoiced energy, unvoiced errors were often made. To improve the performance of the voicing detection a second channel was used to detect the presence of low-frequency information in the input speech. By using logged speech, which made the detection practically insensitive to amplitude of the speech, and low-pass filtering the result, a ratio of low-frequency energy could be obtained. After appropriate rectification and filtering the low-frequency information was weighted with the multilevel voicing output and the result thresholded. The voicing extractor block diagram is shown in Figure 3-21. This proved to give very good vowel discrimination and made no unvoiced errors for higher formant frequency phonemes. The circuit, however, sometimes failed to go into an unvoiced state for plosives and unvoiced fricatives such as θ or f. Total performance of the voicing circuit was judged to be, while not perfect, adequate for synthesis purposes.

Other implementations to affect a voicing decision seem quite promising. Such a method is one that detects the ratio of harmonic to nonharmonic energy by measuring the periodicity of zero-crossing. However, further development would be needed to measure its full potential, and such development was deemed to be beyond the scope of the present program. This voicing detector is discussed more fully in a later subsection.

3.5.2 Detailed Description of Two-Formant Synthesizer

3.5.2.1 Pitch Generator

The pitch generator supplies constant-amplitude rectangular pulses to the switch modulators. For a voiced sound the pitch pulses occur at a periodic rate determined by the pitch command signal. When the voicing command calls for an unvoiced sound, pitch pulses occur at a random rate determined by the noise source. Separate miltivibrators are used for voiced and unvoiced sounds, to allow the relative energy between voiced and unvoiced sounds to be balanced by varying the multivibrator pulse widths. Figure 3-22 shows a block diagram of the pitch generator.

3.5.2.2 Switch Modulators

The switch modulators are simply low offset switching transistors followed by emitter follower buffer amplifiers.

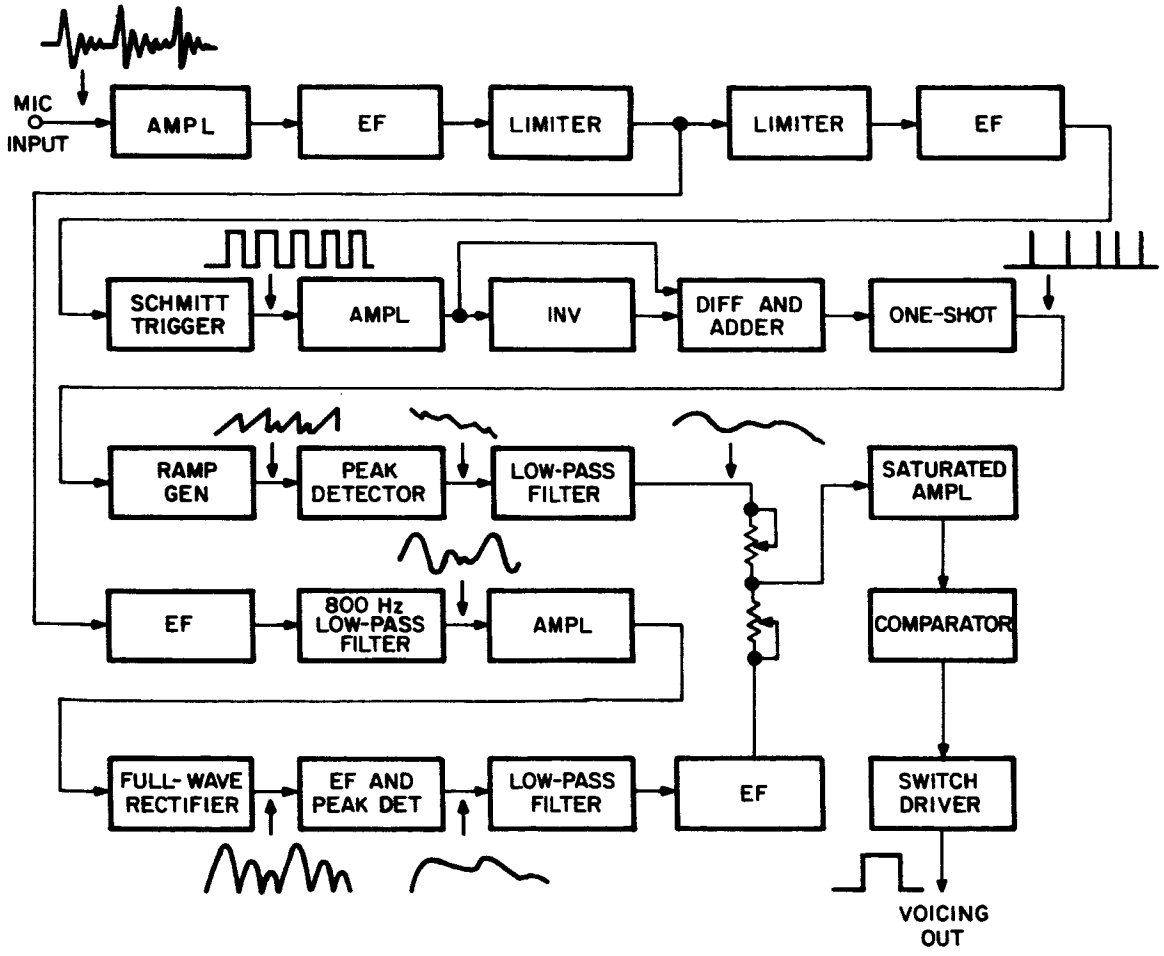


Figure 3-21. Voicing Extractor, Block Diagram

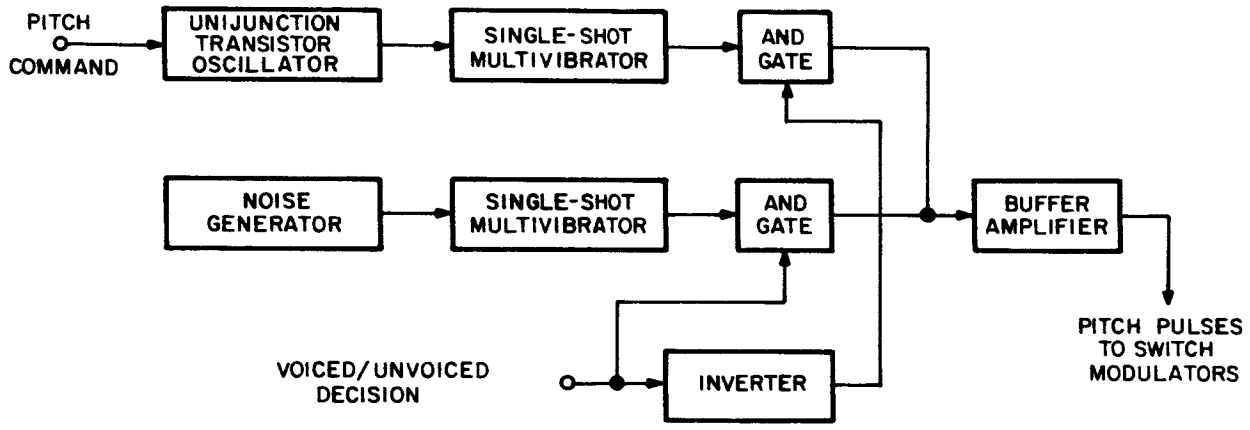


Figure 3-22. Pitch Generator, Block Diagram

3.5.2.3 Active Voltage Tuned Filters (AVTF)

The AVTF's designed for this program and used in the equipment are described fully in Appendix C. Originally, series-tuned analog filters were used but they proved to be too noisy. This problem was investigated and found to be due to the fact that amplifier noise is multiplied by filter Q when a series-resonant filter is used with an amplifier. A new type of AVTF was designed, which is the parallel resonant analog of a single-tuned LC bandpass filter. This filter is 20 ($\log_{10} Q$) db quieter than the series-resonant filter of the same Q.

To simplify problems associated with the signal driver circuitry, a modification of the bandpass AVTF shown in Figure C-6 of Appendix C was used in the equipment. This modification consisted of feeding the pitch pulses into the source of transistor Q5 from a low-impedance source. While the filter retains the lower noise property of the bandpass filter, from the signal input at Q5 source to the output it appears as a peaked low-pass filter. At the time this simplifying modification was made it was assumed that pitch harmonics outside the peaked response region of the AVTF would contribute negligible energy to the synthesized output speech. At present there is reason to doubt the validity of this assumption. It is suggested that further investigations be made into the characteristics of the formant resonators in the human vocal tract and into the characteristics of the AVTF. If the AVTF can be made to be more similar to the human formant resonators in impulse response, a considerable improvement in synthesized speech quality may be obtained.

3.5.2.4 Tracking Low-Pass Filters

These are critically damped low-pass AVTF's which further attenuate pitch harmonics above the formant resonanace. They are further described in Appendix C.

3.5.2.5 Summing Amplifier

The summing amplifier adds the two formant signals and increases their amplitude to about one volt peak-to-peak.

3.6 ANALOG BREADBOARD INVESTIGATIONS

In attempting to improve quality and intelligibility, a number of ideas suggested by speech literature and by workers at Philco were tried. Some of the ideas gave excellent results and are incorporated into the equipment. Others yielded results which were encouraging but they could not be pursued

to completion without interferring with the contract schedule. A list of some of the ideas that were tried and a brief description of the results are given here.

a. Pre-emphasis of high frequencies to obtain better amplitude description of low-level, high-frequency information. This idea has historical basis in channel vocoders. In this equipment it was not perceptually significant.

b. Pitch pulses sloped similar to the human glottal pulse as described by Flanagan.[1] The hoped-for improvement in quality did not occur probably because the formant AVTF does not have the same transfer function as the human formant resonant cavities.

c. Parallel-resonant analog AVTF's as opposed to series-resonant analog AVTF's for the formant filters to reduce electrical noise. This idea worked well and is incorporated in the equipment. Both the series-resonant and parallel-resonant analog AVTF's represent new developments, and invention disclosures No. 25-351 and No. 26-154 have been filed with Philco in favor of NASA.

d. Tracking low-pass filters after the formant AVTF's to further attenuate pitch harmonics above the formants. This improved quality considerably and is incorporated in the equipment.

e. Nasal antiresonance. This was suggested by a report "Formant Tracking Vocoder System."[2] Introduction of a nasal antiresonance was claimed to improve quality. Why it did not in this equipment is not understood, and perhaps the application of nasal antiresonance should receive further study.

---

1. Flanagan, J. L., "Speech Analysis, Synthesis, and Perception." Academic Press, New York, N. Y., 1965, pp. 191-202.

2. "Formant Tracking Vocoder System," U.S. Army Electronics Command final report on Contract No. DA 36-039-AMC-00006(E), August 14, 1962, p. 48.

f. Fixed-frequency third formant with variable amplitude. A 20-Hz bandwidth amplitude parameter extracted from the third formant region was used to synthesize a fixed-frequency third formant at 3.3 kHz. This made some improvement in the vowels I and i but caused degradation of the fricative sounds, and so was discarded.

g. Pitch periodic formant shift: from sonograms it was noticed that the second formant frequency shifted upward from its position immediately after the pitch pulse to a slightly higher frequency just before the next pitch pulse. Using a signal from the pitch generator this effect was implemented in the synthesizer. Perceptually it had no effect, and it was concluded that the human being probably averages the formant frequency over the short interpitch pulse period.

h. Formant Q controlled by formant frequency. This was implemented using formant Q versus frequency data from Peterson and Barney.[1] This seemed to be of moderate importance perceptually but the method of implementation gave so much trouble that the idea was discarded as not practical at this time.

It should be noted that the above eight ideas were tested with the equipment in various stages of perfection. It is probable that some of the more gross imperfections present in the equipment at the time of test obscured whatever improvement these tests may have produced. For this reason, in future work these ideas should not be disregarded, but the more promising ones tried again after some of the more important problems are solved.

---

1. G. Peterson and H. Barney, "Control Methods Used in the Study of the Vowels," JASA, Vol. 24, No. 2, pp. 175-184, March 1952.

3.7 MULTIPLEXING STUDIES

The time-multiplexed output of the narrow-band speech compression system will appear as a PAM (Pulse-Amplitude Modulated) signal. The total information bandwidth, which is 140 Hz including sync, is almost identically equal to the allowable bandwidth (160 Hz) of the transmission link. As a consequence, the problem of data transfer required considerable though which resulted in a study of techniques for the suppression of cross-talk in the demultiplexer. From this study two techniques, the Return To Complement (RTC) system and the Finite Memory Filter (FMF) were breadboarded and evaluated. Ultimately the FMF was chosen as the more desirable technique to use. Both techniques are explained in the following paragraphs, along with the reasons for choosing the FMF system.

3.7.1 Return to Complement (RTC) System

In essence, the RTC approach realizes a much faster effective decay of the trailing edge of a pulse by shifting the final discharge level in an appropriate fashion. This is illustrated in Figure 3-23, for a band-limited PAM channel.

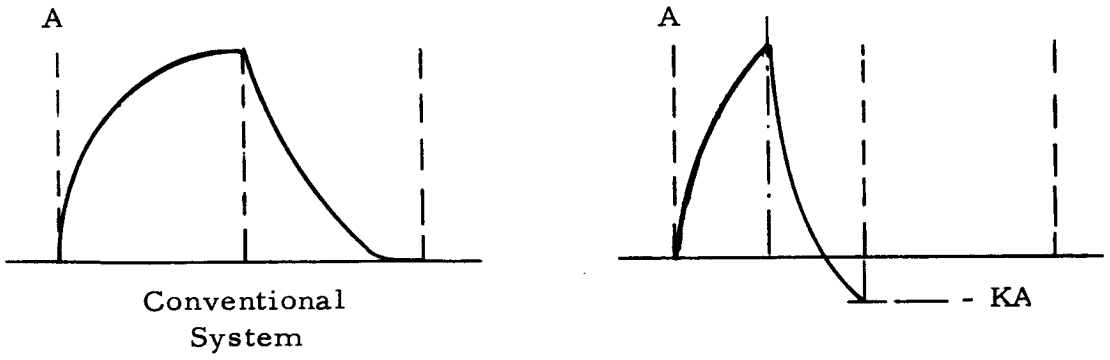


Figure 3-23. Comparison of Pulse Waveforms for Conventional and RTC Systems

Figure 3-24 shows a block diagram of the RTC transmitter for a four-channel system. The analog input (negative) for each channel is applied to a negative analog gate while the inverted signal (to develop the complement) is applied to a positive gate. The gates are switched on by a combined operation of the shift register and the drive pulses. The negative gates operate during

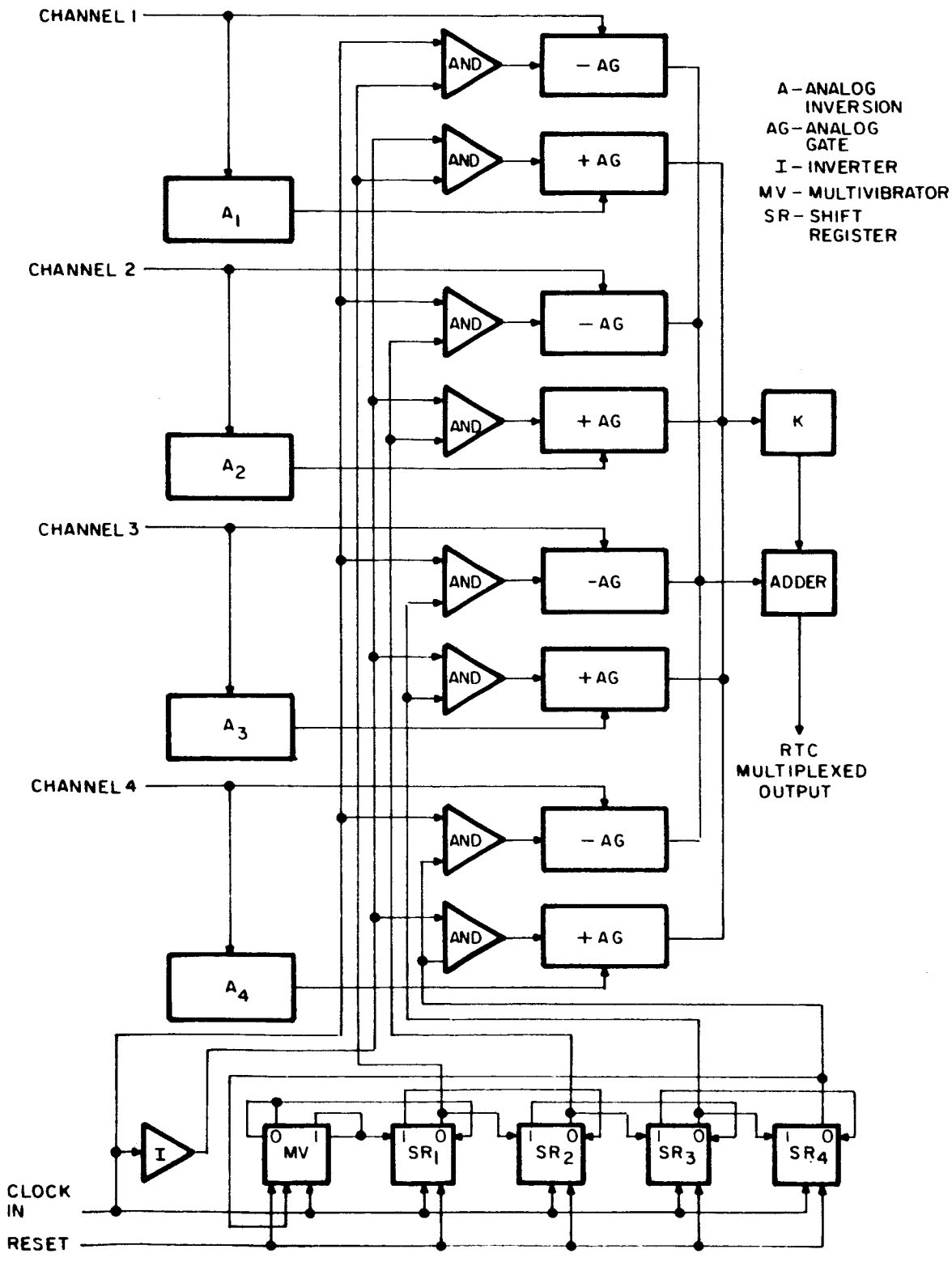


Figure 3-24. RTC Transmitter, Block Diagram

the time when the counter and drive are simultaneously up; this occurs one eighth of the frame time, occupying the first half of a channel. The positive gates operate when the drive is down and the counter output is up. This action is obtained by inverting the clock and processing the inverted clock and counter output by an AND gate.

The eight gated outputs are then fed into an adder after the complement has been attenuated by a factor K (less than 1). The output of the adder is therefore the RTC multiplexed output suitable for modulation and transmission.

Figure 3-25 shows a block diagram of the RTC receiver. The receiver consists of a 6-db-per-octave LPF whose 3-db point depends on the previously mentioned factor K, and analog gates and processing circuits for demultiplexing.

3.7.1.1 RTC Disadvantages

There are several inherent disadvantages in this system, even though it theoretically eliminates all crosstalk. The primary disadvantage lies in the restrictions on sampling rate. There is a 160-Hz bandwidth restriction imposed. This is assumed to be a sharp filter. However, the RTC assumes a low-pass filter with only a 6 db/octave rolloff. This implies that the 6 db/octave rolloff LPF located in the receiver must have a cutoff frequency much lower than 160 Hz. According to the developed formula

$$K = e^{-B/T_O},$$

if we lower the cutoff frequency we must either lower the sampling frequency or increase the power output by using a larger K. Either choice is undesirable. It has been found in fact that with K of 0.7 the sampling rate is only 60 Hz for acceptable crosstalk. This means a 120-Hz bandwidth, or 40 Hz of the 160 Hz allotted is unused.

The second difficulty in this system is the fact that the transmitter must contain additional gates and associated circuitry involved in deriving the complement function and adding it to the output signal.

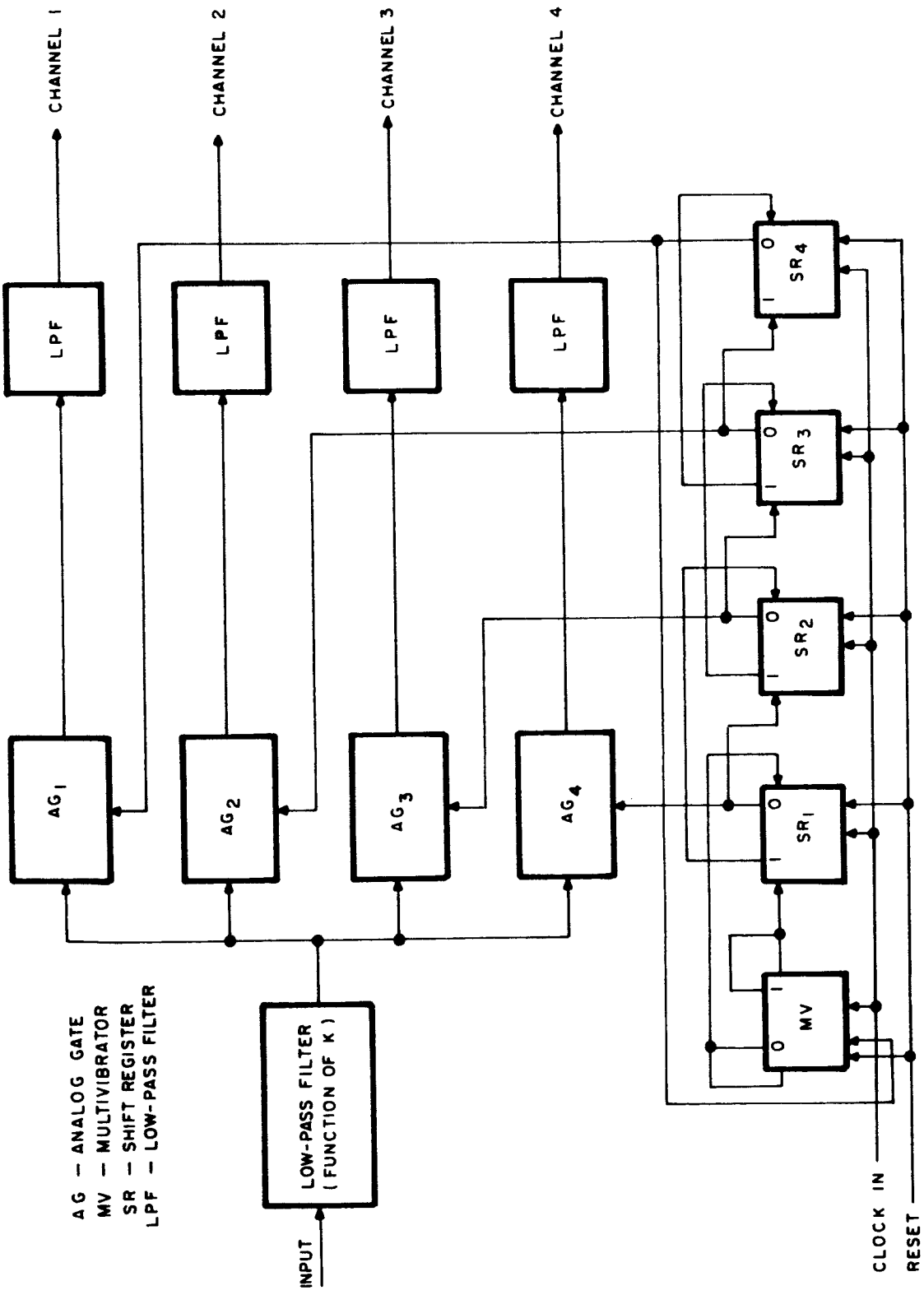


Figure 3-25. RTC Receiver, Block Diagram

3.7.2 Finite Memory Filter (FMF) System

The second system studied is the Finite Memory Filter. The FMF is essentially made up of a tapped delay line whose taps feed back delayed information into a network which cancels any crosstalk present from the previous channels. A simplified block diagram is shown in Figure 3-26. This system has the advantage of canceling crosstalk in any number of channels, rather than in only one channel as in the RTC approach.

