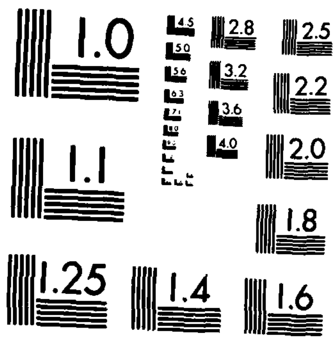
AD-A125 728 ADAPTIVE-BANDWIDTH AUDIO INPUT FOR THE NARROWBAND LINEAR PREDICTIVE CODER(U) NAVAL RESEARCH LAB WASHINGTON DC A.F THORNHILL ET AL. 18 MAR 83 1/1

UNCLASSIFIED NRL-MR-5036 F/G 17/2 NL

END

FILMED

DTIC


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NRL Memorandum Report 5036	2. GOVT ACCESSION NO. AD-A125728	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ADAPTIVE-BANDWIDTH AUDIO INPUT FOR THE NARROWBAND LINEAR PREDICTIVE CODER		5. TYPE OF REPORT & PERIOD COVERED Final report on one phase of a continuing NRL problem. 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) A.F. Thornhill and H.O. Murphy		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Research Laboratory Washington, DC 20375		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NAVELEX X0734CC; X1099CC; 75-0114-0-3
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Electronic Systems Command Washington, DC 20360		12. REPORT DATE March 18, 1983 13. NUMBER OF PAGES 20
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Adaptive filters
Voice intelligibility
Linear predictive coding

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Digital voice communication systems using linear predictive coders lose intelligibility through failure to reproduce certain speech sounds. Sibilants may have spectra falling entirely outside the input passband of the digital processor. An adaptive analog input filter allows sibilant energy to enter the processor. Although distortion is produced, intelligibility is improved.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

CONTENTS

ADAPTIVE-BANDWIDTH AUDIO INPUT FOR THE NARROWBAND LINEAR PREDICTIVE CODER

INTRODUCTION

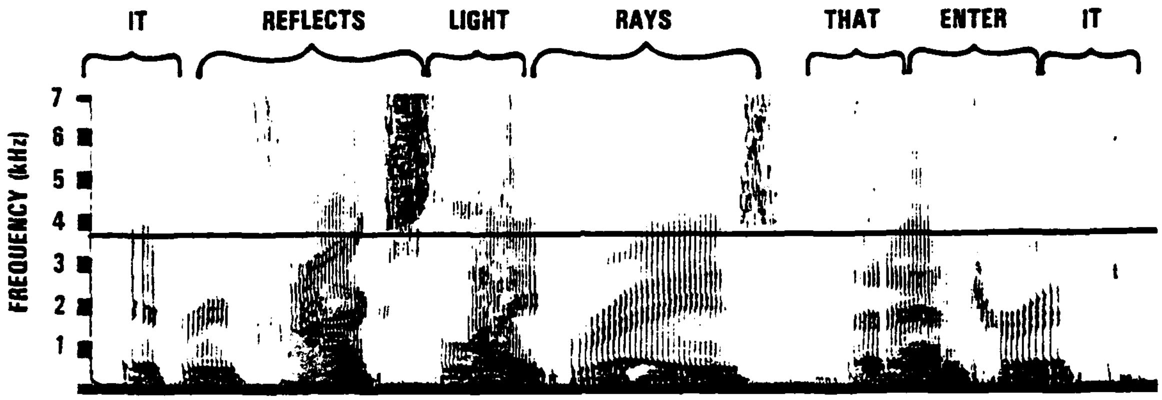
This report describes an analog method for controlling the audio input bandwidth preceding a narrowband linear predictive coder (LPC). The LPC is a digital processor which operates on sampled voice waveforms to produce a narrowband digital signal. The narrowband digital signal can be transmitted over voice-grade landlines or radio channels and reconstructed at the receiving end by a synthesis technique.

The analog voice waveform is sampled at an 8 kHz rate in the current narrowband LPC, so input frequencies higher than 4 kHz cannot be recognized following the sampler. A frequency, F, above 4 kHz will be interpreted as (8000 - F) Hz. To prevent interpretation errors caused by this characteristic of sampled data systems, the sampler is preceded by a low-pass filter (LPF) to remove frequencies higher than one-half the sampling frequency. The filter is also known as an anti-aliasing filter. Most filters have a finite slope of the stopband attenuation characteristic, so the corner (cut-off) frequency of the filter is usually set somewhat lower than one-half the sampling frequency. The corner frequency specified for the current narrowband LPC is 3.6 kHz, and it is not adjustable.

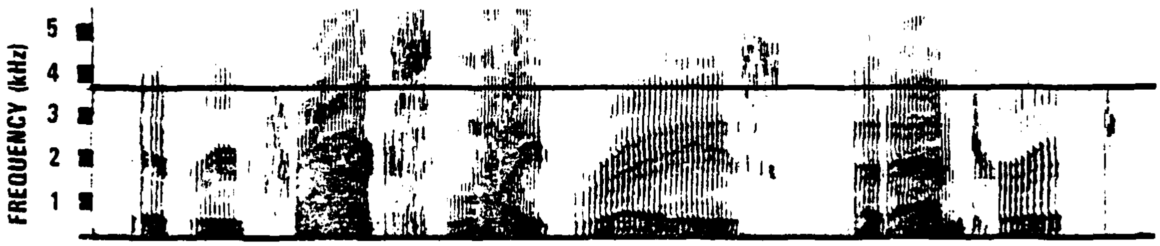
Current narrowband LPC systems have lower intelligibility than conventional, non-processing, systems of similar bandwidth. Research at NRL [1] has discovered a number of the reasons for reduced intelligibility with the LPC. One of these is that the spectral components of some important speech sounds, chiefly sibilants, are almost entirely above the corner frequency of the 3.6 kHz LPF. This causes the LPC to miss, or misrepresent, these sounds. The problem is not as severe in conventional telephone systems because sampling and digital processing are not used, and because the carbon microphone in standard telephone handsets is a non-linear device which spreads the speech spectra so that some of the high frequency energy is shifted into the system passband. Figure 1 shows the speech spectra for a short phrase at the outputs of a dynamic microphone and a carbon microphone. Note that the spectra for the /s/ in "reflects" and "rays" are almost entirely above 3.6 kHz for the dynamic microphone output, but extend well below 3.6 kHz for carbon microphone output. Unfortunately, the distortion produced by the carbon microphone produces unsatisfactory results with the narrowband LPC.

An alternative approach to spreading high-frequency sound spectra into the passband is to deliberately introduce aliasing. This is accomplished by increasing the audio input bandwidth to above one-half the sampling frequency when speech contains primarily high-frequency spectral components. Figure 2 shows spectra for a phrase before and after such processing. An adaptive-bandwidth filter may be implemented digitally [1] or by analog means, as presented in this report.

Manuscript approved January 18, 1983.

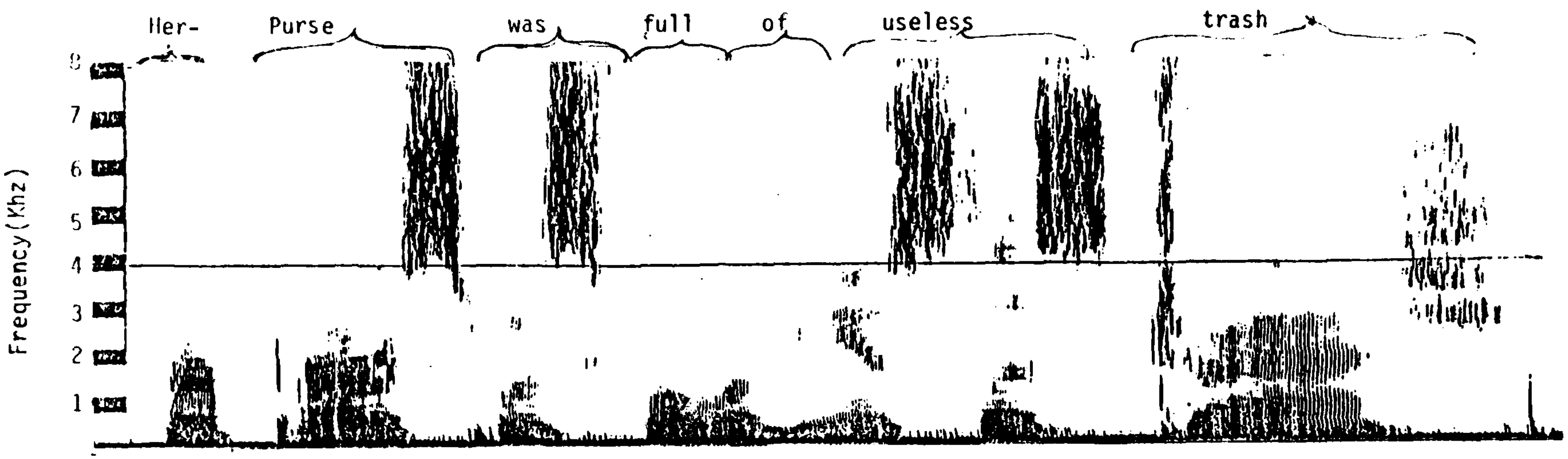


(A) DYNAMIC MICROPHONE OUTPUT



(B) CARBON MICROPHONE OUTPUT

Fig. 1 — Speech spectra from dynamic and carbon microphone outputs



(A) UNPROCESSED SPEECH (INPUT)