With the FMF, all the complex circuitry is located in the receiver, and only a clock, five flip-flops and four negative analog gates are located in the transmitter. In addition, power is saved by virture of the fact the complement is not transmitted.

Since delay lines have fixed delays, it was decided to build a digital approximation of the delay line so various sampling rates could be tried. A block diagram of this approach is shown in Figure 3-27. The multiplexed input is stored in the adder and at the same time applied to series of sample-and-hold circuits. The clock input, operating at the channel rate, operates the fourth S & H on the first positive excusion, transferring the voltage from the third sample-and-hold to the fourth. The third S & H operates at a time delay later, equal to the width of the single-shot in the fourth sample-and-hold, transferring voltage from the second sample-and-hold to the third. This process continues back to the first sample-and-hold until such time when the fourth sample is activated again. The reason for the delay in transference of voltage from one sample-and-hold to the next is that information is destroyed if simultaneous transference is used. This operation occurs four times per frame, effectively simulating a delay line, if narrow timing pulses are used with a maximum delay of one frame.

The outputs of each sample-and-hold are inverted and attenuated, and added back to the input. The amount of attenuation is a function of the number of taps and the delay line.

3.7.3 Detailed FMF Discussion

3.7.3.1 PAM Multiplexing Technique

The multiplexing system employed is a time-division multiplexing system in which each parameter occupies the interconnecting link exclusively for a given period of time. This period of time is $1/N\ f_s$, where N is the number of multiplexing channels and f_s is the sampling rate. The speed

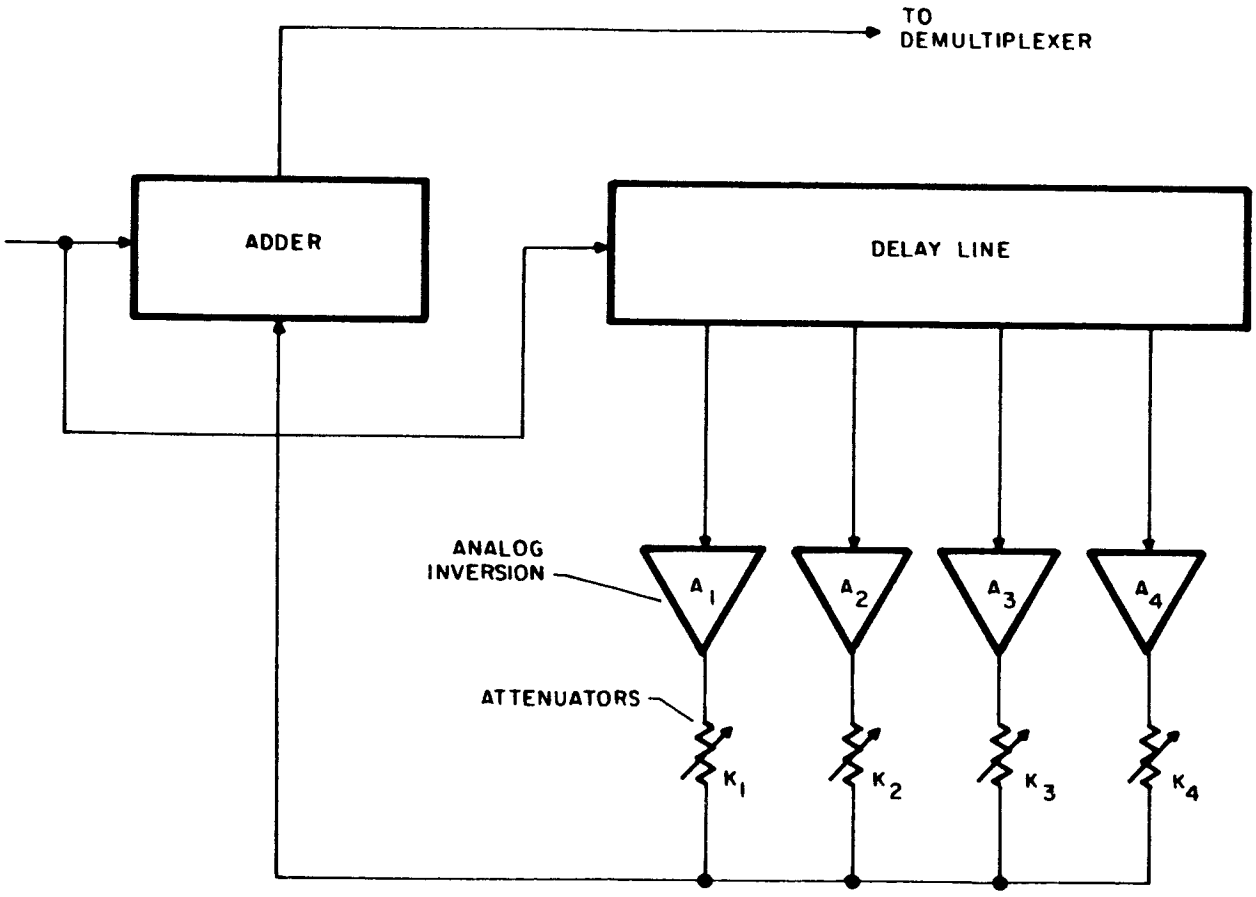


Figure 3-26. FMF, Simplified Block Diagram

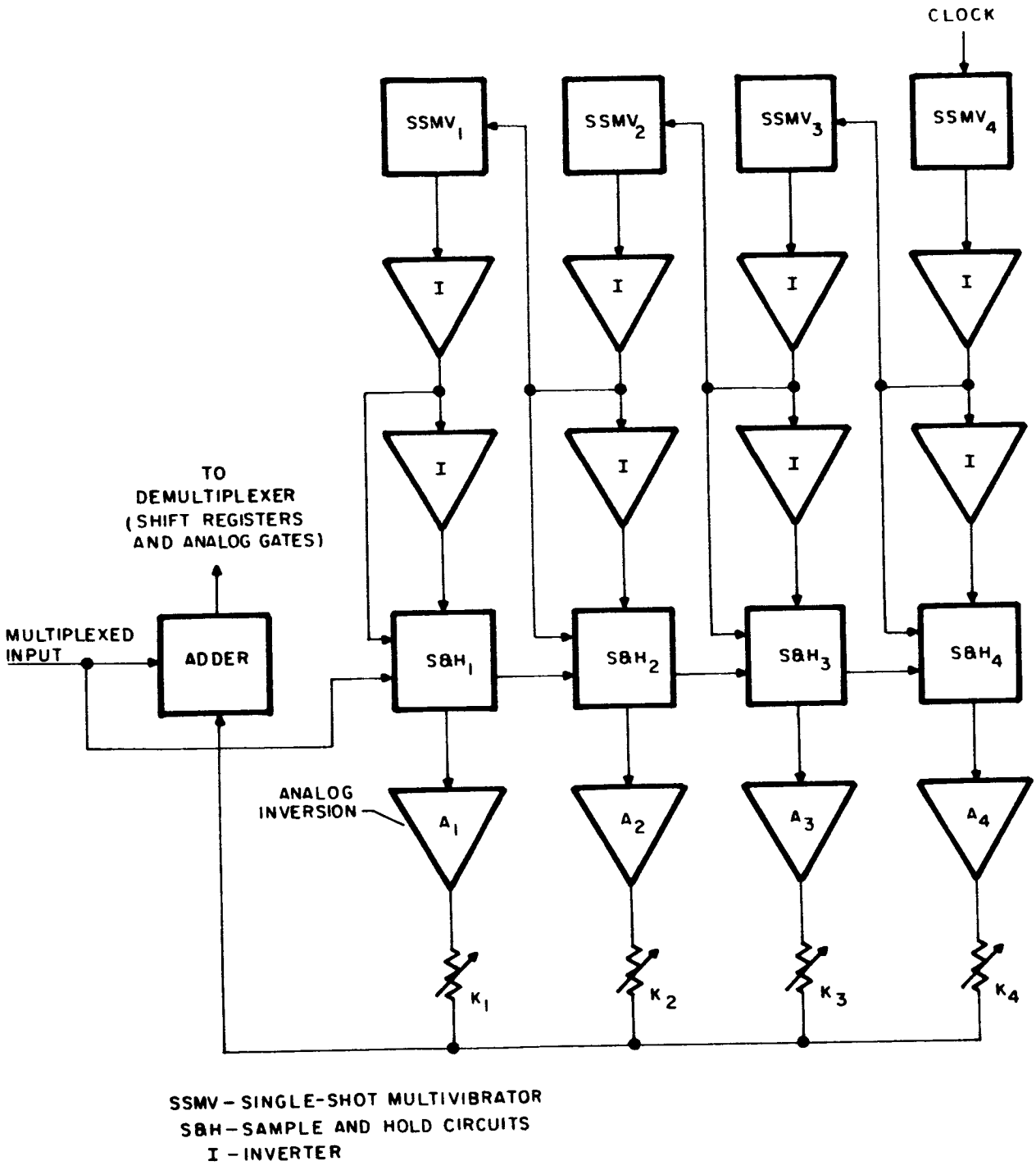


Figure 3-27. FMF, Digital Implementation

of sampling, or commutation, is determined by the Nyquist rate and will be twice the highest frequency component contained in the channel. Since each channel is band limited to 20 hertz and there are five information channels plus two sync channels, the frame rate is then 40 hertz, and the minimum channel rate of change is 280 hertz. In the receiver the sync pulse is extracted and the PAM signal is passed through a Finite Memory Filter and the demultiplexing gates, where the information contained in each channel is distributed to its respective synthesizer input.

It should be noted here that total transmitted bandwidth is 140 Hz filtered to 160 Hz. However, more bandwidth compression can be realized by transmitting sync only during times of speech silence. This requires a more stable clock than that which is used in the present multiplex system. If the 4-Hz tap described from the Apollo Central Timing Equipment described in Section XIX of Apollo Training Telecommunication Subsystem Study Guide, Course Number A-630 were used, an effective 100-Hz bandwidth of processed speech would be realized since the two 20-Hz sync channels would be dropped from the information frame. Sync would then be effectively accomplished in the following manner. Upon detecting that there was silence by the zero level of the amplitude channels, the sync channel would be superimposed on channel five of the five-channel multiplexer. During sync transmission the information from the analyzer would not be transmitted since the control gates (except for channel five) would be disabled. The sync pulse would be threshold detected and selected in the same manner as described below and the receiver would be rephased with the transmitter.

3.7.3.2 PAM Multiplexing Operation

The block diagram of the multiplex/demultiplex system is shown in Figure 3-28. The multiplexer consists of a clock and counter to derive the basic channel rate; a control register and gates to derive the steering pulses that control the multiplex gates which gate each parameter from the speech analyzer onto the interconnecting link after it is low-pass filtered to 160 Hz.

In the receiver, the filtered, PAM signal is sent directly to the sync-extracting circuits and a finite memory filter. The derived sync pulse phase locks the receiver clock by clearing the digital binary counter associated with the crystal-controlled clock oscillator. Then, as in the multiplexer, the demultiplexing gates are selected so that each channel is selected and distributed to the proper synthesizer input. Before the channel information is distributed it is corrected for channel crosstalk in the finite memory filter.

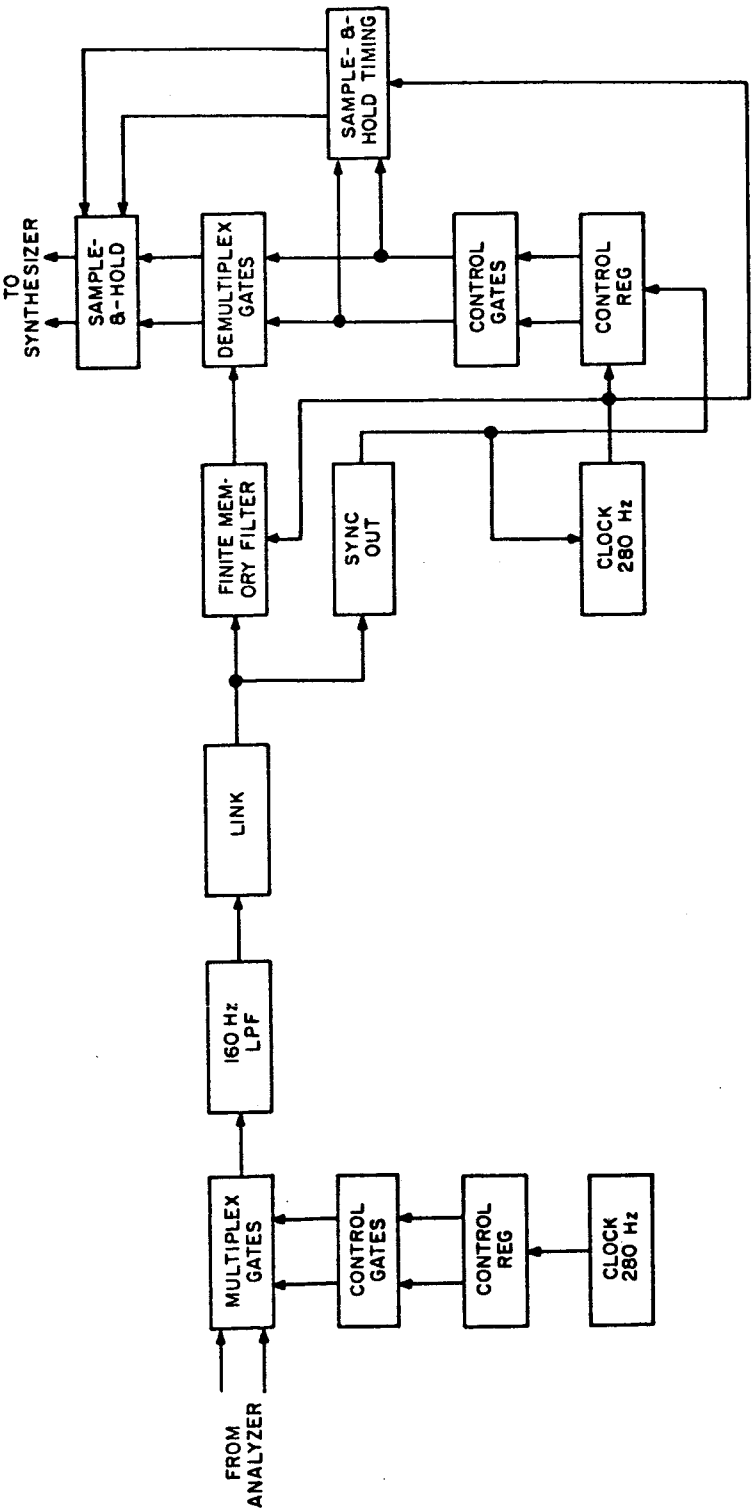


Figure 3-28. Multiplex/Demultiplex Block Diagram

3.7.3.3 Multiplex and Demultiplex Clock and Counter

The multiplex clock and counter is shown in Figure 3-29 and consists of a 32-kHz cyrstal-controlled oscillator having a 70 PPM/°C temperature coefficient and a 1.3 volt minimum sinusoidal output into a 4.7 kilohm load. This output is shaped to drive milliwatt logic arranged as a binary counter, which counts to 112 to provide a 286-hertz square-wave output which, after shaping and buffering, yields 200 nsec trigger pulses at a 286-pulse-per-second rate for the channel control register.

The demultiplexer clock and counter shown in Figure 3-30 is basically the same as that used in the multiplexer except for the additional inverter on the output to provide both positive and negative triggers necessary to drive the channel-control register, finite memory filter timing circuits, sample-and-hold timing circuit, and sync circuits. This counter also has a clear line used for phase locking the receiver clock with the received signal. The internal logic of the PL909 and PL910 is shown in Figure 3-31, and Figure 3-32 shows the internal logic of the PL913 and PL939.

3.7.3.4 Channel Control Register and Gates

The channel control register and gates for the multiplexer are shown in Figure 3-33. The control register is a four-bit milliwatt-logic ring counter using the PL913. The selection gates select one channel gate at a time so that each channel exclusively occupies the interconnecting link for a period between control register trigger pulses. The internal logic of the standard gate used for multiplexing and demultiplexing is shown in Figure 3-34. The clamp supply serves the dual purpose of clamping each channel input to a +4 volt level, as well as supplying the +5 volt level for the sync channel.

The demultiplexer channel control register and gates shown in Figure 3-35 are basically the same as those used in the multiplexer except for the addition of the reset circuit and trigger delay. These are necessary for proper phasing of the steering gates with the received signal.

3.7.3.5 Sync Extractor

The sync extractor shown in Figure 3-36 operates in the following manner: The PAM signal from the interconnecting link is passed through a threshold detector. The threshold detector produces an output when the PAM signal exceeds a given threshold level. This will occur when the sync