(B) SPEECH WITH OUTPUT SELECTIVE ALIASING (OUTPUT)

Fig. 2 — Input and output spectra of new front-end processor

The purpose of the adaptive-bandwidth filter is to make the narrowband LPC sounds more intelligible without increasing the LPC bandwidth or requiring changes in existing analysis hardware or software. Tests have shown that the use of the adaptive filter improves the intelligibility scores by a few points. More significantly, speech becomes more readily understandable with more clearly audible sibilant sounds.

The controlled aliasing technique can be used also with band-limited analog systems. Normally, analog communication systems have no aliasing mechanism, so adaptive filtering of the voice input would have little effect, but if the adaptive filter is followed by a sampler operating at twice the system bandwidth, to produce aliasing, and a LPF to smooth the output, effects similar to those obtained with digital processing should result.

APPROACH

Several approaches to development of the adaptive-bandwidth audio unit were considered. To duplicate the computer simulation, a digital processing system would be necessary (and one was suggested in [1]), but a substantial development effort would have been required. It was decided that an analog unit conforming to the block diagram (Figure 3) would be a sufficiently close approximation to the computer simulation to permit intelligibility testing.

In Figure 3, the main signal path runs across the top of the diagram, and the control components are shown below the main path. The input section and the high-pass filter (HPF) duplicate functions already present in the normal narrowband LPC analog input section, and deliver an audio signal of correct amplitude to the main signal path and to the control circuits. The main signal path runs from the HPF to a controllable low-pass filter (LPF2) and an output section which drives the analog-to-digital converter (ADC) in the LPC. LPF2 [2] is an integrated-circuit, switched-capacitor low-pass filter. It is a seven-pole, six-zero elliptic design with less than 0.2 dB passband ripple and more than 75 dB out-of-band rejection. The slope of the attenuation characteristic above the corner frequency approaches 100 dB/octave. The corner frequency of the LPF is set by the clock frequency supplied to the circuit (the corner frequency is 1/100 the clock frequency).

The control circuits, below the main signal path, provide simple spectrum analysis of the audio signal, and control the clock frequency to LPF2. There are two signal paths in the control section. One is bandwidth-limited by LPF1, which has a corner frequency between 3.5 kHz and 4 kHz. The other path is not filtered, but has an adjustable attenuator. The signals in the two paths drive two peak detectors, which produce d.c. outputs proportional to the peak-to-peak voltages at the detector inputs. The two detector outputs go to a comparator and electronic switch. The switch controls the clock frequency to LPF2. When the wideband signal has greater amplitude than the narrowband signal, the comparator switches the higher clock frequency to LPF2. When the wideband signal is no greater than the narrowband signal, the LPF2 clock is at the lower frequency (used also by LPF1). The adjustable attenuator in the wideband control path allows the threshold ratio of wideband/narrowband amplitudes, at which switching of LPF2 to the wideband mode occurs, to be set as desired.