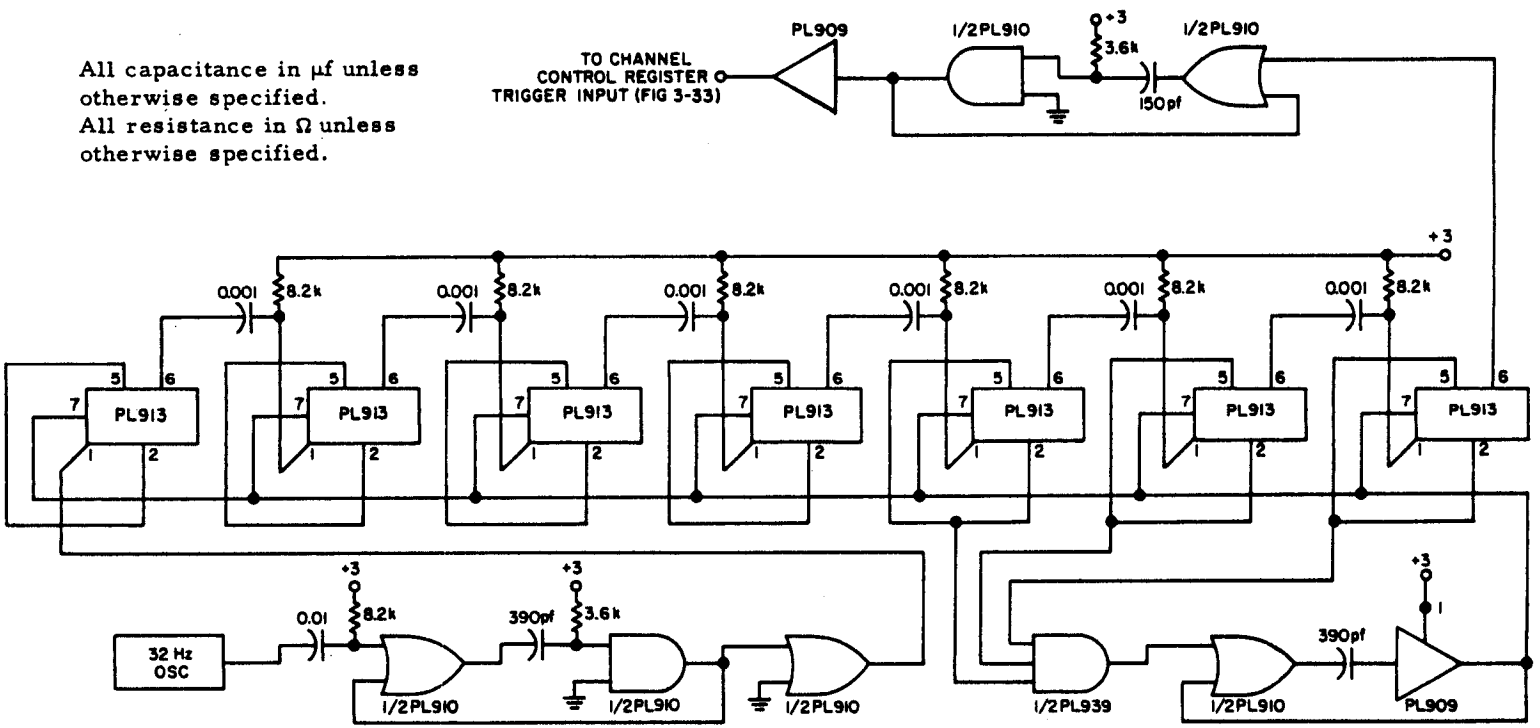


Figure 3-29. Multiplex Clock

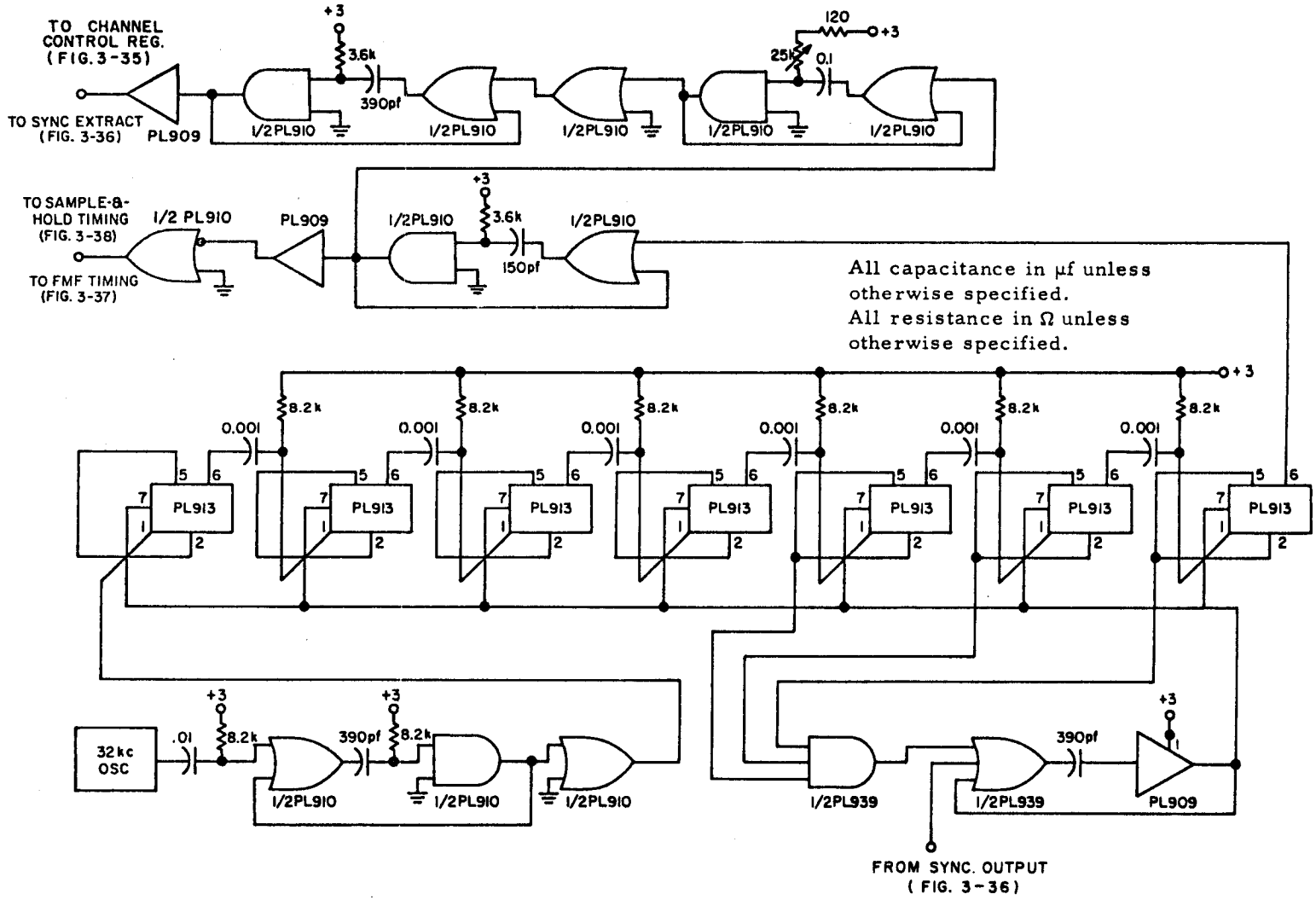


Figure 3-30. Demultiplex Clock

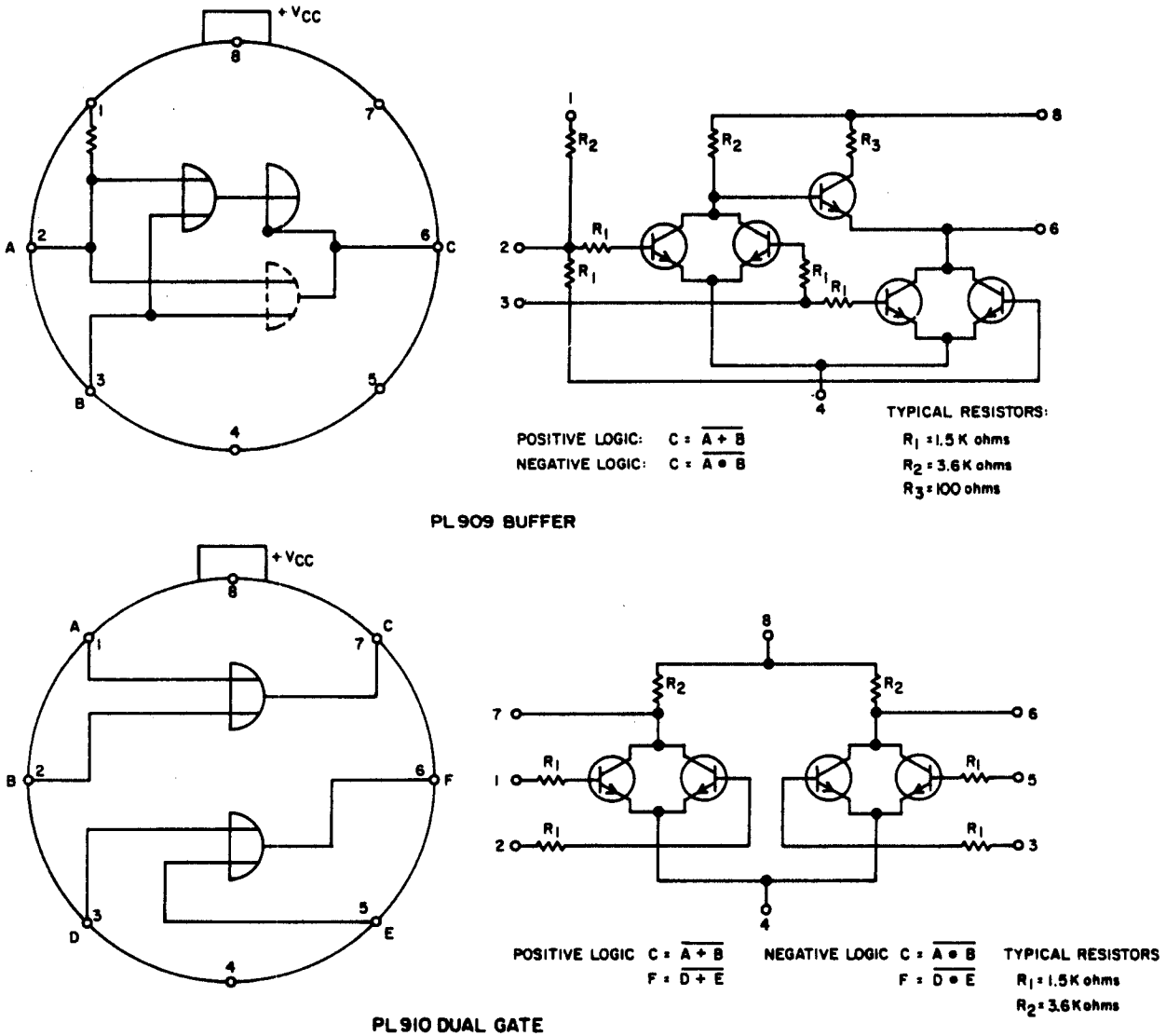


Figure 3-31. Milliwatt Micrologic PL909 and PL910

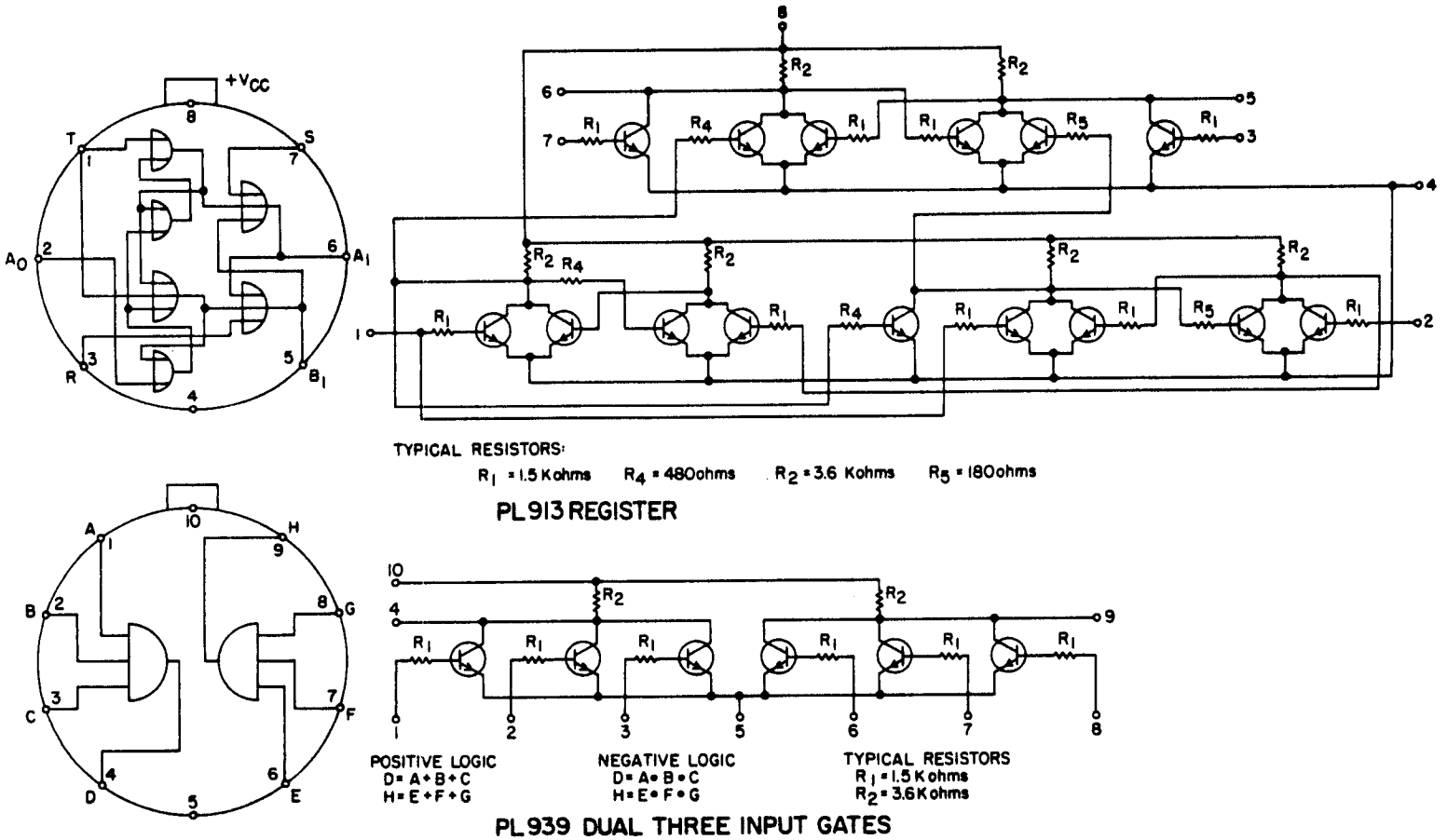


Figure 3-32. Milliwatt Micrologic PL913 and PL939

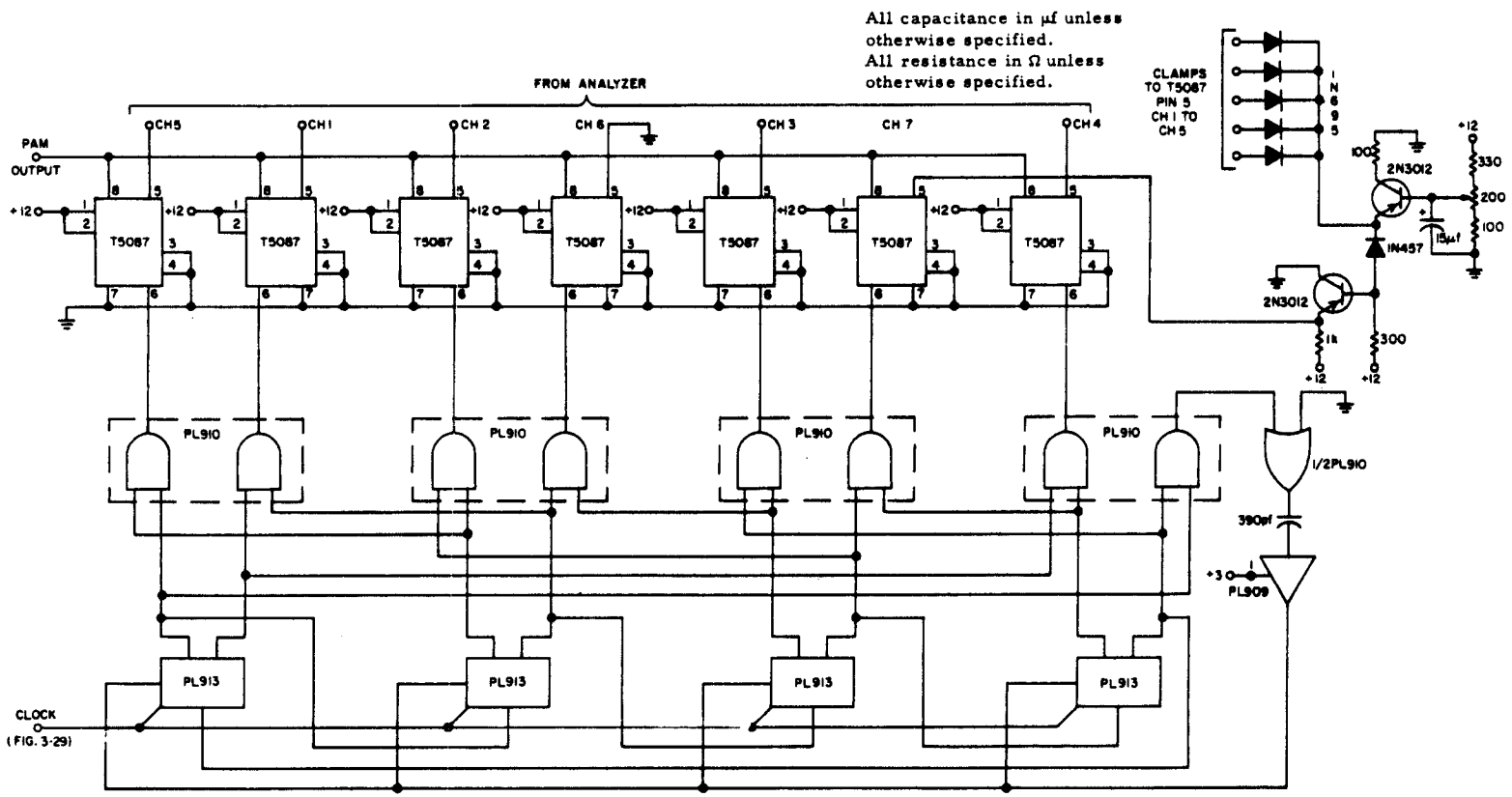


Figure 3-33. Channel Control Register and Gates — Multiplexer

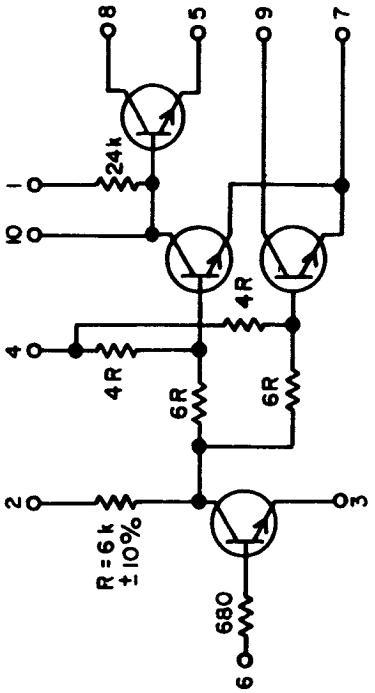


Figure 3-34. Standard Gate

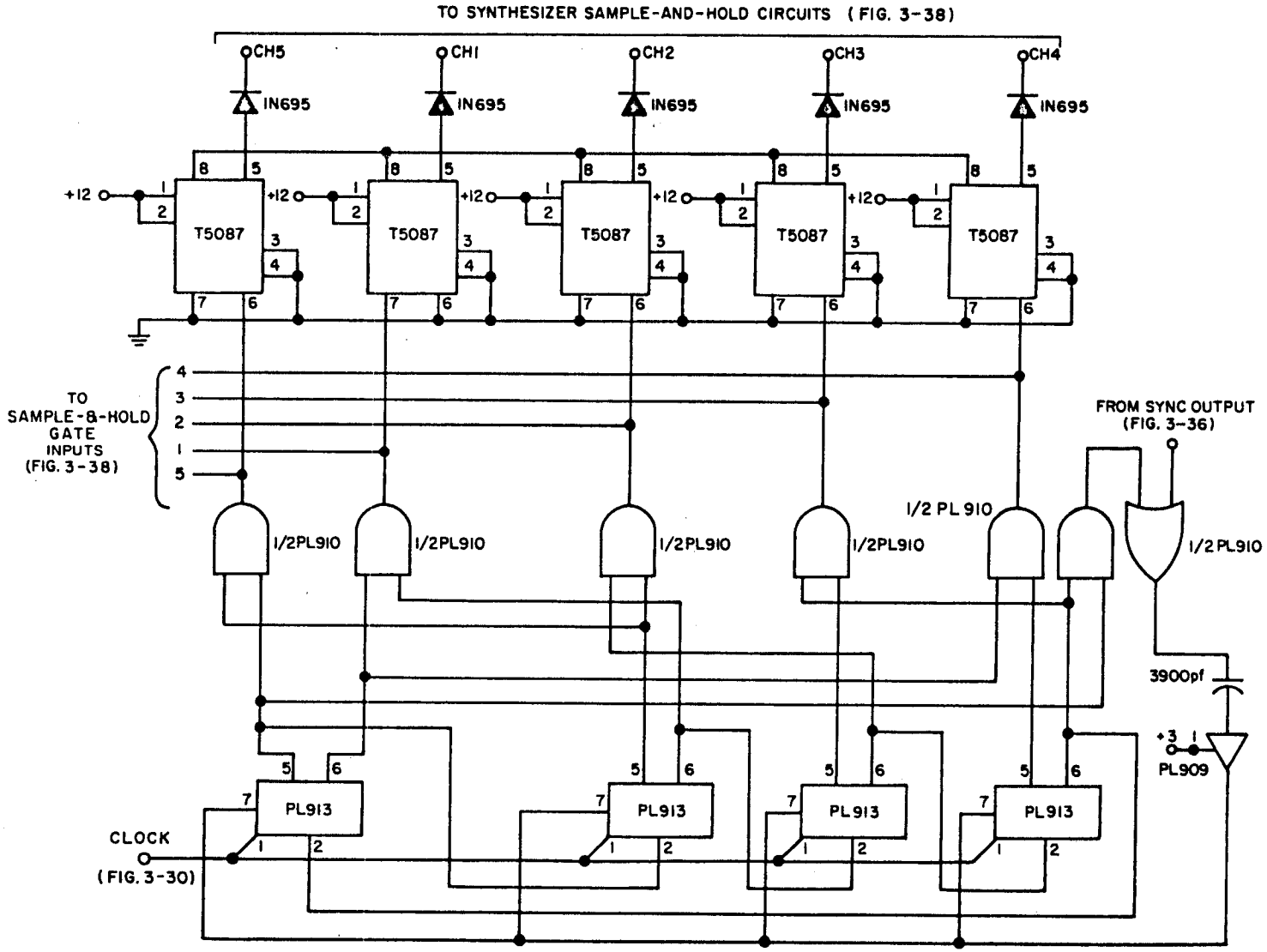


Figure 3-35. Channel Control Register and Gates — Demultiplexer

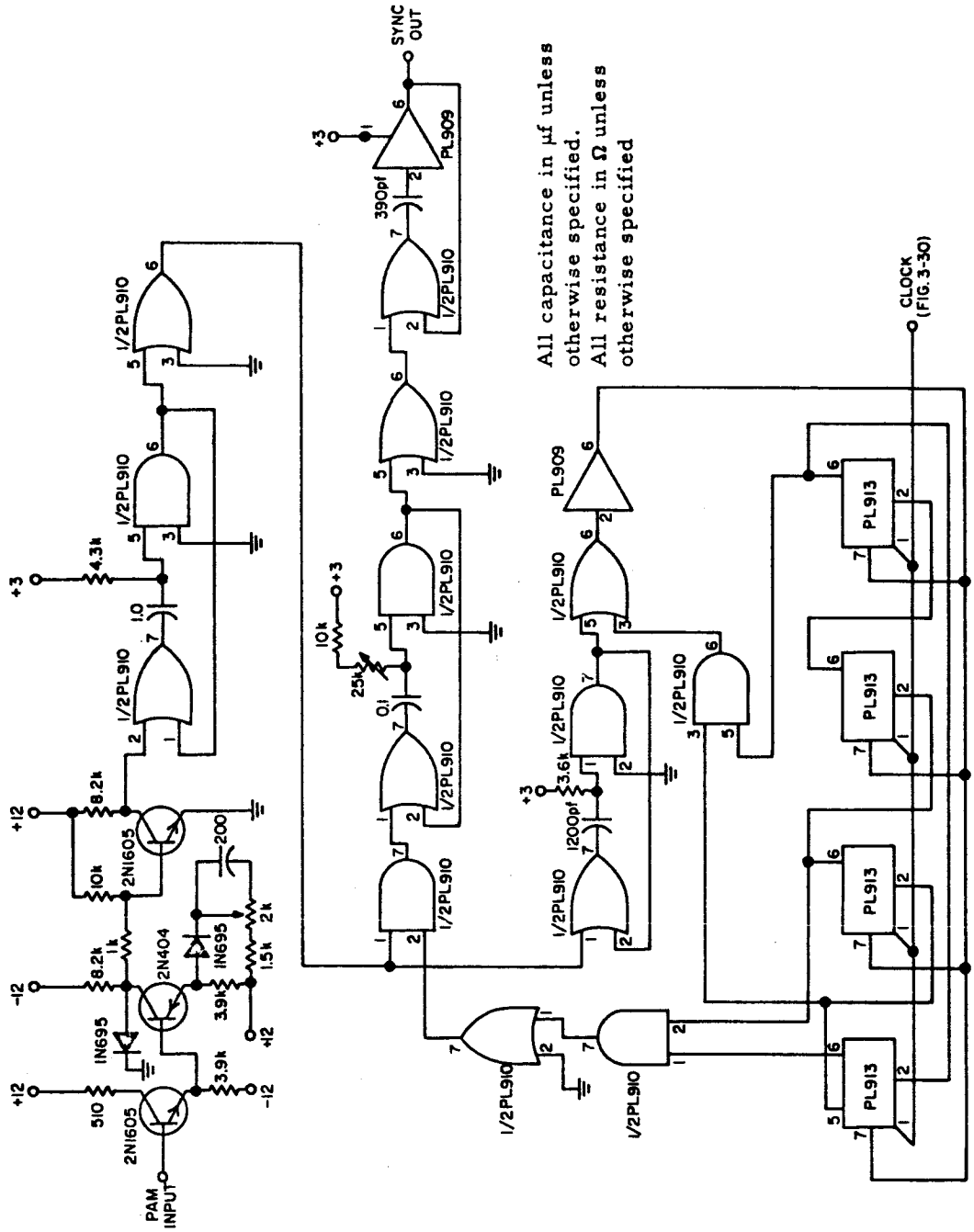


Figure 3-36. Sync Control

channel, which is larger in amplitude than the information channels, is present at the threshold detector or when a noise pulse greater than the threshold level is present. Time stability of the extracted sync pulse is obtained by preceding the sync channel with a blank channel of zero amplitude so that the time from the beginning of the sync channel to threshold is constant. In order to distinguish the sync pulse produced from large amplitude noise, a counter is used. This counter is reset by the pulse which may either be a sync pulse or noise, and the counter then counts seven channels by timing pulses from the system clock at the channel rate. If the reset pulse was a true sync pulse the counter output would permit the next pulse occurring at the sync channel time to pass through an AND gate and this pulse would be the pulse used to rephase the receiver. If this second pulse did not occur at the sync channel time the counter would be reset and the process restarted. An operative sync pulse would not be produced until two sync channels without high-level noise between them were seen by the sync circuits. The output of the sync AND gate is then delayed and shaped to be used by the receiver clock and registers.

3.7.3.6 Finite Memory Filter

The Finite Memory Filter shown in Figure 3-37 is a channel cross-talk correction device, used to correct the error in a given channel caused by the information in the preceding channel due to the bandwidth limitations of the interconnecting link. The band-limited PAM signal is applied to two adders and a sample-and-hold circuit. Consider that the sample-and-hold circuit samples a channel at the latest time permitted by the timing tolerances of the system. This level is subtracted from the changing information after the sample pulse. The difference between the signal levels of the two channels is then attenuated and added to the incoming signal. The amount of attenuation is determined by the frequency limitations of the incoming signal and the clock rate, and, for close-tolerance clocks and filters, it is essentially constant. After this attenuated difference between the previously sampled channel and present channel is added to the present channel, the PAM input is sent to the demultiplex gates where it is gated to the sample-and-hold circuit corresponding to the synthesizer input selected to receive the information. The information is then sampled and held at the synthesizer input. After this sample time the input sample-and-hold again samples the channel to apply the correction to the next channel, and the process is repeated.

The sample-and-hold timing gates shown in Figure 3-38 steer the sample pulses for the channel sample-and-hold circuits to the sample-and-hold circuits corresponding to the channel being selected.