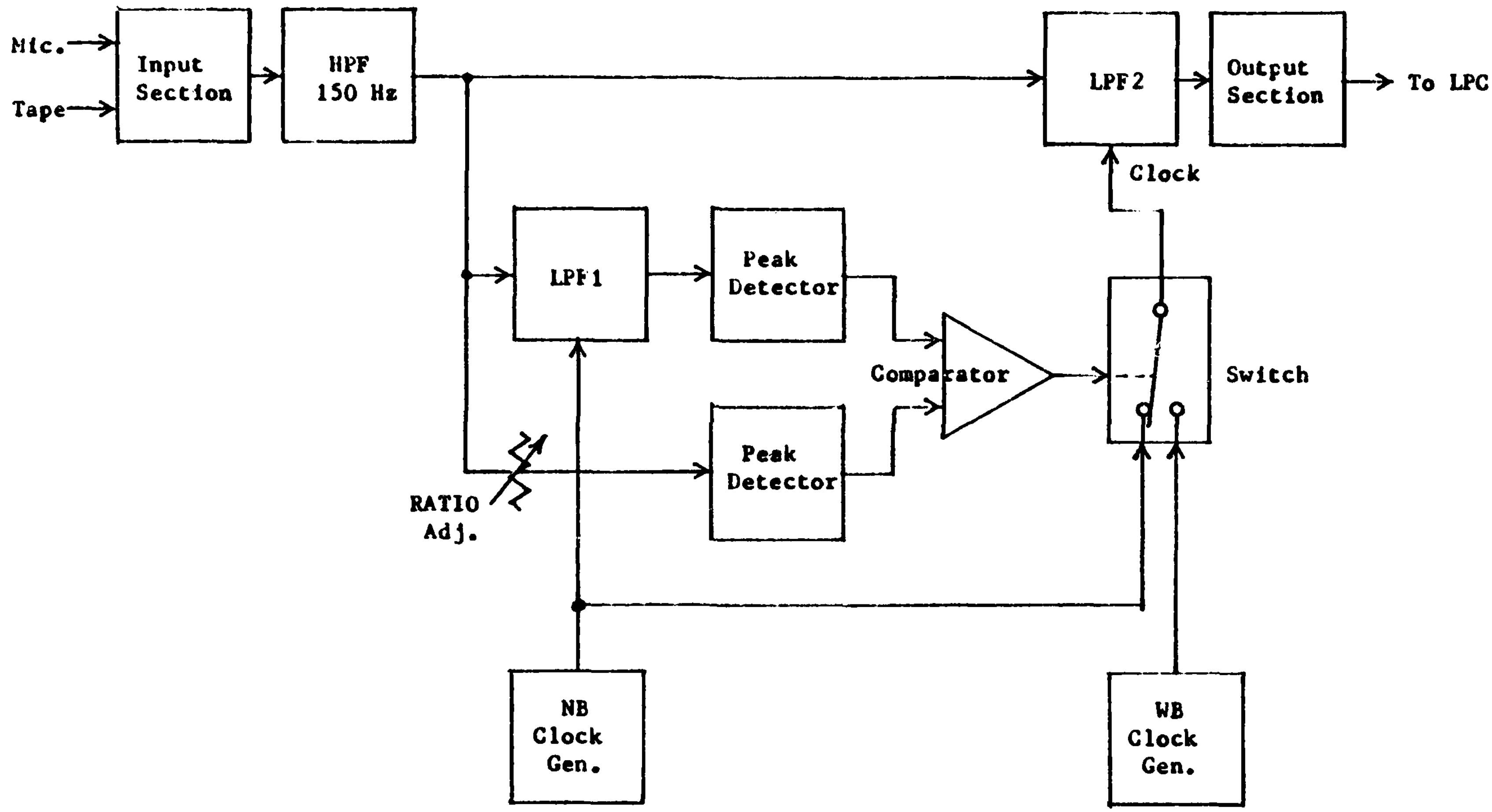


Fig. 3 — Adaptive audio unit block diagram

Two clock generators are used, each adjustable in frequency. The low-frequency generator, used by LPF1 and by LPF2 in the narrowband mode, is variable from 300 kHz to 400 kHz (corresponding to filter corner frequencies of 3 kHz to 4 kHz). The high-frequency generator, used by LPF2 in the wideband mode, is variable from 350 kHz to 700 kHz.

EXPERIMENTS

An experimental audio unit conforming to the block diagram has been built. A detailed description of the circuits is in the appendix. Experimental use of the audio unit indicated that critical elements in its effectiveness are the narrowband and wideband corner frequencies and the attack and decay times of the peak detectors. Provision had been made for control of corner frequencies, but the detector time constants were not adjustable without component changes.

As originally designed, the detector time constants were short, in an attempt to minimize the control delay. Unfortunately, voiced sounds, which do not require bandwidth switching, contain short intervals in which only high-frequencies are present. These sounds caused bandwidth switching for short periods (200 microseconds) at the pitch frequency of the voiced sound. Increasing the attack time constants of both peak detectors prevented most of the false triggering, and extending the decay times produced further improvement. Because of the different bandwidths of the two control paths, it appears that the wideband detector should have a slightly slower attack and faster decay than the corresponding characteristics for the narrowband detector. Further work on time constants and, possibly, a better sensing method should be done.

Diagnostic Rhyme Tests (DRT) [3] were made for three conditions of the audio input unit. These tests used tape recordings of speakers reading lists of words. The words occur in pairs which differ in only one sound. Pair differences are chosen to test various speech characteristics as the test progresses. Intelligibility scores are produced by listeners for each of the characteristics.