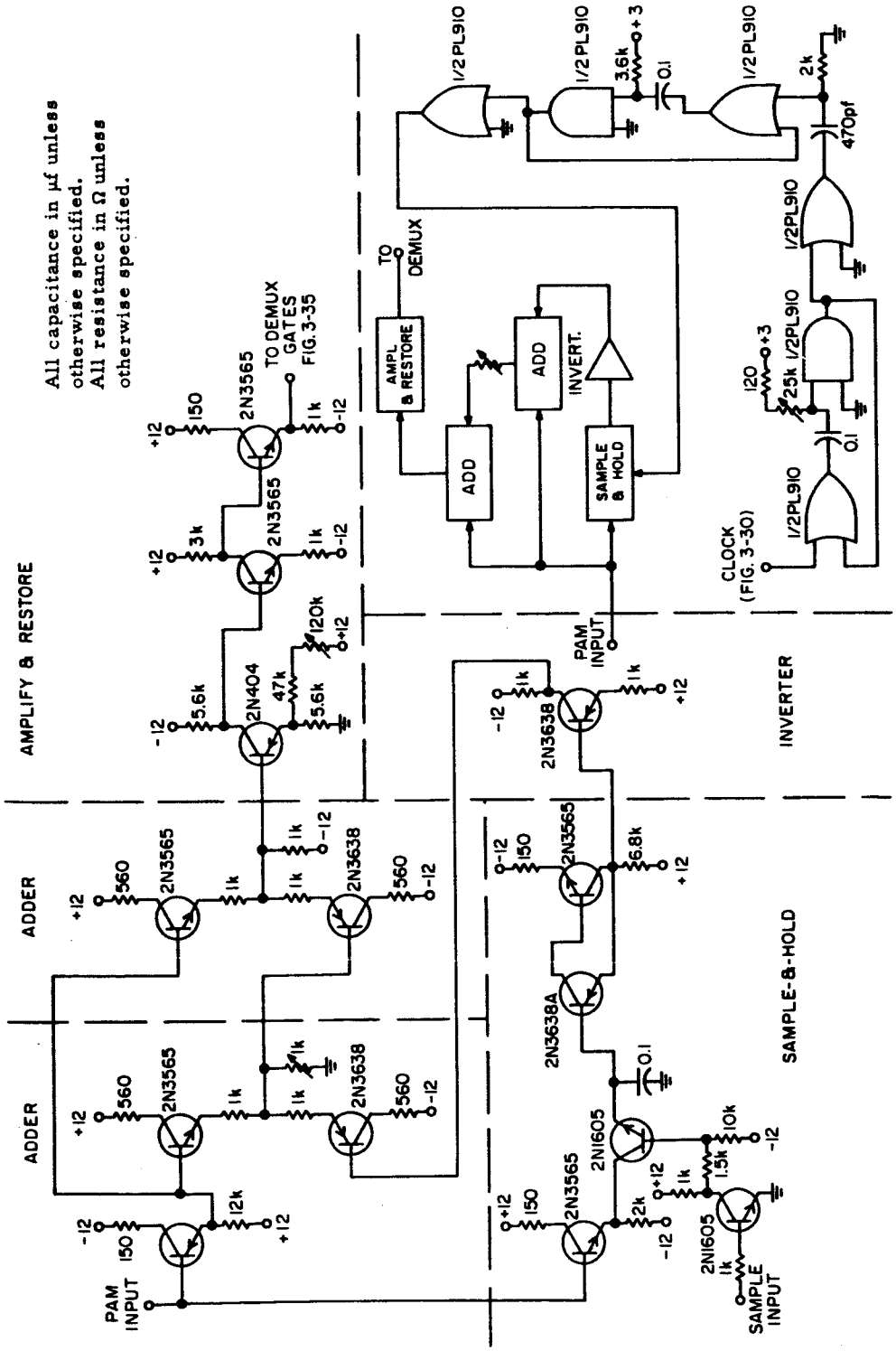


Figure 3-37. Finite Memory Filter

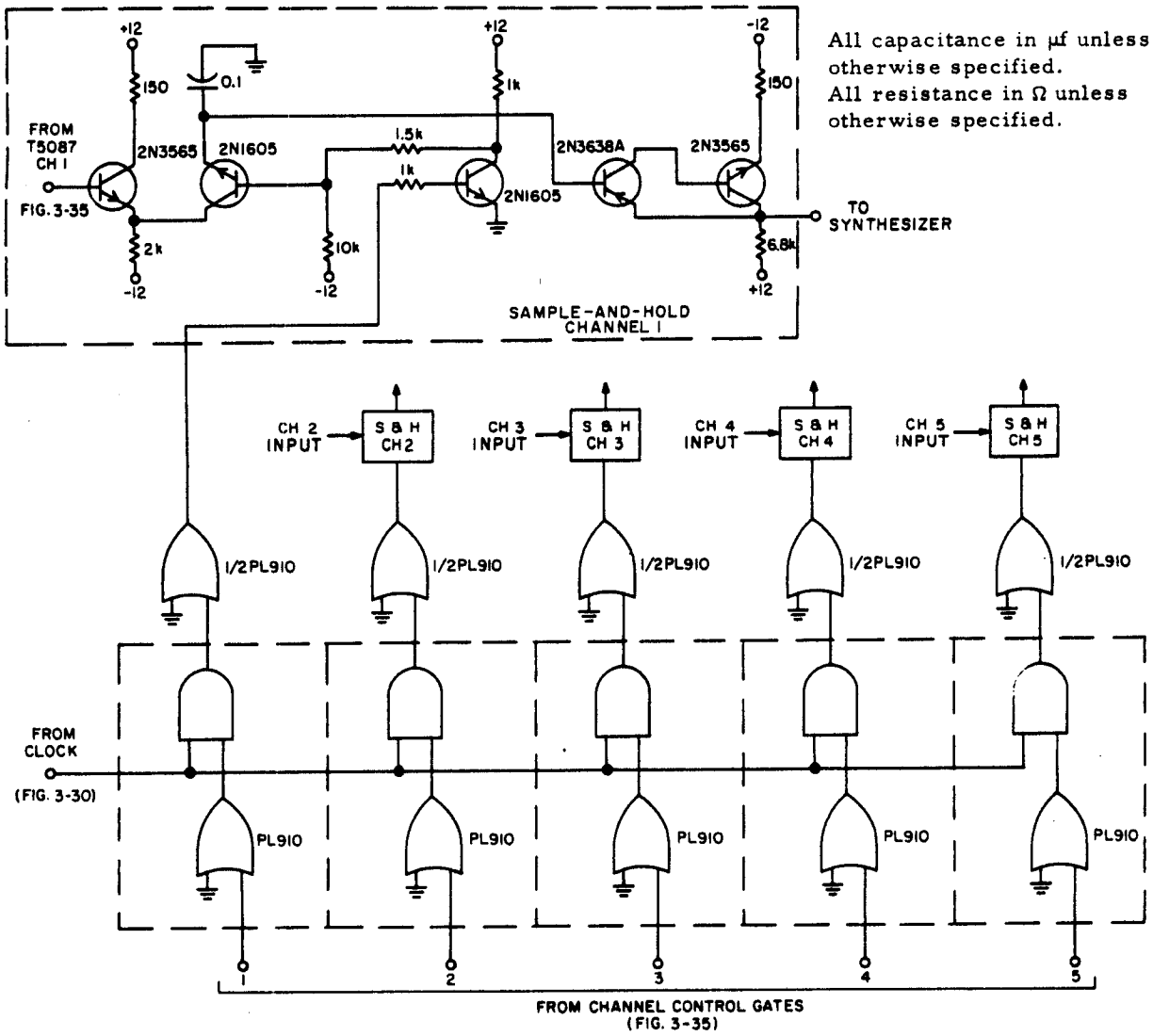


Figure 3-38. Channel Sample-and-Hold Timing

The multiplex/demultiplex and finite memory filter timing are shown in Figures 3-39 and 3-40. The filtered PAM signal is phase shifted approximately one channel, as shown in the top picture of Figure 3-39. The bottom picture of Figure 3-39 shows the input and output of the finite memory filter along with the sync pulse. With the output of the finite memory filter are shown the channel sample pulse with the finite memory filter sample pulse and the output from the number one channel demultiplex gate. The timing is such that the channel sample-and-hold pulse precedes the finite memory filter sample pulse which occurs at the end of the channel, which is the peak of the sync pulse in the finite memory filter output trace in Figure 3-40. The demultiplex gate is open in sufficient time so that the channel sample pulse samples the information at the same time during each frame.

3.7.3.7 Channel Contents

Each channel in the multiplexer receives the following parameters from the analyzer, and the corresponding channel in the demultiplexer distributes them to the corresponding synthesizer input:

Channel 1: First Formant Amplitude

Channel 2: Second Formant Amplitude

Channel 3: First Formant Frequency

Channel 4: Second Formant Frequency

Channel 5: Pitch/Voice

Channel 6: Blank-zero amplitude

Channel 7: Sync

3.7.3.8 Multiplexer Logic

The logic elements used in the multiplexer and demultiplexer control are the Philco Milliwatt Micrologic family, which is a compatible set of microelectronic integrated digital logic circuits capable of performing complex NAND/NOR logic functions. The elements are manufactured by planar epitaxial diffusion techniques on a monolithic silicon chip.

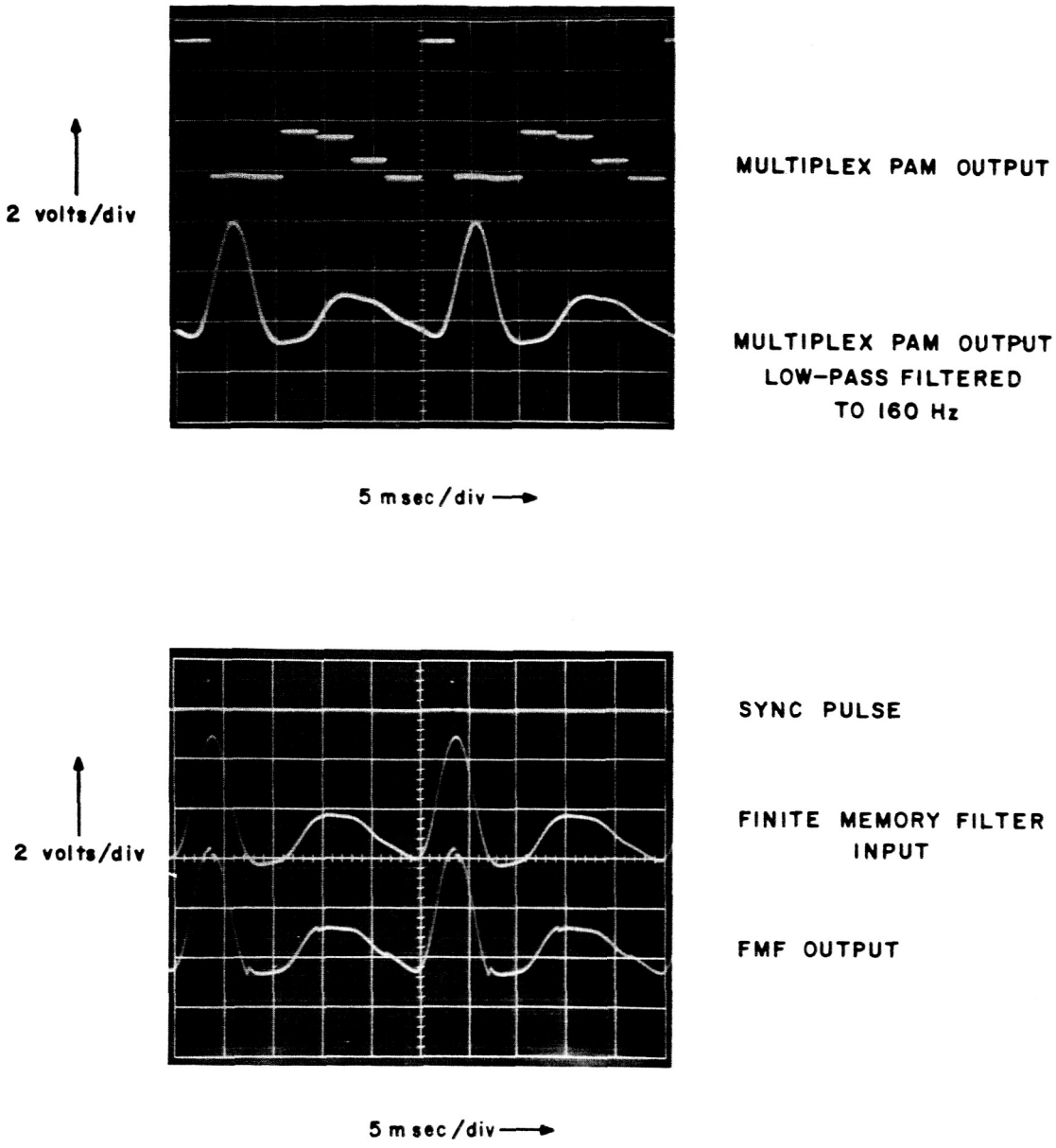


Figure 3-39. Multiplexer and Sync Timing

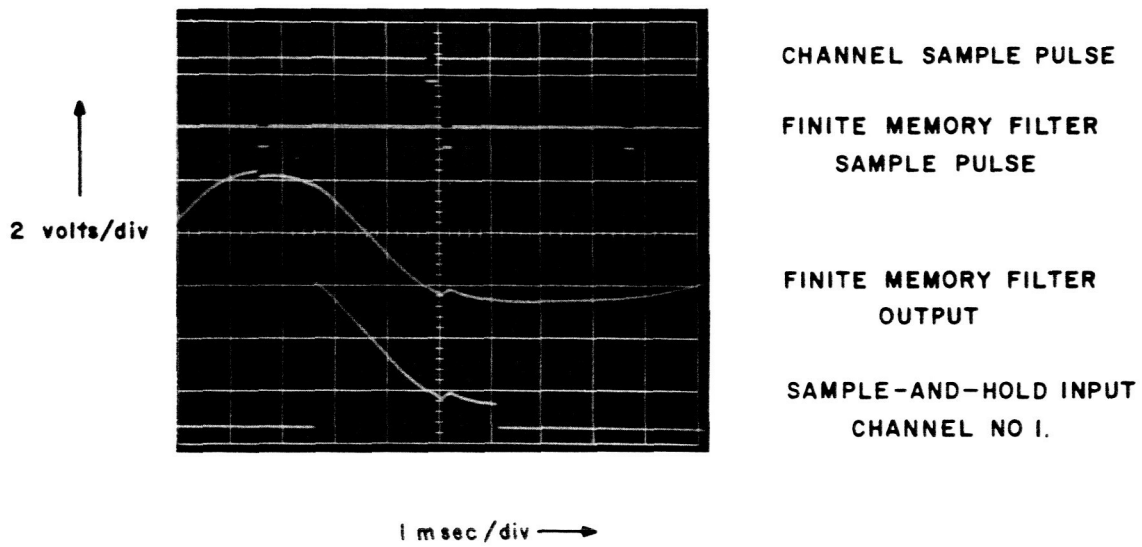


Figure 3-40. Demultiplexer Sample-and-Hold/Channel Timing

The Milliwatt Micrologic elements are intended for use in high-speed logic systems at low dc power dissipations. A typical propagation delay of 40 nanoseconds at a power dissipation of 2 mw may be obtained from the basic gate circuit.

A brief description of each of the Milliwatt Micrologic elements used in the multiplexer and demultiplexer follows:

a. The PL909 Milliwatt Micrologic Buffer is a low-impedance driver having considerably higher fanout capability than the basic Milliwatt Micrologic gate. An internal timing resistor is available to permit capacitive coupling for use in monostable and astable multivibrators and in pulse differentiation.

   AVERAGE POWER DISSIPATION (+25°C): 10 mw at 50% duty cycle.

b. The PL910 Milliwatt Micrologic Dual Gate is a dual two-input NAND/NOR gate which may be used as a set-reset flip-flop, a double inverter, a pair of inverters, or with the PL921 Expander to provide dual NAND/NOR with increased fan-in.

   AVERAGE POWER DISSIPATION (+25°C): 4 mw.

c. The PL913 Milliwatt Micrologic Register is a set-reset Gated "D" flip-flop for use in shift register or counter applications. When T is high, direct set and reset inputs, S and R, control the output state. If S and R are low, the state present at A_o will be stored when T goes from high to low. A change in A_o while T is low will not affect the output.

   AVERAGE POWER DISSIPATION (+25°C): 15 mw.

d. The PL939 Milliwatt Micrologic Dual Three-Input Gate is a NAND/NOR gate which may be used as a set-reset flip-flop, a non-inverting gate, a pair of inverters, or with the PL921 Expander to provide dual NAND/NOR with increased fan-in.

   AVERAGE POWER DISSIPATION (+25°C): 4 mw.

The multiplexing and demultiplexing gate is the Philco integrated circuit type T5087 Standard Gate Module with normal and complementary outputs.

SECTION 4

CONCLUSIONS AND RECOMMENDATIONS

The 160-Hz bandwidth, two formant speech communicator developed on this contract has shown phonetically balanced intelligibility test scores of 70 - 80 percent in a multiplexed connection (see Figure 4-1). For intelligibility in the order of 75 percent a signal-to-noise ratio of +20 db referred to a 3.2 kHz link bandwidth is required. This required S/N ratio has two causes:

a. When the speech parameter signals (formant frequency, formant amplitude and pitch) become noisy, the intelligibility degrades rapidly. It is believed that the reason is this: Normal full spectrum speech contains great redundancy and it is this redundancy that allows voice communication in very noisy channels (87 percent phonetically balanced test score in +6 db signal-to-noise ratio channel)*. When this redundancy is removed the remaining information must be accurate (noise free) or intelligibility is lost.

b. While there are some minor operational defects in the multiplexer-demultiplexer system the major problem here is in the type of coding, pulse amplitude modulation, (PAM) that was chosen. Because of the requirement for noise free signals at the synthesizer a better choice of coding would have been pulse code modulation (PCM) or some other system which exhibits a noise threshold similar to a frequency modulation link. Using delta modulation, the system would operate at +6 db signal-to-noise ratio with essentially no degradation (error rate 0.01) and provide a 12.7 db transmitter power saving and about 80 percent PB word intelligibility.**

As a result of experience gained on this contract and on other current Philco speech signal processing contracts, a number of improvement areas in the analog signal processing are seen to exist. Four areas offering large improvement are:

* "Some Comparisons Between Rhyme and PB-Word Intelligibility Tests," K. D. Kryter et. al. Journal of Acoustical Society of America June 1965, p 1146.

** For further discussion refer to Task 6 report "Space Mission Applications Study."

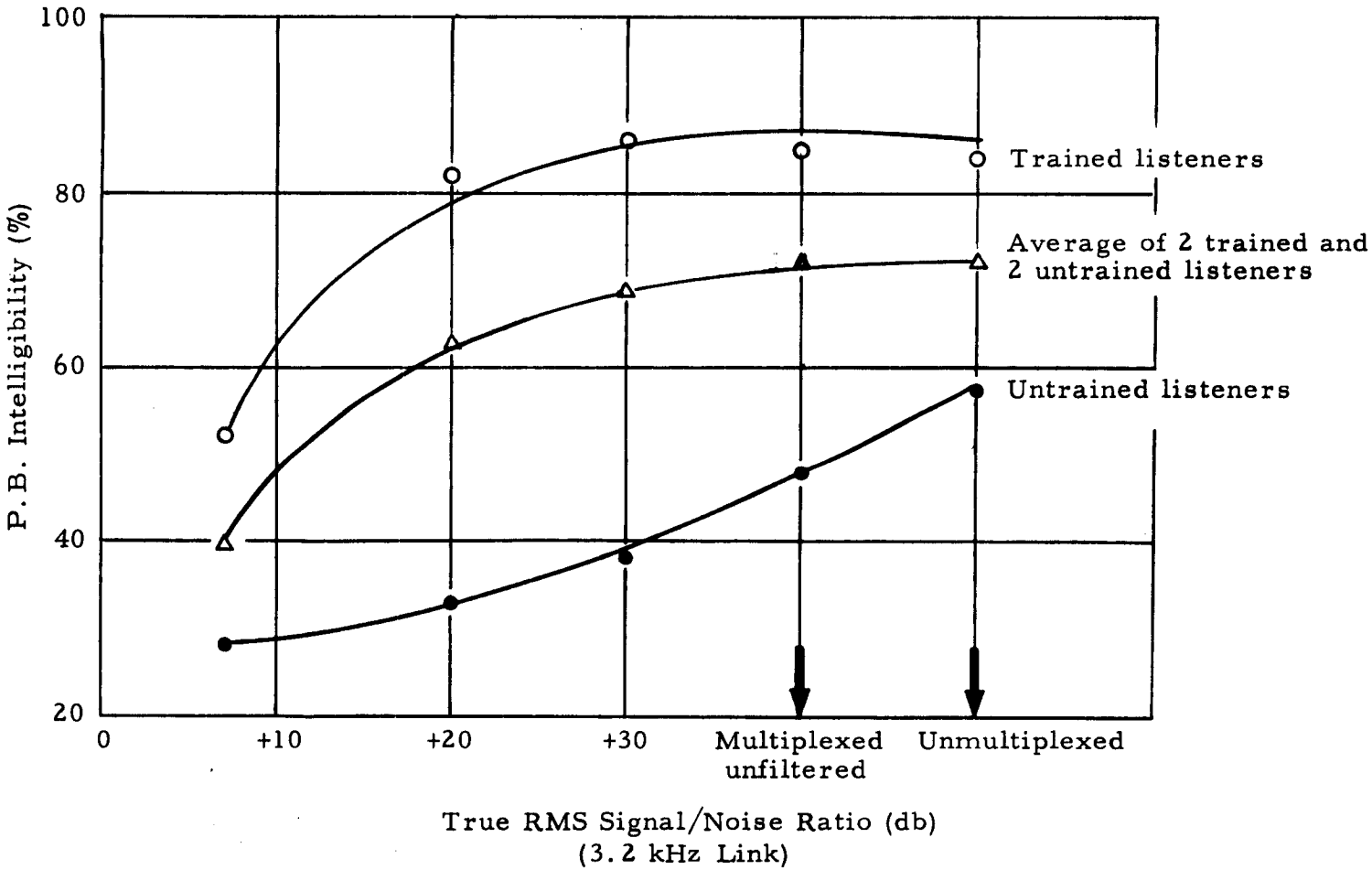


Figure 4-1. P. B. Intelligibility Versus Signal/Noise

(1) Simpler formant frequency analyzers based upon pulse summation in a low-pass filter. This would save about 38 transistors and associated circuitry in the analyzer and make alignment much simpler.

(2) New voicing extractor developed on a current Philco speech recognition contract. This extractor seems more accurate than the one presently used and is simpler (using 17 less transistors).

(3) Frequency domain pitch extractor similar to the type used in Philco KY/537, KY/585 channel vocoders. There would probably be no significant simplification but decreased speaker sensitivity is expected.

(4) Improved dynamic range in the amplitude extractors would improve performance on nasal consonants. By using microcircuits the improvement would also save six transistors in the amplitude extractors.

An improvement in synthesized voice quality may be obtained by a change in the pitch pulse drive to the formant filters in the synthesizer. This would result in a slight increase in synthesizer complexity — perhaps 6-8 transistors.

As explained in the body of the report the original design concept centered on the Single Equivalent Formant (SEF) concept. While the SEF parameter was proven capable of carrying vowel information, it was not successful in transmitting and synthesizing perceptually acceptable diphthongs, semi-vowels and liquids. The reason for this is believed to be that the SEF parameter contains the transient information in an implied manner which is not adaptable to the simple synthesis technique now being used. It was judged that successful solution of the SEF transient problem was not probable within the program, so the two-formant system approach was followed instead of the SEF concept. This resulted in a wider bandwidth system but one which satisfied the program requirements and was completed on time.

While it may yet be possible to build extremely low bandwidth speech communications systems (about 60 Hz) using the SEF concept and sophisticated synthesis techniques based upon a priori knowledge of the transient phonemes, such an idea will require additional study to be proven.

APPENDIX A

IDENTIFICATION OF KEY TECHNICAL PERSONNEL

A.1 NAMES, TITLES, AND MAN-HOURS

The personnel, and the approximate number of man-hours worked on this contract are listed below.

Engineer	Title	Hours
H. Drucker	Engineer, Electronics	200
D. M. Jurenko	Group Supervisor and Project Leader	600
J. Loe	Senior Engineer	1250
J. Schultz	Project Engineer	250
C. Teacher	Supervisory Engineer, Research Section	150

APPENDIX B

HANDSET CHARACTERISTICS

The handset used on this program is equipped to provide noise-canceling performance, and was originally developed to meet the requirements of the Soft-Talk System (484L). More specifically, the noise-canceling microphone was designed to operate in the environment of jet aircraft in conjunction with speech-processing equipments.

The following is a list of specifications governing the noise-canceling dynamic microphone:

Electrical

a. The frequency response of the microphone shall be ±3 db from 70 Hz to 4000 Hz with reference at 1000 Hz.

b. The microphone shall possess noise-canceling properties over the frequency range from 70 Hz to 4000 Hz. The noise-canceling properties over the range from 70 Hz to 300 Hz shall be approximately equal to the noise-canceling properties over the frequency range from 300 Hz to 4000 Hz.

c. In an ambient noise level of approximately 85 db of jet aircraft noise, as measured with a General Radio Sound Level Pressure Meter, Model 1551C or equivalent, the average talker signal-to-noise ratio shall be greater than 25 db.

d. The microphone amplifier shall be designed to operate from existing power supply voltage of -12 volts dc.

e. The microphone shall discriminate against unwanted sounds arriving from a distance in favor of sounds emanated from a close source and shall have directional properties that will result in rejection of noise originating at distances over one (1) foot.

f. The minimum output level of the microphone shall be -50 db.

The noise-canceling dynamic microphone and handset assembly shall meet or exceed the following environmental conditions:

a. Temperature – The handset shall be capable of satisfactory operation over the temperature range from -12°C to +50°C. It shall withstand a storage temperature of -62°C to +71°C.

b. Altitude – The handset shall be capable of withstanding (nonoperating) a reduced pressure of 3.44 inches of mercury equivalent to 50,000 feet altitude. Satisfactory operation is required in a reduced pressure of 20.58 inches of Hg equivalent to 10,000 feet altitude.

c. Humidity – The handset shall have satisfactory operation in a humid atmosphere from 0 to 98 percent at an ambient temperature of +50°C.

d. Shock and Vibration – As encountered during movement and handling through military and commercial transportation systems.

The frequency response of the actual noise-canceling microphone is given in Figure B-1. The design comprises a wide-range low-distortion dynamic microphone and integrated transistor amplifier. It provides a speech output level of approximately -20 dbm and can drive any load from 100 ohms to infinity.

The Stanford Research Institute, under subcontract to the Philco Corporation, performed a detailed evaluation of the delivered noise-canceling assembly, The results of that evaluation show a S/N improvement of approximately 17.6 db over the standard microphone used.

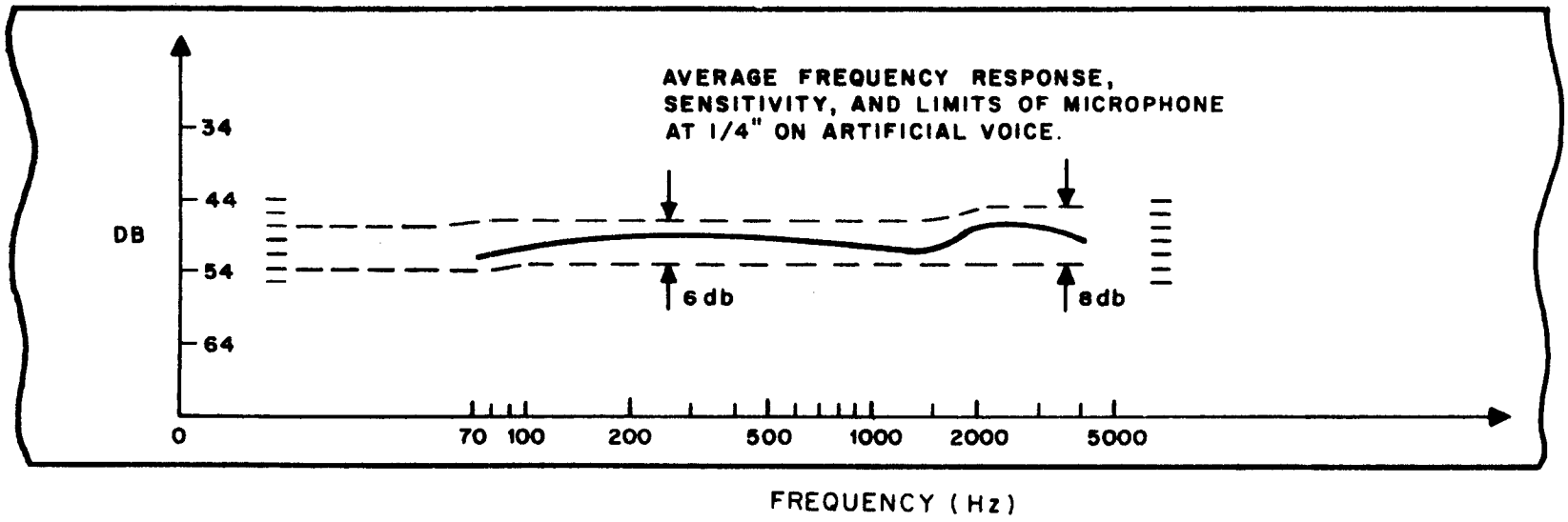


Figure B-1. Frequency Response of Noise-Canceling Microphone

APPENDIX C

A FAMILY OF VOLTAGE TUNABLE FILTERS

C.1 INTRODUCTION

Designed for the audio frequency range but applicable to several hundred kilohertz, the family of voltage tunable filters consists of a low-pass filter, a high-pass filter, and a bandpass filter. Stable Q's up to 100 are practical. Since no inductors are used these filters are adaptable to microminiature construction techniques. Filter frequency is linearly related to control voltage over a 20:1 range.

C.2 VOLTAGE TUNING

The variable element in all three types of tunable filters described here is a resistor in the form of a field-effect transistor. As described in the literature,[1] an F.E.T. can be used as a voltage-variable resistor if the voltage impressed between the source and drain is kept small. Under this condition the source-drain resistance of some types of commercial F.E.T.'s are closely approximated by:

$$R_{SD} = \frac{R_o}{1 - K_1 V_{gs}}$$

where:

R_o = resistance at zero gate bias

K_1 = a constant dependent upon the FET type

V_{gs} = gate to source voltage

Fortunately this inverse relationship between gate voltage and source-drain resistance fits in with common active-filter designs which

1. Sherwin, James S., "FET's as Voltage Controlled Resistors," Solid State Design, August 1965; for further discussion.

are electrically equivalent to the familiar single LC low-pass, high-pass, and bandpass filter.

C.3 LOW-PASS FILTER

The basic active low-pass filter is shown in Figure C-1.

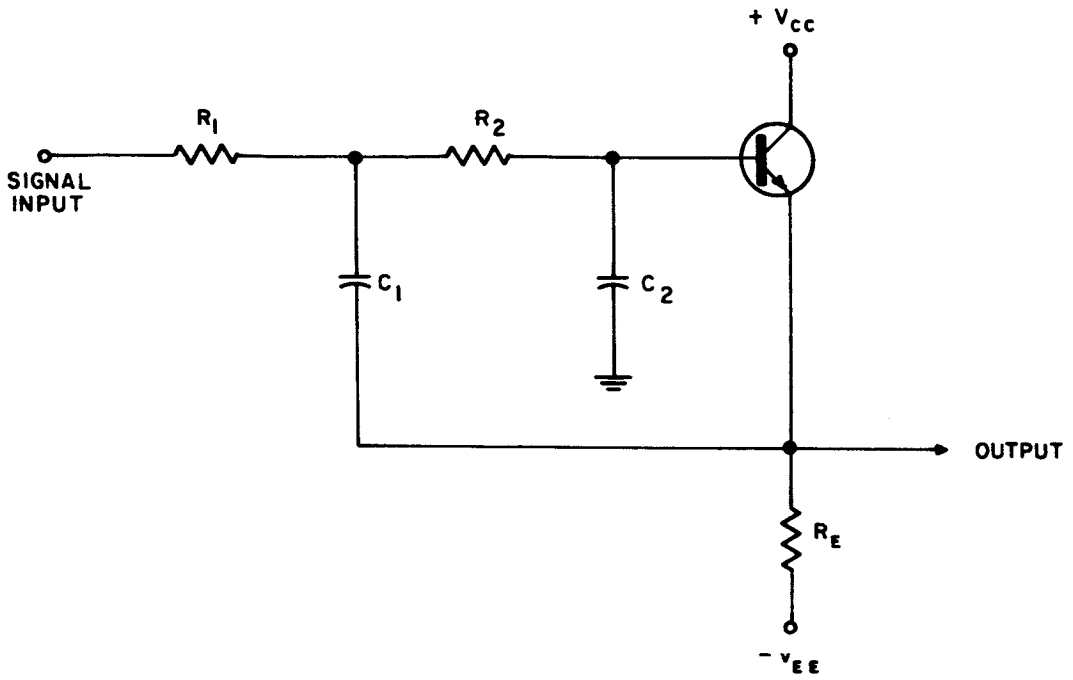


Figure C-1. Basic Active Low-Pass Filter

This filter can have a peaked or unpeaked response depending upon the damping factor, as shown in Figure C-2.

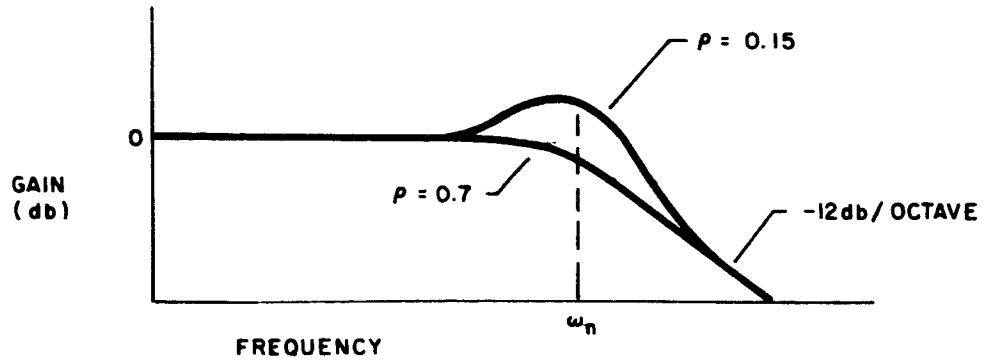


Figure C-2. Low-Pass Filter Response

In Figure C-3, the field-effect transistors have been substituted for R_1 and R_2 to give an active voltage-tunable filter (AVTF).

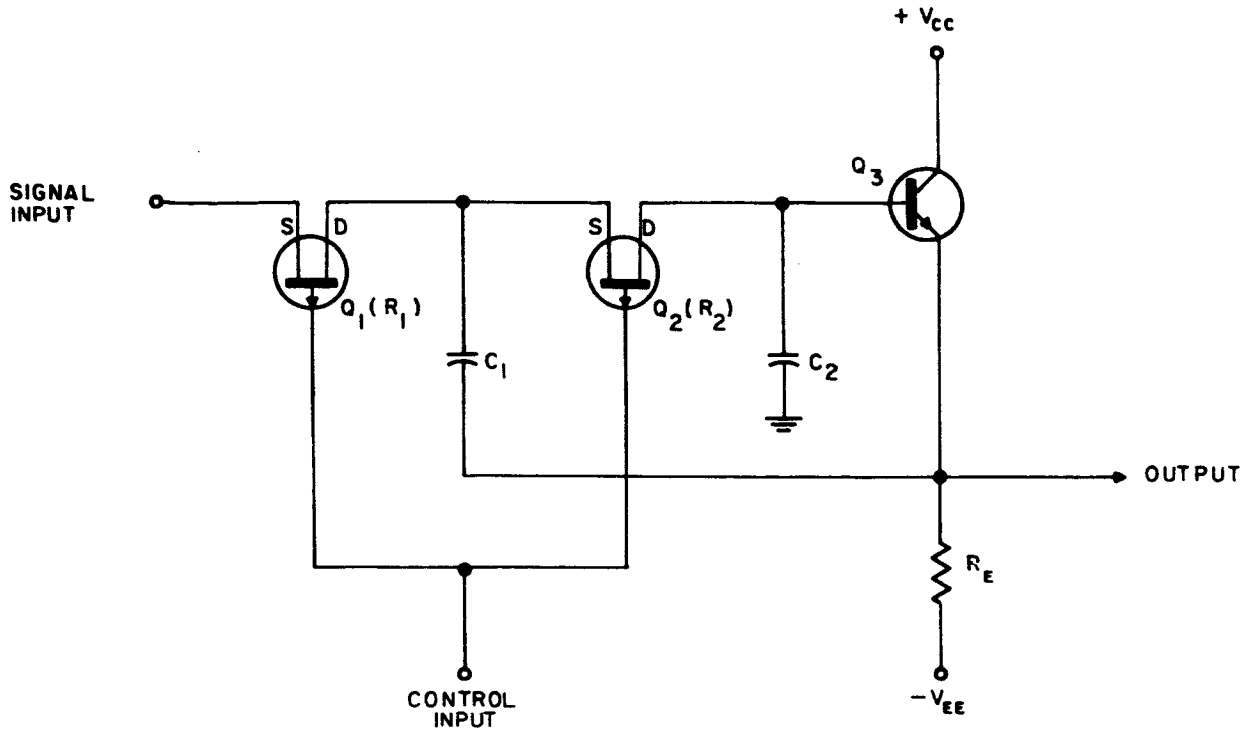


Figure C-3. Active Voltage-Tuned Filter (AVTF) Low-Pass Configuration

C.3.1 Low-Pass Design Equations

R_{IN} (input impedance) of emitter follower > 10 X_{C_2} (1)

R_{OUT} (output impedance) of emitter follower < 0.1 X_{C_1}

$$R_1 = R_2 \quad (2)$$

$$W_n = \frac{1}{RMC_2} = \text{resonant or corner frequency} \quad (3)$$

$$M^2 = \frac{C_1}{C_2}$$

$$\rho = \frac{1}{M} + \frac{M}{2}\left(\frac{1}{1 - K_2}\right) = \text{damping factor} \quad (4)$$

K_2 = voltage gain of emitter-follower (5)

By substituting the FET equation into the corner frequency Equation (3) for the low-pass filter the corner frequency is seen to be a linear function of the FET gate control voltage:

$$W_n = \frac{1 - K_1 V_{gs}}{R_o M C_2}$$

C.4 HIGH-PASS FILTER

By substituting FET's for resistors in a high-pass active filter a voltage-tunable high-pass filter can be built as shown in Figure C-4.

C.4.1 High-Pass Design Equations

Emitter follower: $R_{IN} > 10R_2$; $R_{OUT} < 0.1R_1$ (6)

$$C_1 = C_2 \quad (7)$$

$$W_n = \frac{1}{C\lambda R_1} = \text{corner frequency} \quad (8)$$

$$\rho = \frac{1}{\lambda} + \frac{\lambda}{2}(1 - K_2) = \text{damping factor} \quad (9)$$

$$\lambda^2 = \frac{R_2}{R_1} \quad (10)$$

$$K_2 = \text{gain of emitter follower} \quad (11)$$

As a practical consideration the FET gates are operated at the same control voltage making $R_1 = R_2$. This constrains the filter to a damping factor of 1. This restraint can be loosened somewhat by assuming $C_1 \neq C_2$. Doing this forces a return to some of the more basic equations of active filter design as developed by Davis.[1]

1. Davis, Steven J., "A Simple and Optimum Design of Active Filters," Philco internal report, October 1965.

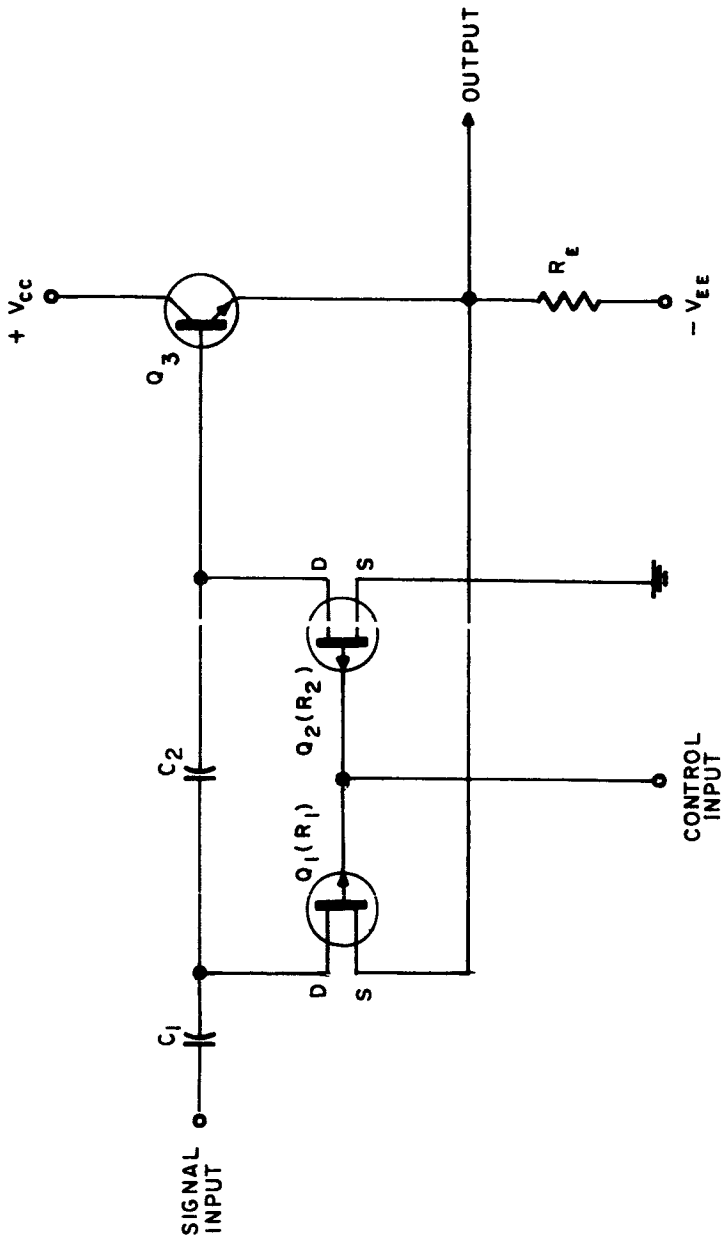


Figure C-4. High-Pass AVTF

C.4.2 High-Pass Exact Design

$$\text{Emitter follower: } R_{IN} > 10R_2;\ R_{OUT} < 0.1R_1 \tag{12}$$

$$W_n^2 = \frac{1}{R_1 R_2 C_1 C_2} = \text{corner frequency} \tag{13}$$

$$\rho = \frac{1+M^2}{2\lambda M} + \frac{\lambda}{2M}(1-K_2) = \text{damping factor} \tag{14}$$

$$M^2 = \frac{C_1}{C_2} \tag{15}$$

$$\lambda^2 = \frac{R_2}{R_1} \tag{16}$$

$$K_2 = \text{gain of emitter follower} \tag{17}$$

Even after going to this length, however, we can see from the design equations that ρ can only have values of one or greater.

As for the low-pass filter, corner frequency is linearly related to control voltage.

C.5 BANDPASS FILTER

The bandpass active voltage-tuned filter is basically quite different from the low-pass and high-pass filters. A simplified fixed-tuned version, shown in Figure C-5, consists of two cascade stages of amplification with overall gain approaching two. A third emitter-follower stage provides a low output impedance. The frequency-selective networks are a low-pass filter comprised of R_1 and C_1 and a high-pass filter formed by R_2 and C_2.

The products R_1C_1 and R_2C_2 are equal, since this results in the lowest required gain for a given damping factor. Capacitor C_3 and resistor R_3 provide ac coupling and dc return to the base of Q_1.

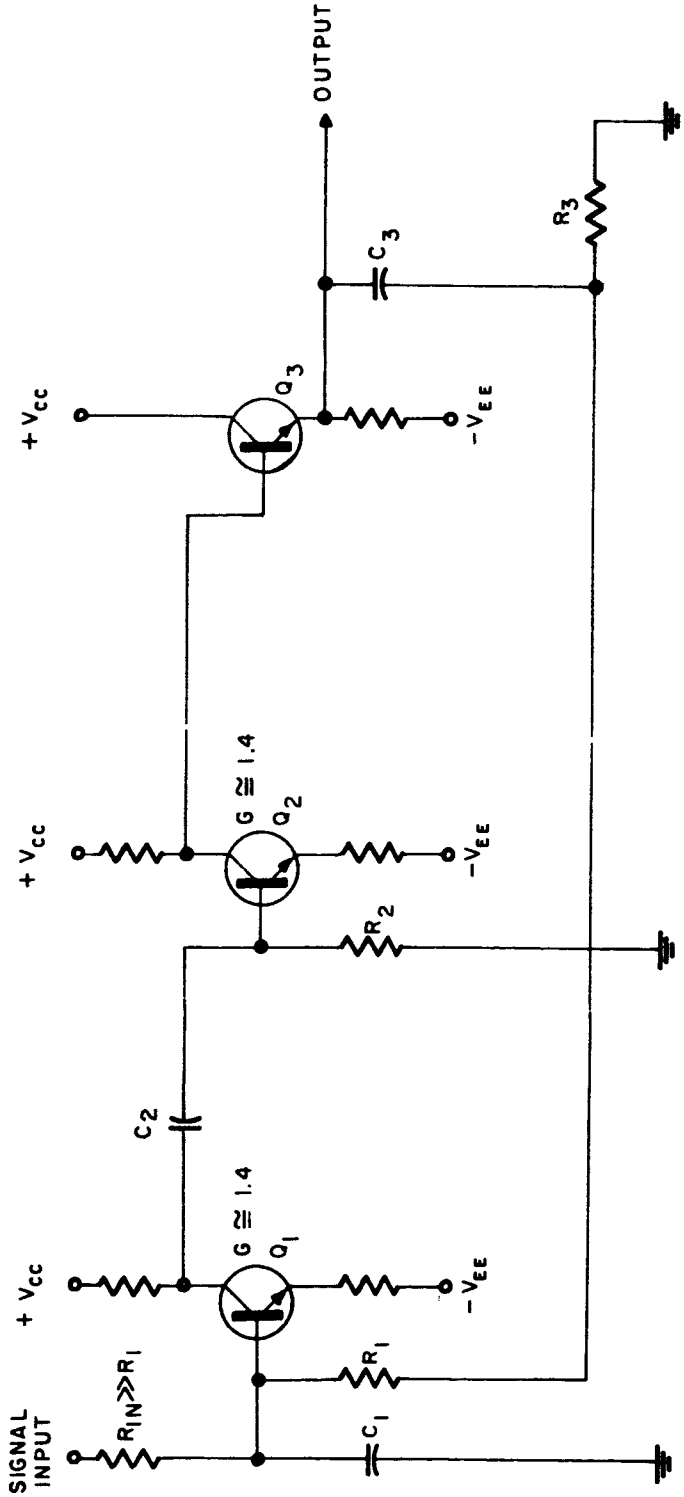


Figure C-5. Basic Active Bandpass Filter

As in the high-pass and low-pass AVTF's, the bandpass AVTF is obtained by replacing the frequency-determining resistors with field-effect transistors. Figure C-6 shows the voltage tunable bandpass filter.

C.5.1 Bandpass Design Equations

$$R_1 = R_2, \; C_1 = C_2 \tag{18}$$

$$W_n = \frac{1}{RC} = \text{resonant frequency} \tag{19}$$

$$Q = \frac{1}{1-0.25G_1^2 G_2^2} = \text{quality factor} \tag{20}$$

G_1 and G_2 are the voltage gains of Q_1 and Q_2 respectively (21)

$$R_{IN}\, Q_1 > 10 X_{C_1}, \; R_{IN}\, Q_2 > 10 X_{C_2} \tag{22}$$

Quality factor (Q) is used here rather than ρ, since Q is more commonly used with bandpass filters.

Again, resonant frequency is a linear function of control voltage.

C.6 TUNING RATE

The rate at which corner or resonant frequency can be changed is dependent upon the rate at which the FET drain-source resistance can be changed. High-frequency parameters for FET's are not usually available; however, the tuning rate for an audio-frequency filter has been measured to be in excess of 4×10^8 hertz per second.

C.7 PRACTICAL CONSIDERATIONS

The field-effect transistors used in the active voltage-tuned filters are units matched for operation over a 20:1 frequency range to the following specifications:

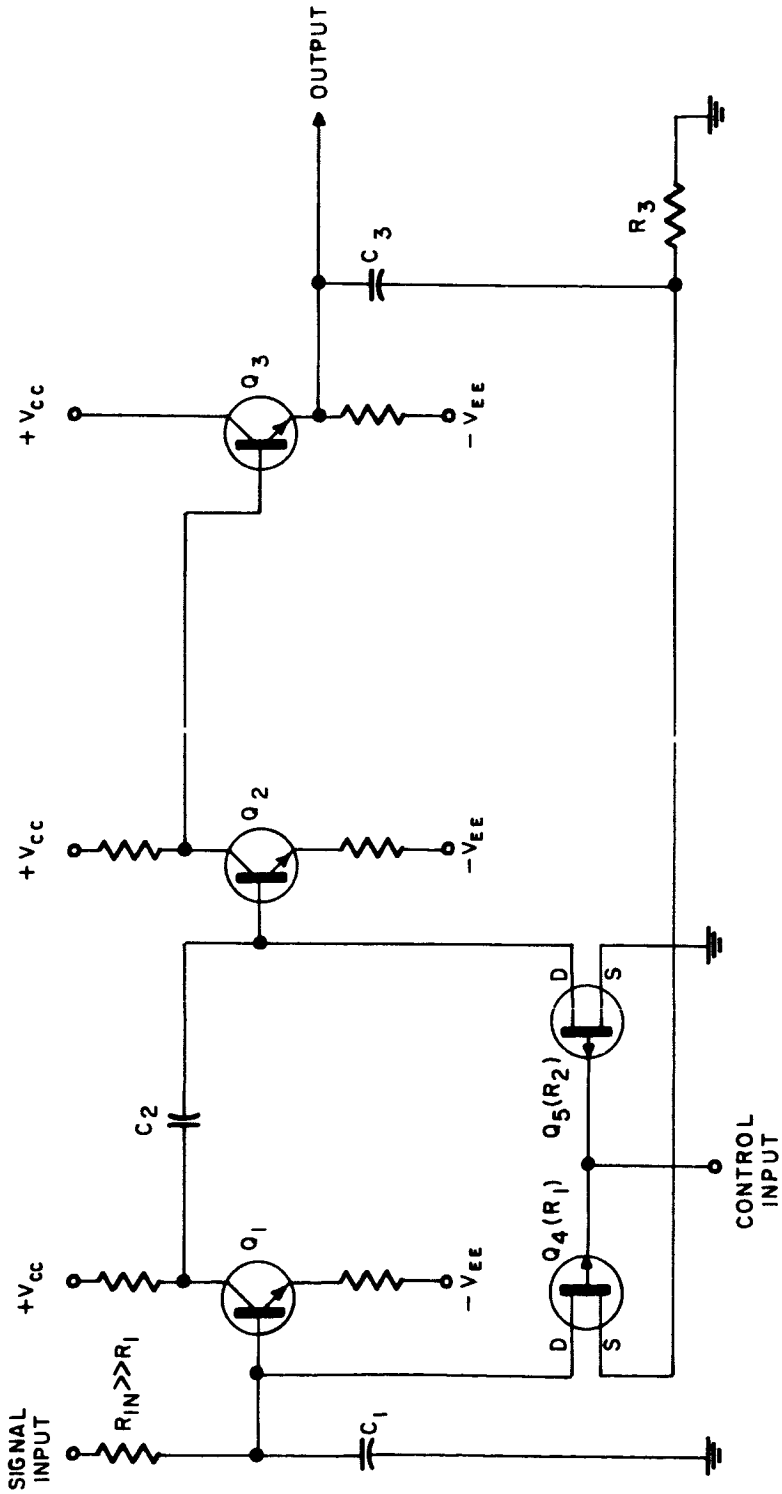


Figure C-6. AVTF Bandpass

a. Resistance at zero gate voltage (R_o): matched within 5 percent

b. Resistance at $4R_o$: matched within 5 percent

c. Resistance at $20R_o$: matched within 5 percent

Pairs matched to these specifications are available from Siliconix, Inc., as part number SU700. They are not expensive compared to voltage-variable inductors.

Maximum peak-to-peak output signal swing is limited by the FET's to about 1 volt. At greater signal swing the output signal begins to be distorted.

When very low damping factor is required it becomes difficult to satisfy the requirements of high amplifier input impedance and output impedance. Also, slight amplifier gain variations strongly affect damping when damping is low, as can be seen from the design equations.

To overcome the problems arising in the low-damping case a high-gain operational amplifier connected for the appropriate gain has been found to be a very useful substitute for the amplifying transistors.

C.8 SUMMARY

The filter techniques described here offer many advantages over more conventional voltage-variable (more accurately current-variable) inductors used in electronically tunable filters. Some of the advantages are:

a. Readily adapted to microminiature construction

b. High Q obtainable at audio frequencies

c. High permissible tuning rate

d. No feedthrough from the control input to output

e. Linear control of resonant or corner frequency with control voltage

f. Control signal need supply no power

g. Wide tuning range

h. Moderate cost.

APPENDIX D

BREADBOARD OPERATING INSTRUCTIONS AND CIRCUIT DIAGRAMS

D.1 INTRODUCTION

This appendix contains brief breadboard operating instructions, and circuit diagrams for the analog portion of the system.

D.2 BREADBOARD OPERATING INSTRUCTIONS

D.2.1 Setup

The analyzer and synthesizer both operate on plus and minus 12.0 volts. Power supply voltages should be initially set to 12.0 volts $\pm$ 50 mv with a differential voltmeter, such as Fluke model 801, and checked occasionally to maintain accuracy.

Using a dual-trace dc-coupled oscilloscope, the parameter tracking from multiplexer input to demultiplexer output should be checked. If alignment is required it is obtained by adjusting the demultiplexer dc level control and Finite Memory Filter (FMF) control.

To check alignment of the analog portions of the analyzer/synthesizer the ten vowel phonemes (see Figure 3-6) are spoken, one by one, into the analyzer, and the synthesized formant frequency measured on an oscilloscope. To change synthesized frequency the offset control in the analyzer output shapers is adjusted for optimum tracking for all ten vowel phonemes.

Once the analyzer/synthesizer and multiplexer/demultiplexer tracking is complete the balance between first and second formant is adjusted by the mixer control in the output amplifier to suit the listener.

D.2.2 Speaking Level

Connect a dual-trace dc-coupled oscilloscope to the demultiplexer amplitude parameter outputs. The speaking level should be adjusted so that the amplitudes do not limit. The speaker should speak slowly and distinctly and should minimize amplitude inflections in his voice.

D.3 CIRCUIT DIAGRAMS

Figures D-1 through D-9 pertain to the analyzer, D-10 is the link filter and is mounted on the analyzer chassis. Figures D-11 through D-16 pertain to the synthesizer. Figures D-17 and D-18 show the plus and minus 10-volt regulators used in both the analyzer and the synthesizer.

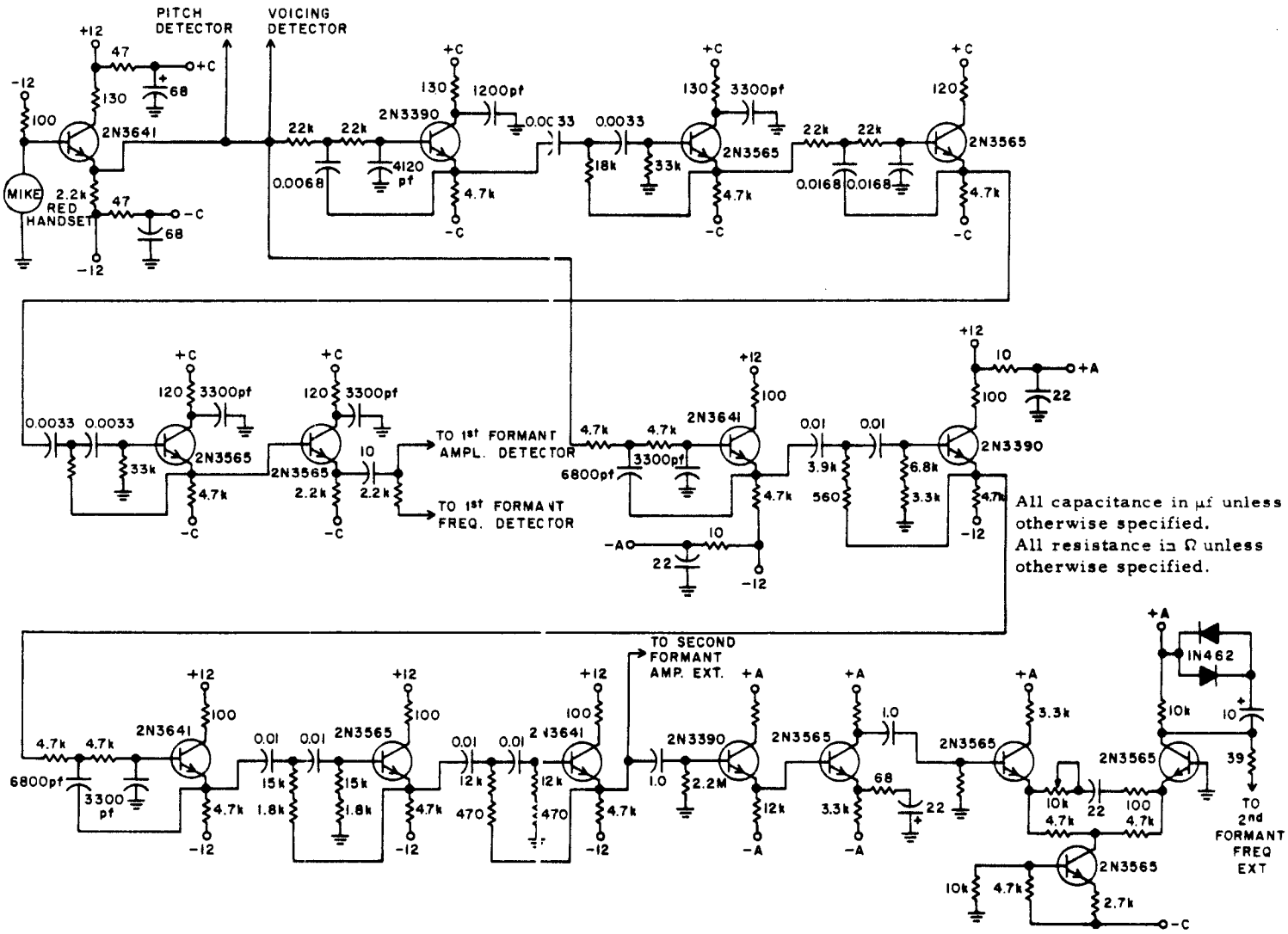


Figure D-1. First- and Second-Formant Input Bandpass Filters

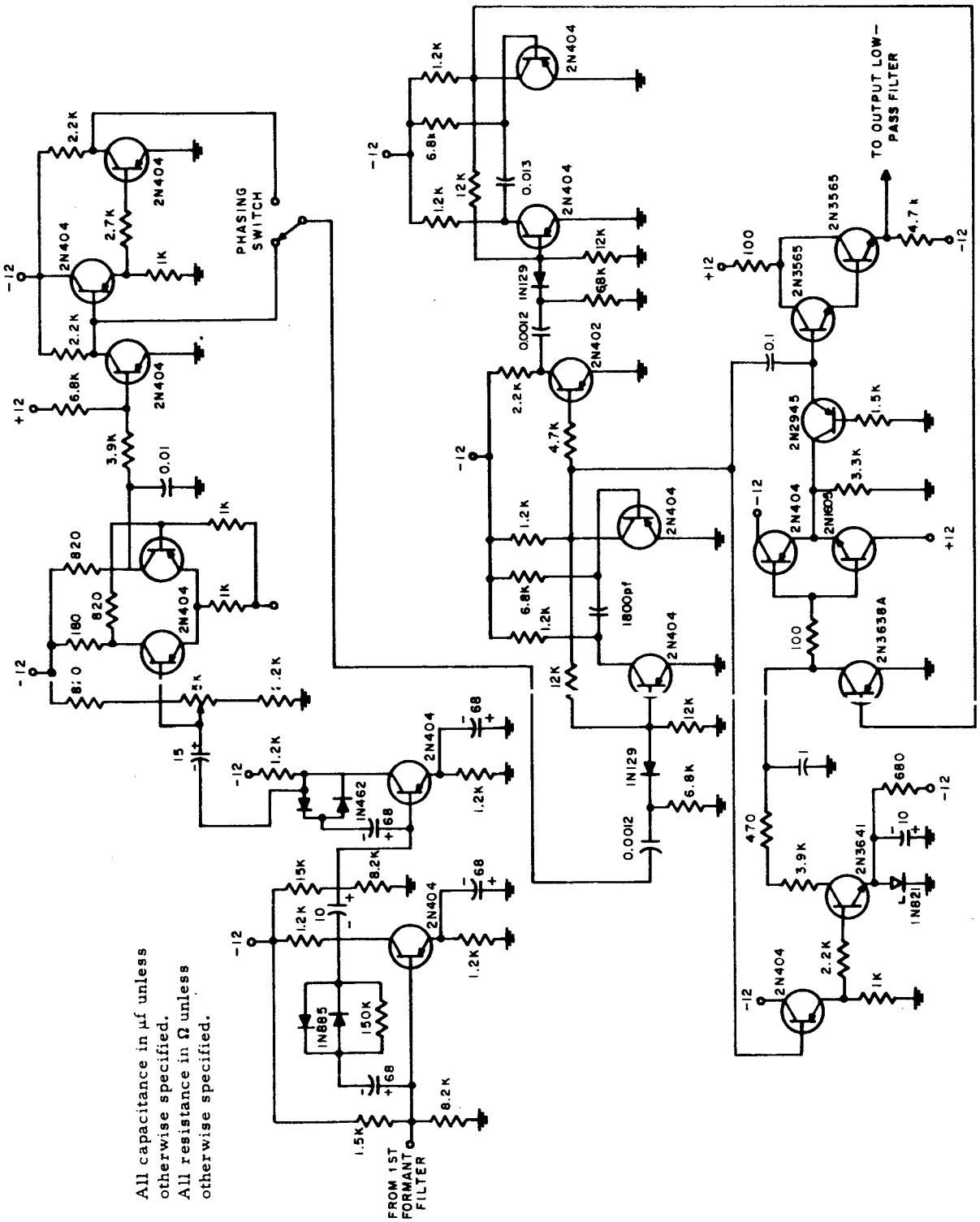


Figure D-2. First Formant Period Detector

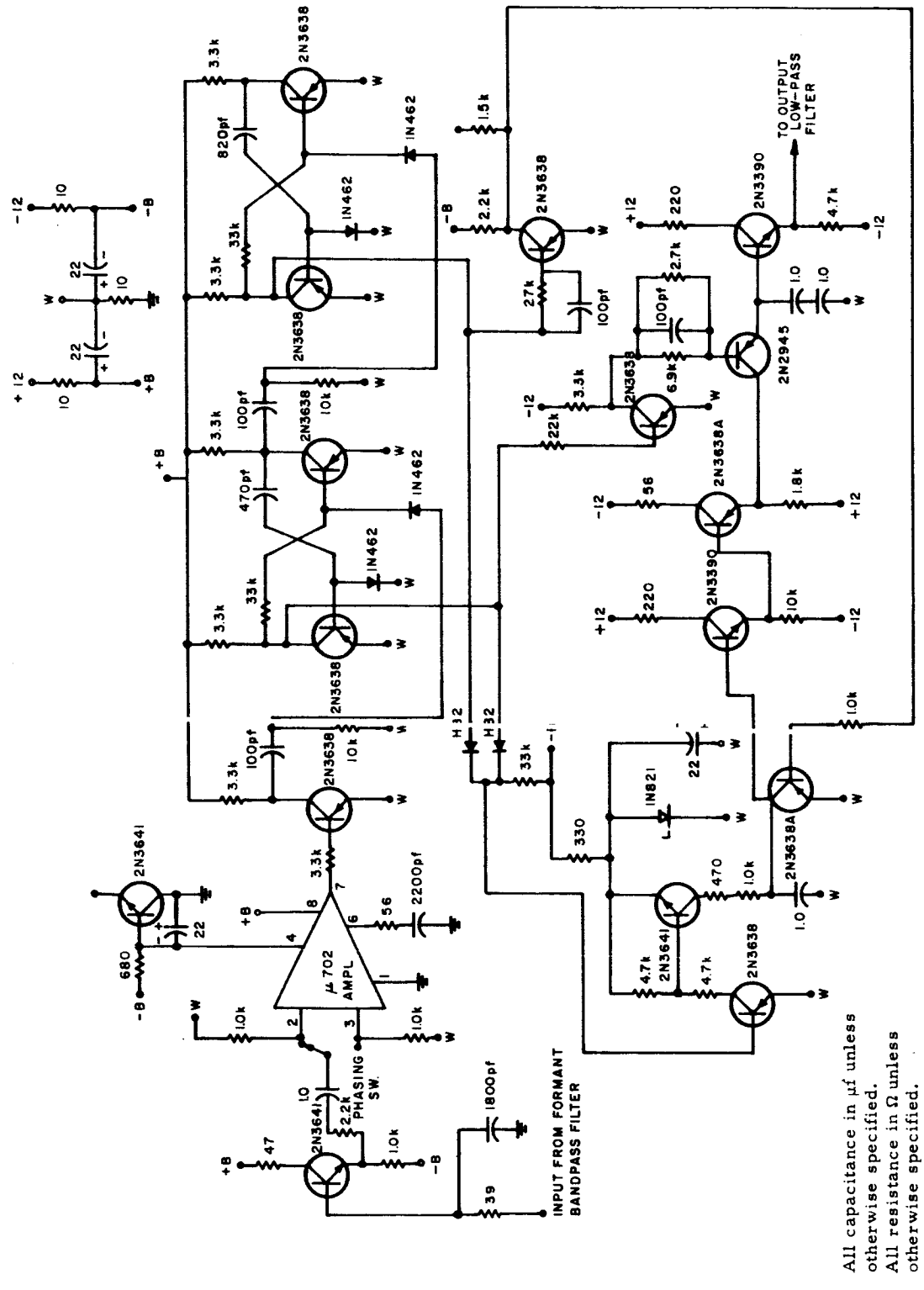


Figure D-3. Second Formant Period Detector

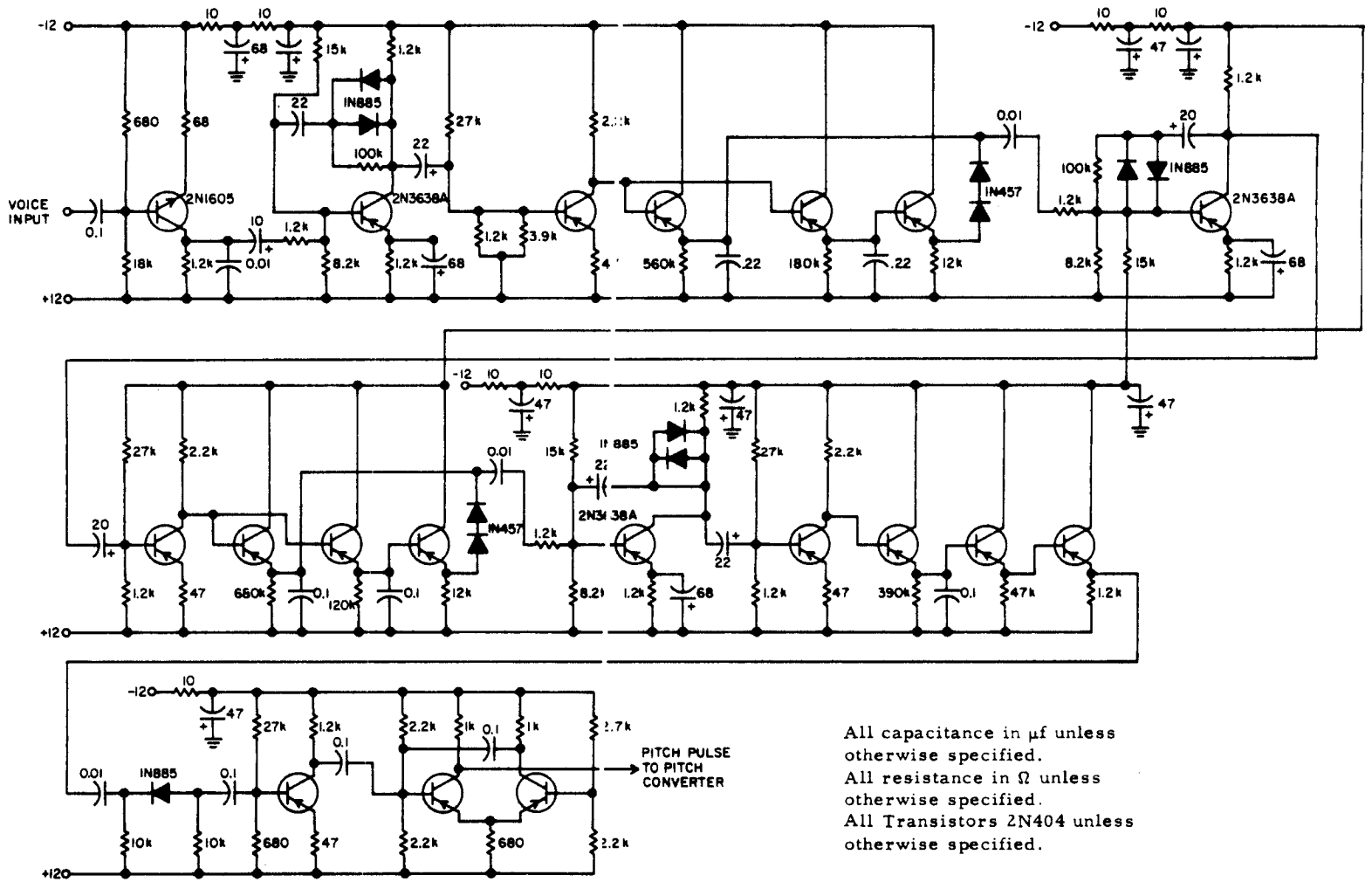


Figure D-4. Pitch Extractor

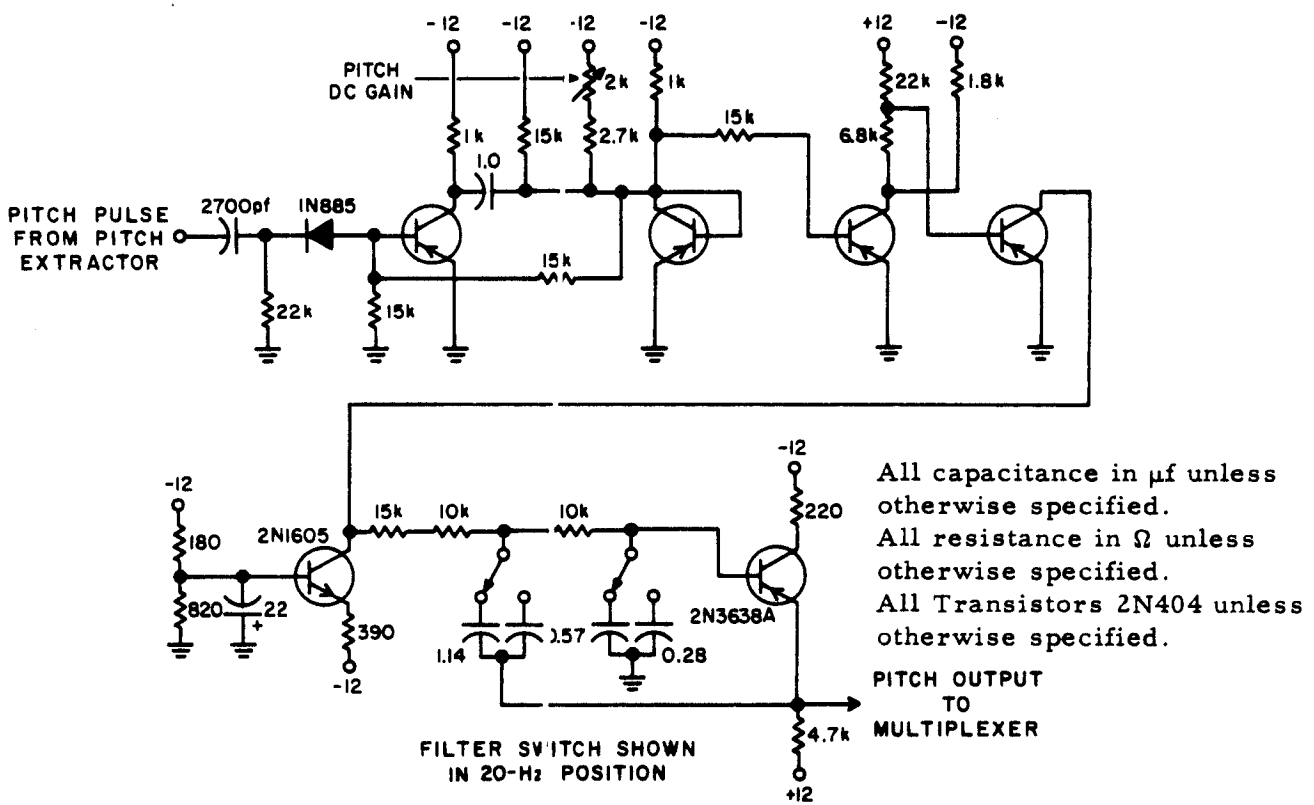
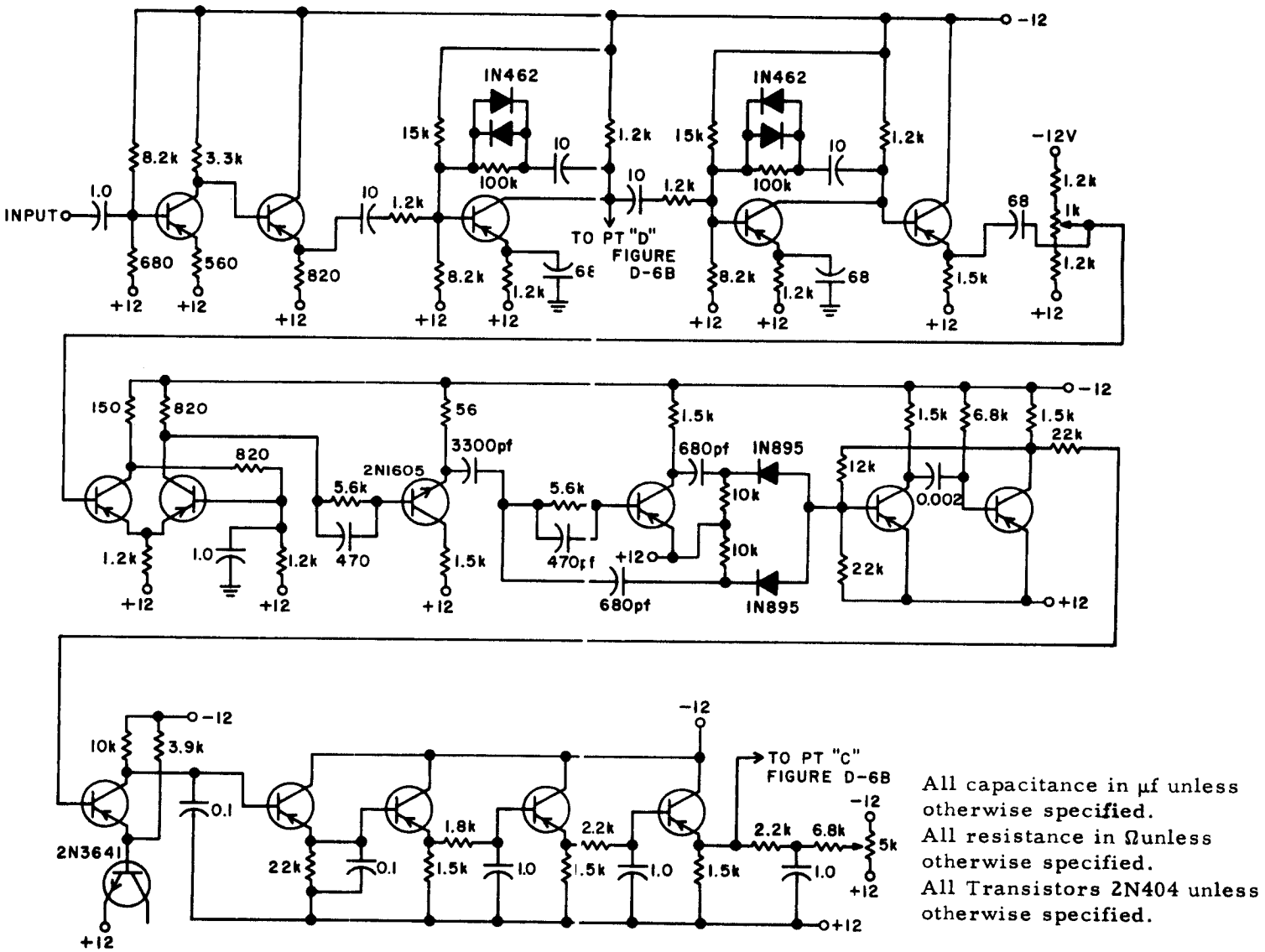


Figure D-5. Pitch-Pulse-to-dc Converter



All capacitance in μf unless otherwise specified.
All resistance in Ωunless otherwise specified.
All Transistors 2N404 unless otherwise specified.

Figure D-6A. Voicing (Part A)

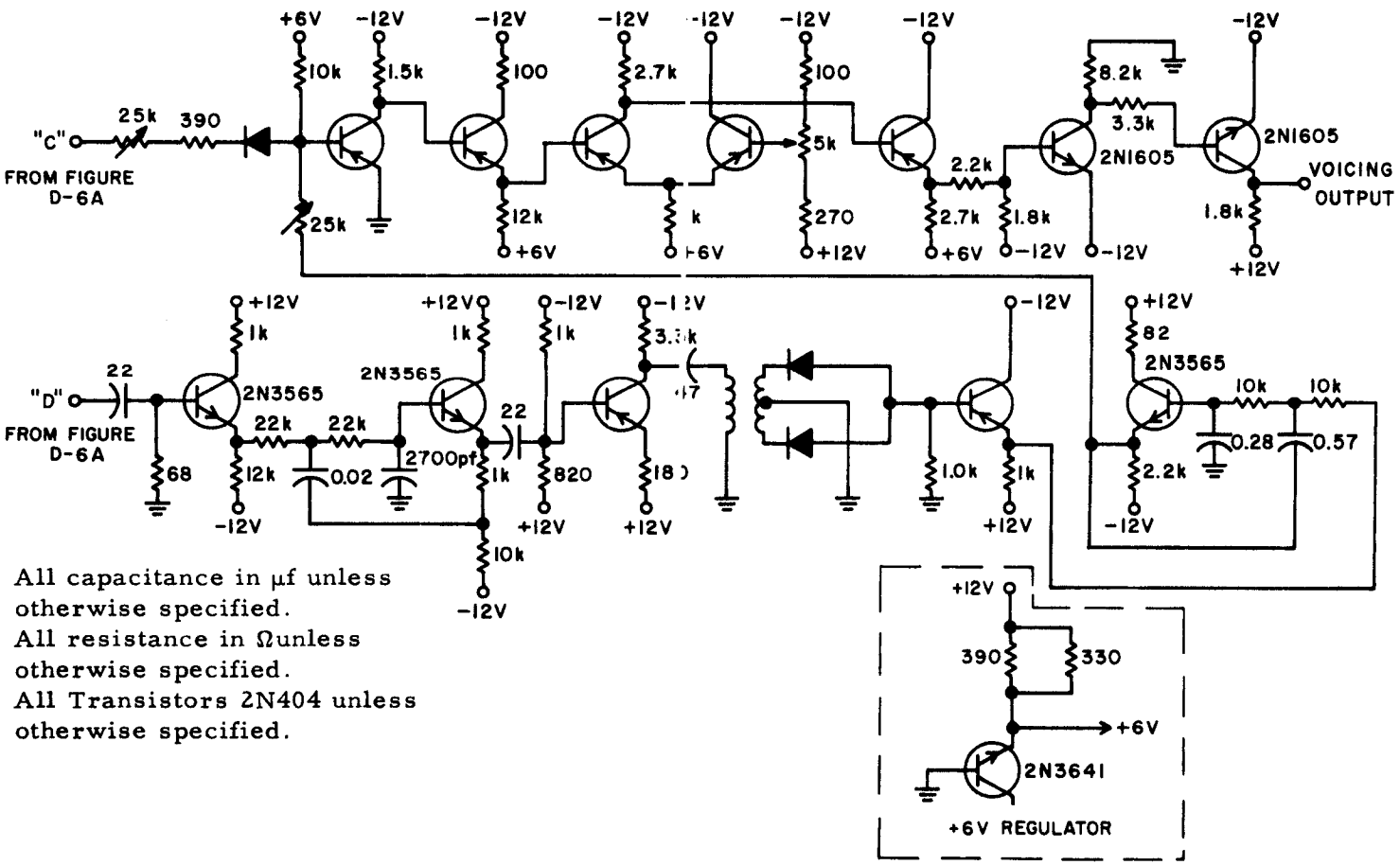
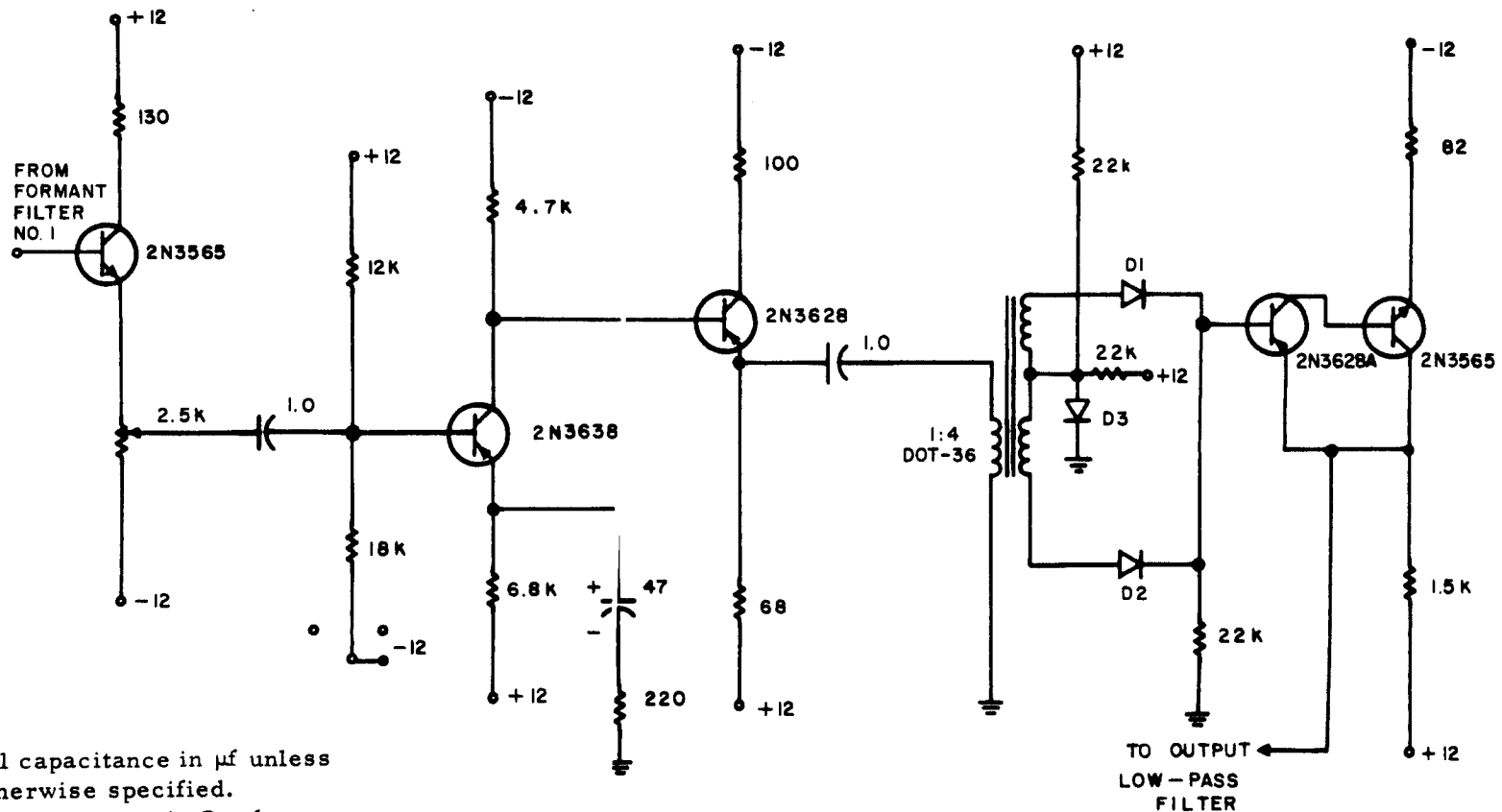
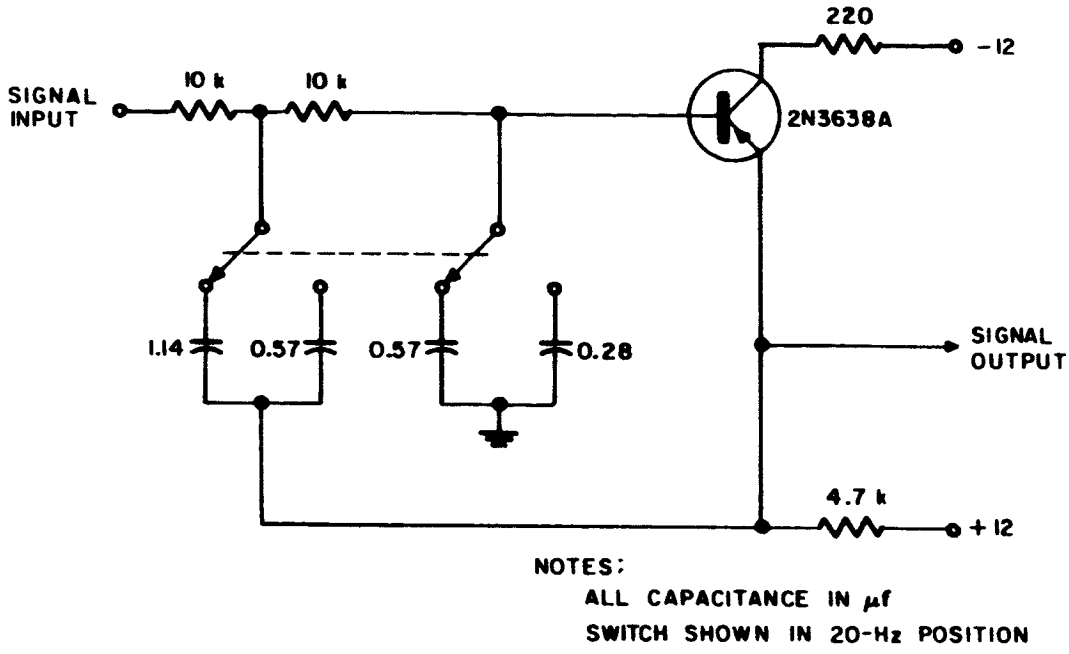


Figure D-6B. Voicing (Part B)

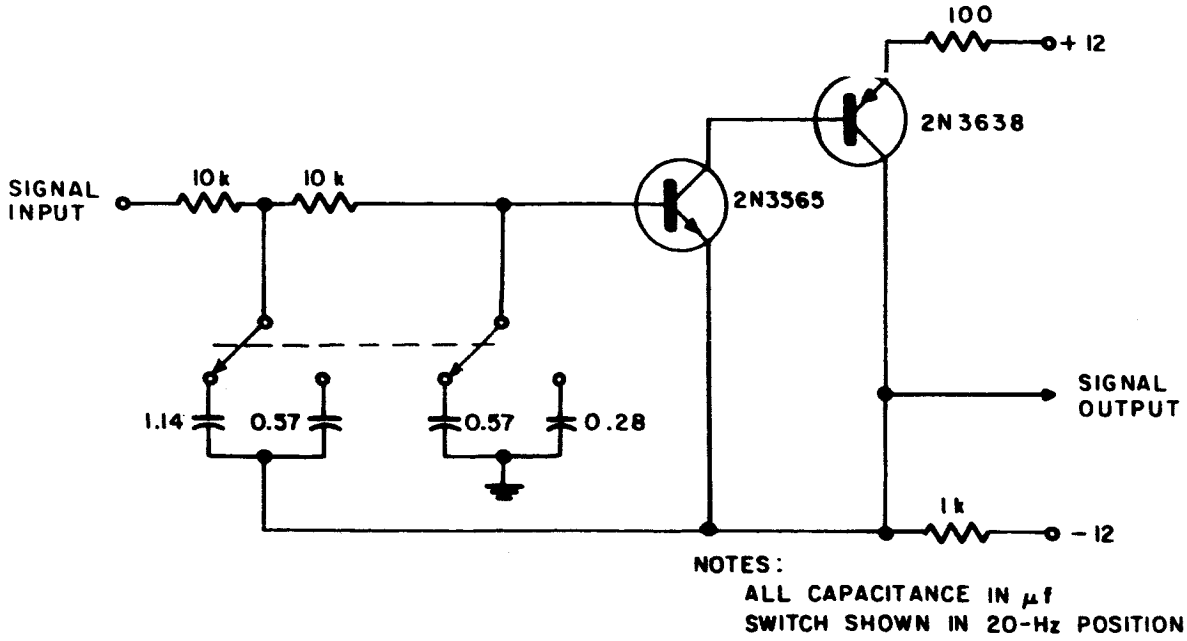


All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-7. Amplitude Detector

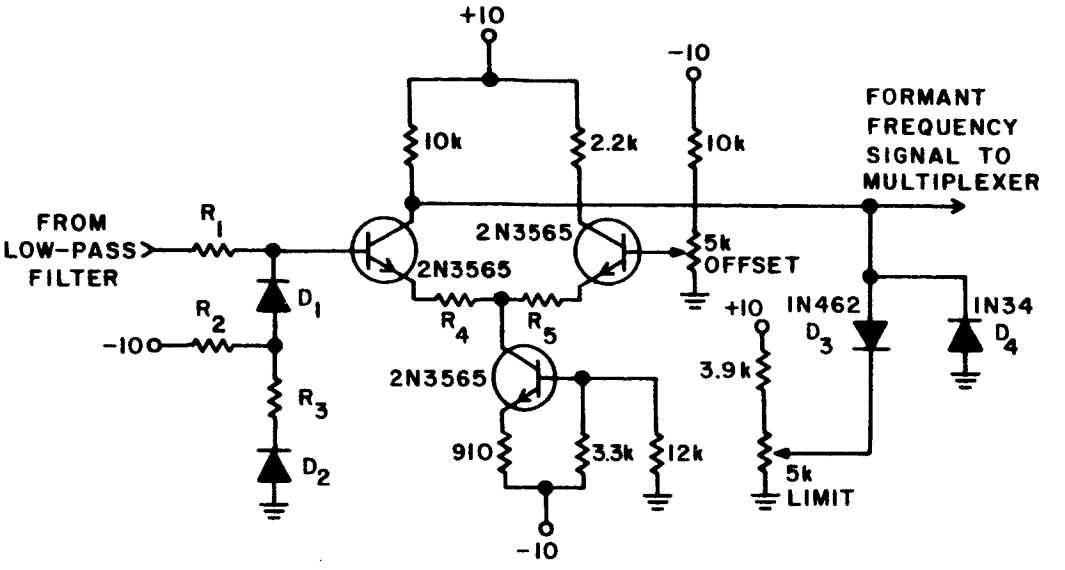


a. Low Z_{IN} Output Filter



b. High Z_{IN} Output Filter

Figure D-8. Output Filters



	First Formant	Second Formant
R_1	22K	22K
R_2	39K ‖ 470K	47K
R_3	2.2K + 3.3K	2.2K + 3.9K
R_4	3.9K ‖ 3.9K	3.9K
R_5	3.9K ‖ 3.9K	3.9K
D_1	1N462	1N462
D_2	————	1N457

All ± 5%

All resistance in Ω unless otherwise specified.

Figure D-9. Output Shaper

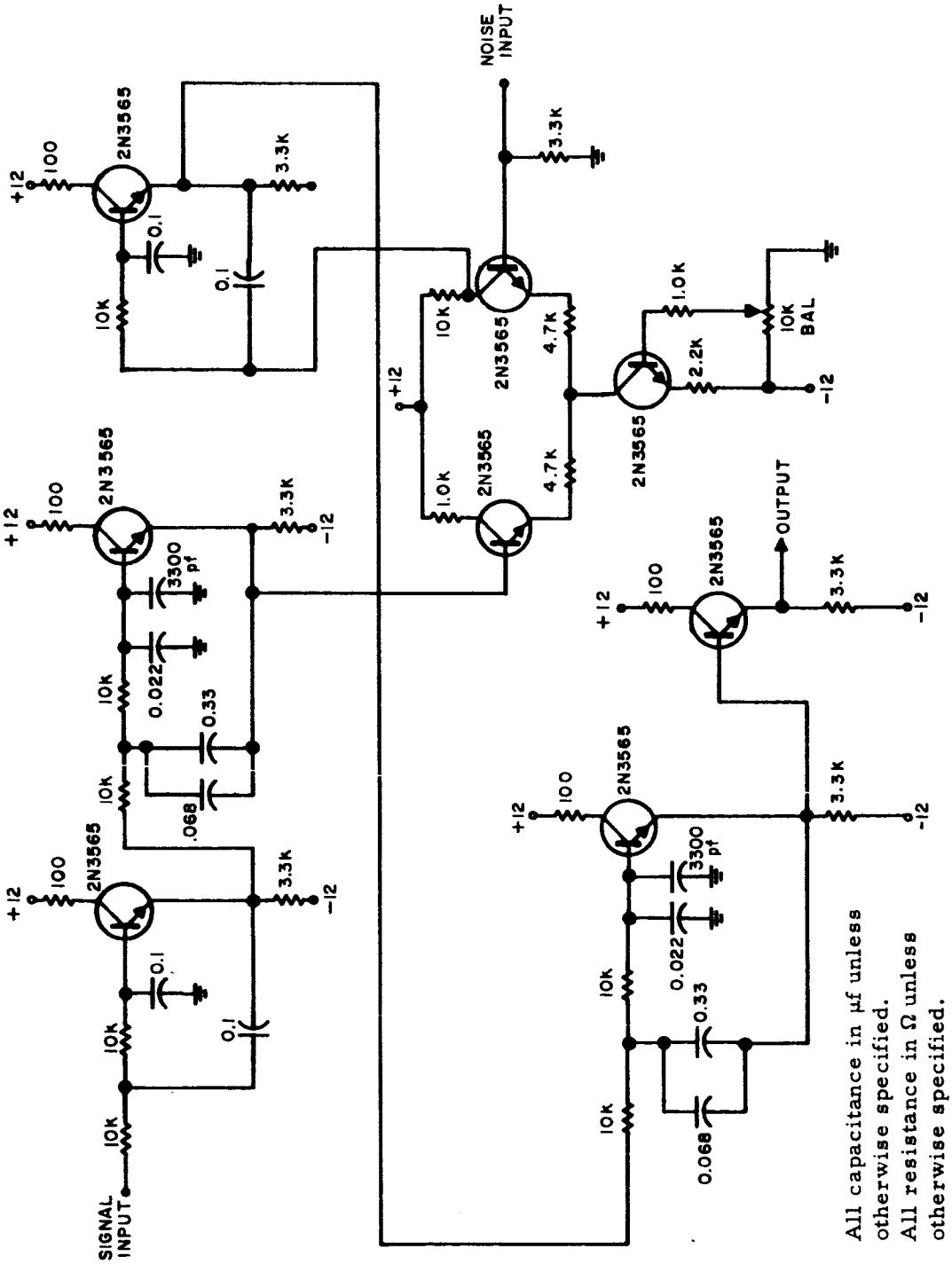


Figure D-10. Link-Simulation Filter

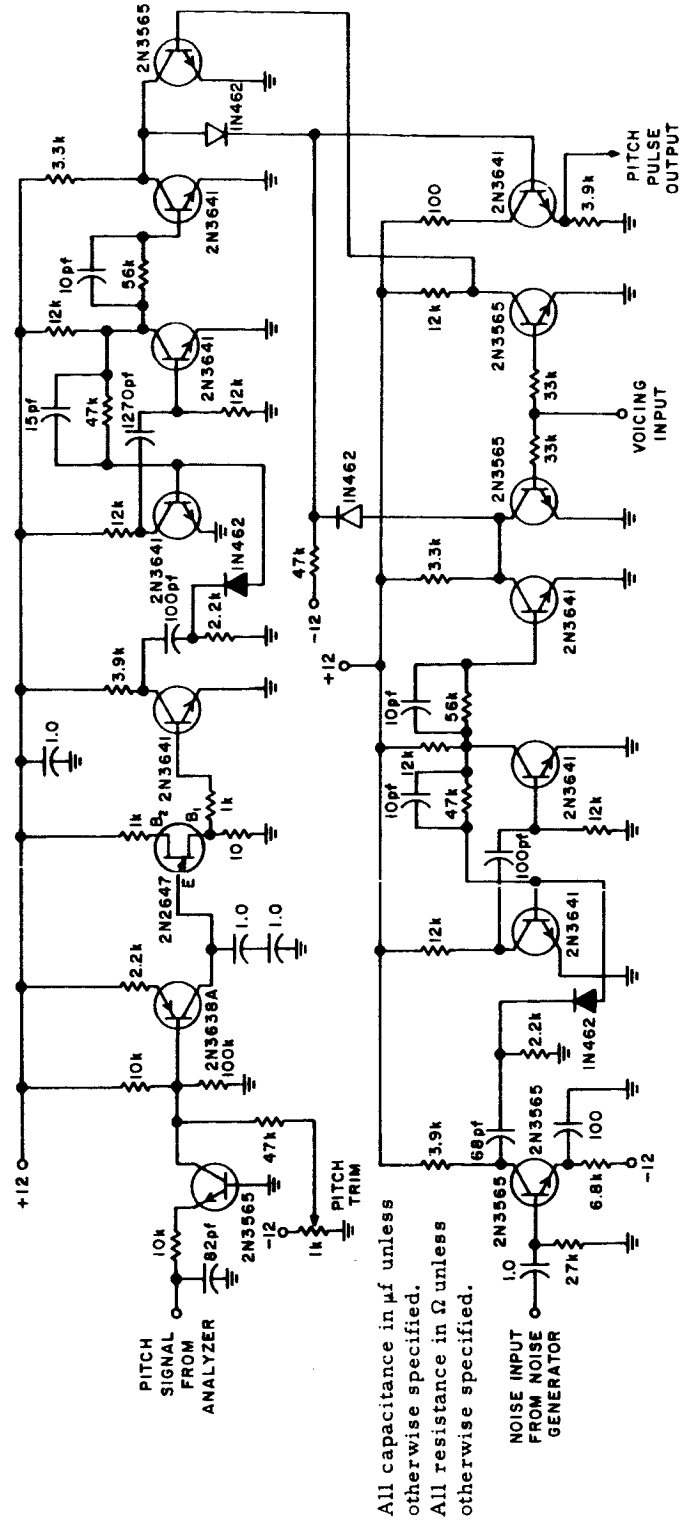
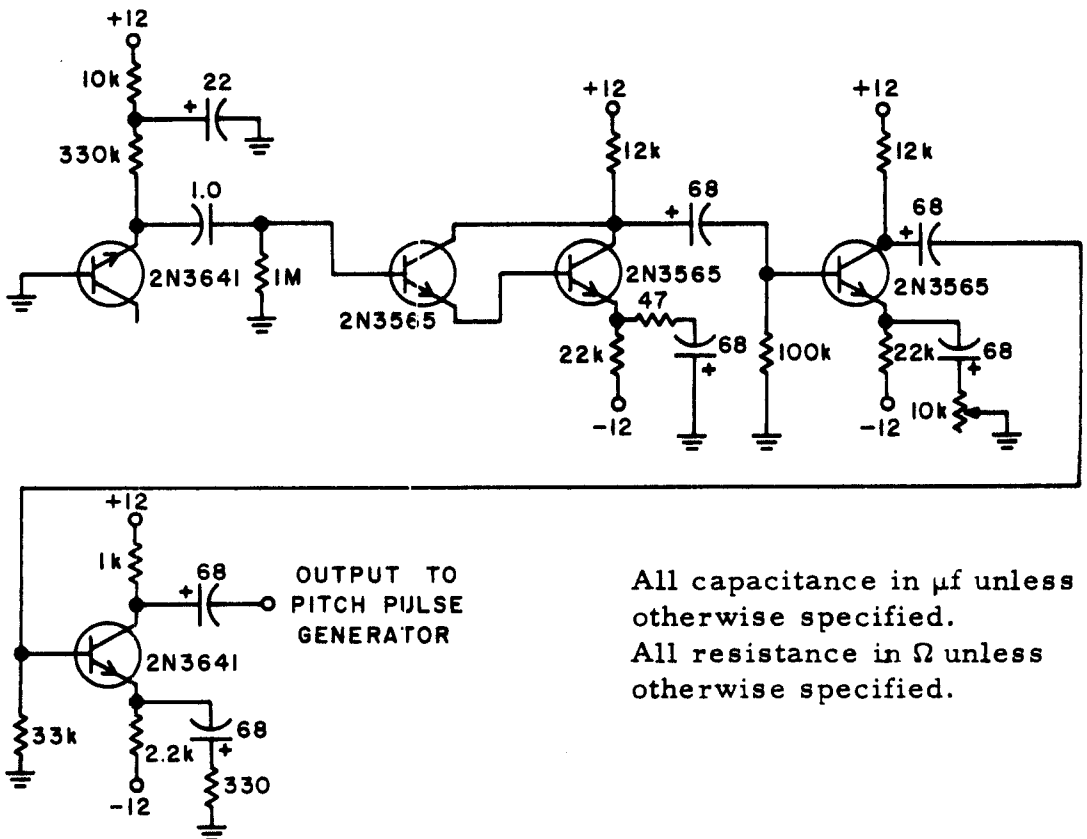


Figure D-11. Pitch Generator



All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-12. Noise Generator

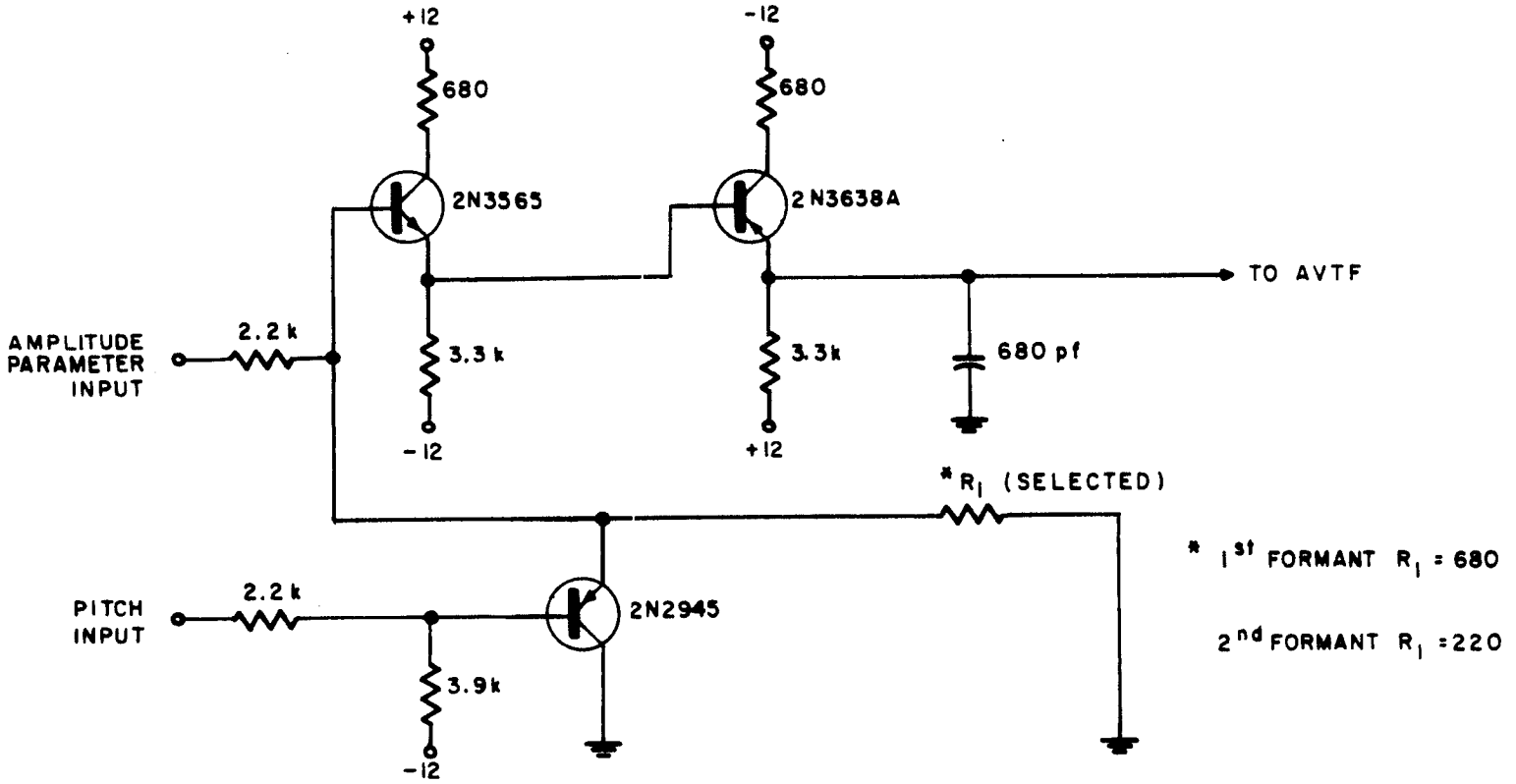
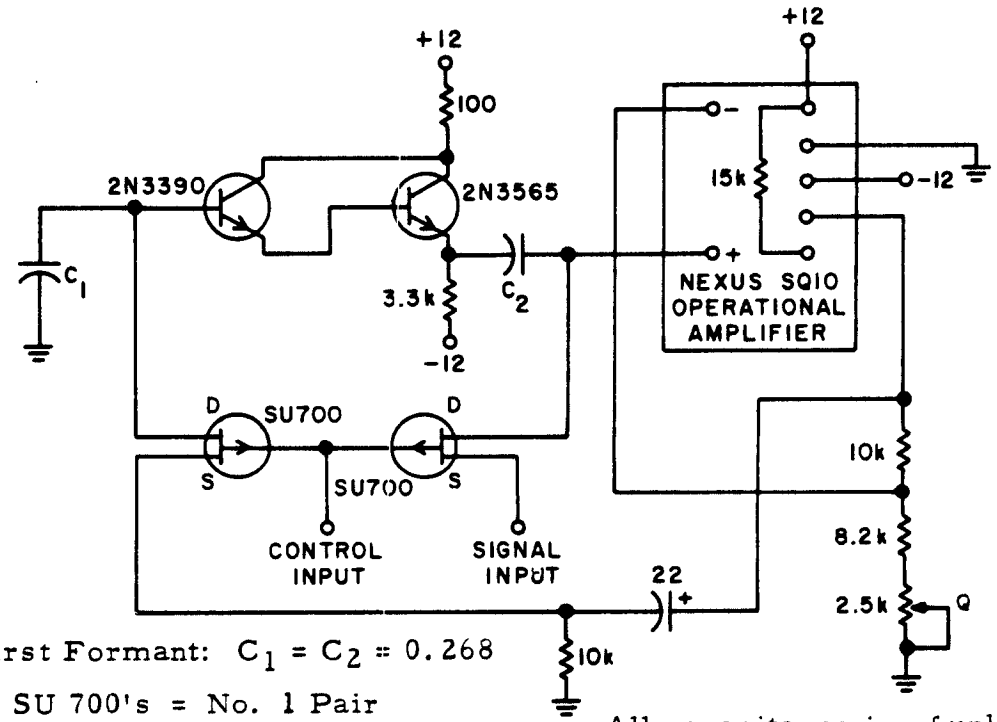


Figure D-13. Switch Modulator



First Formant: $C_1 = C_2 = 0.268$

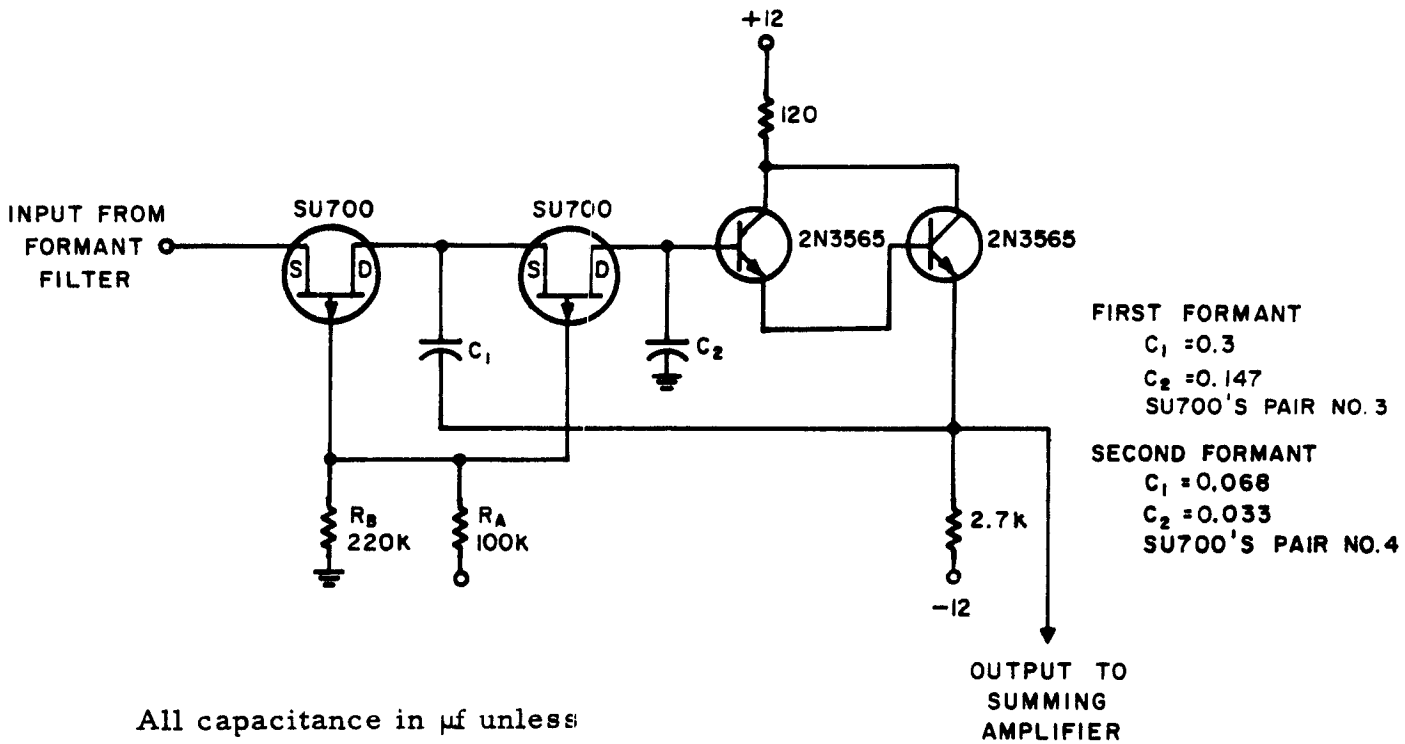
SU 700's = No. 1 Pair

Second Formant: $C_1 = C_2 = 0.0578$

SU 700's = No. 2 Pair

All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-14. Parallel-Resonant Active Voltage-Tuned Filter (A. V. T. F.)



All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-15. Output-Tracking A.V.T.F.

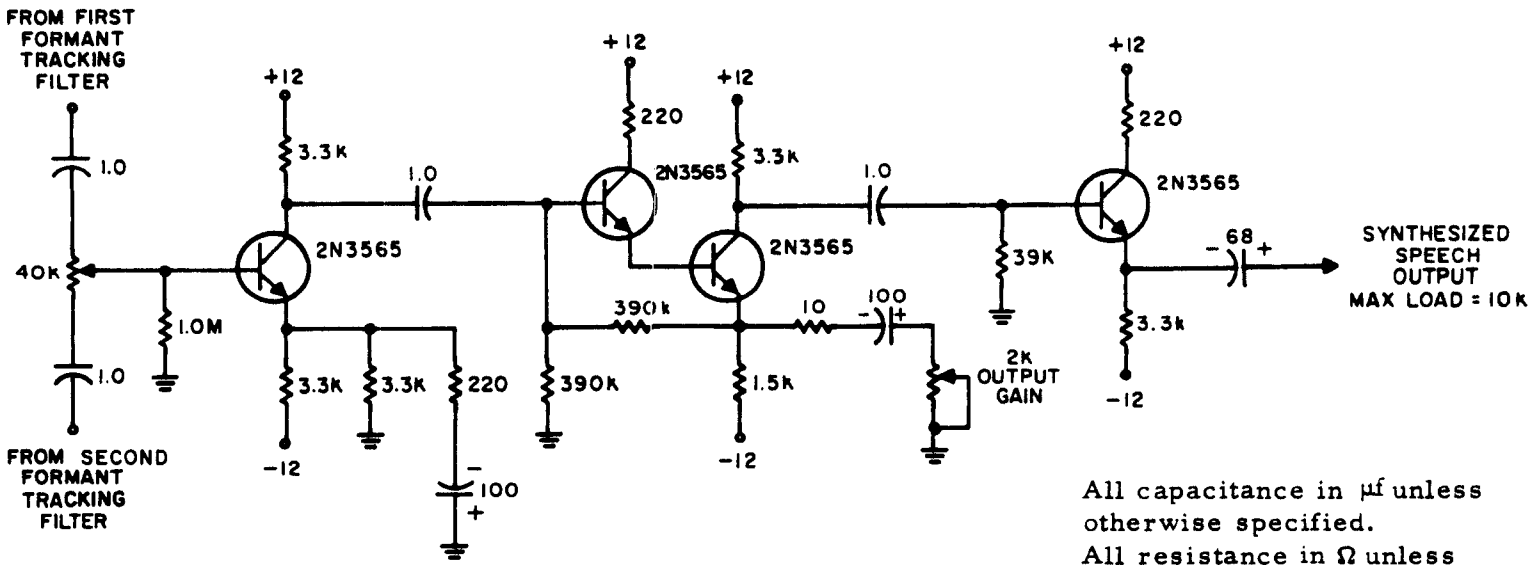
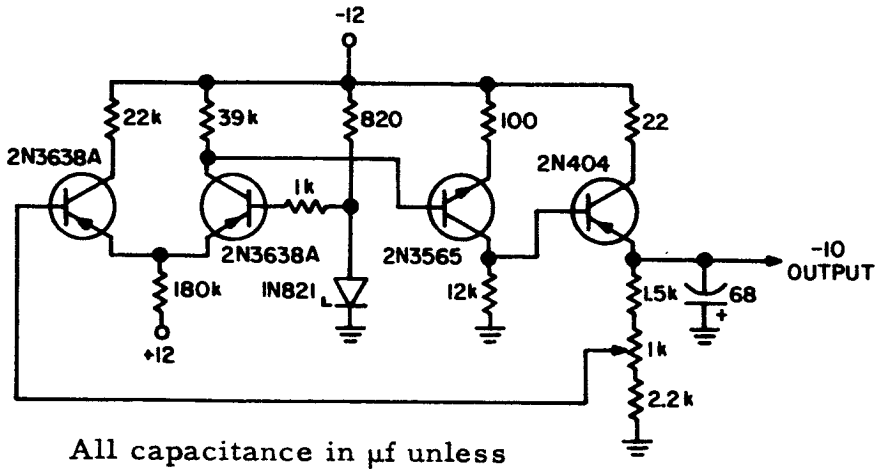
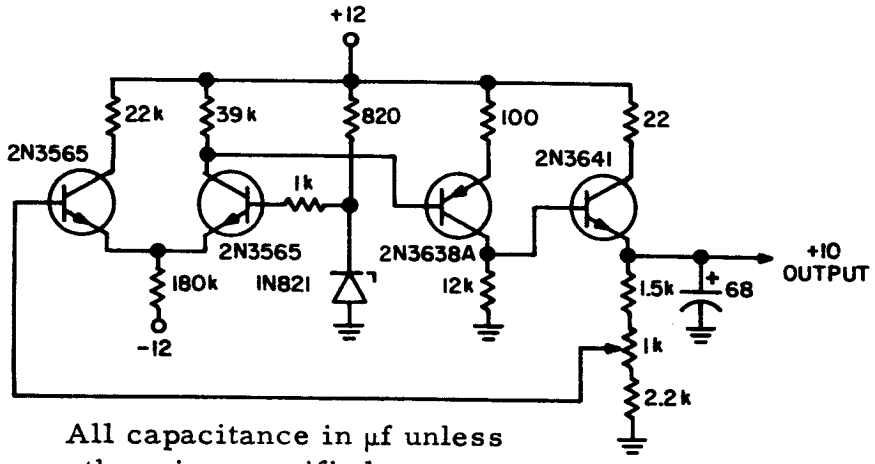


Figure D-16. Summing Amplifier



All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-17. -10 Volt Regulator



All capacitance in μf unless otherwise specified.
All resistance in Ω unless otherwise specified.

Figure D-18. +10 Volt Regulator