Tests were run, with a female speaker, for fixed 3.6 kHz and 4 kHz corner frequencies (bandwidth switching inhibited) and for adaptive bandwidth switching between corner frequencies of 4 kHz and 6 kHz. The DRT intelligibility scores are tabulated below.

Table 1

Characteristic	3.6 kHz	4 kHz	Switched BW
Voicing	74.2	80.5	88
Nasality	98.4	98.4	96.1
Sustention	67.2	72.7	77.1
Sibilation	78.1	80.5	82.6
Graveness	68.7	72.2	71.1
Compactness	88.3	92.2	88.8
Average	79.2	82.8	84.0

Additional tests with other speakers would produce somewhat different results, but the trends indicated by the scores in the table are probably valid. The first two columns of scores show the improvement in intelligibility when bandwidth is increased; no significant aliasing was present during these tests. The third column of scores used switching between 4 kHz and 6 kHz corner frequencies, and input frequencies above 4 kHz would have been shifted by the processor below 4 kHz. Differences in the second and third score columns indicate changes in intelligibility caused by the bandwidth switching process. There were significant improvements in three characteristics, and losses in three. Two of the losses were in characteristics with scores near or above 90, while all of the improvements were in characteristics with scores in the 70's and 80's, so it appears that the intelligibility improvement through use of the adaptive-bandwidth technique is significant, and probably greater than the improvement indicated by the average scores.

CONCLUSIONS

The experiments with the adaptive audio unit have demonstrated the benefits both of wider fixed bandwidth audio input to the LPC and of adaptive switched bandwidth. The equipment used was designed to be inexpensive and easy to build; it is not the only way to implement the desired function, nor the best. Reference 1 outlines a digital method for input bandwidth variation, and many hybrid analog/digital techniques are also possible. A final decision on a production version of an LPC with variable input bandwidth will depend on details of the processor design, and is beyond the scope of this investigation. It should be noted, however, that some form of analog bandpass filter will always be necessary ahead of the digital processor, and that the filters used in the experimental unit are cheap, small (8-pin DIP) and have excellent characteristics (flat response in the passband and very steep cutoff in the stopband). The filters used do not meet military environmental specifications, and it is not known at present whether military grade units will be produced.

ACKNOWLEDGMENTS

The authors thank George Kang of NRL for the original concept and overall direction of the project, and Mark Lidd of SPSI for his assistance in the experiments.

REFERENCES

1. G. S. Kang and S. S. Everett, "Improvement of the Narrowband Linear Predictive Coder, Part I: Analysis Improvements", Naval Research Laboratory, Washington, D. C., NRL Report 8645, 1982.

2. EG&G RETICON Preliminary Data Sheet: R5609 Low-Pass Filter.

3. W. D. Voiers, A. D. Sharpley, and C. J. Hehmsoth, "Research on Diagnostic Evaluation of Speech Intelligibility", Report AFCRL-72-0694, Jan. 1973

Appendix

ADAPTIVE-BANDWIDTH AUDIO UNIT

GENERAL DESCRIPTION

This unit will be installed between a source of analog audio, such as a tape recorder output or a handset microphone, and the analog-to-digital converter (ADC) of a linear-predictive-coder (LPC) digital audio system.

Normally, the unit will act the same as the original audio input section of the LPC system, but whenever the audio input signal contains significant energy at frequencies above the normal audio passband AND little or no energy within the normal passband, then the passband of the adaptive unit will switch to a wideband mode until normal energy distribution is resumed. This will have the effect of allowing isolated high frequency bursts (fricatives, sibilants, etc.) to reach the LPC ADC. The 8000 Hz sampling rate of the LPC A/D converter will cause the apparent frequencies of inputs above 4000 Hz to be folded below 4000 Hz in the final LPC audio output, so frequency distortion will occur, but sound output will be present to represent all the original audio inputs.

Reference to the block diagram, Figure A1, will help clarify the operation of the unit. Starting from the left, there is an audio input section with inputs for microphone and tape. The input section is followed by a 150 Hz high-pass filter (HPF), which removes low frequency line noise and unnecessary audio energy from the input signal. The signal path then splits. The main path leads to a low-pass filter (LPF2) and an output section which provides suitable drive to the LPC ADC. Normally, LPF2 is set to provide the normal cut-off frequency in the range from 3600 Hz to 4000 Hz. The cut-off frequency of LPF2 is switchable, under control of the control circuits shown below the main path on the diagram. The full-bandwidth audio signal goes, in addition to the main path, to an additional low-pass filter (LPF1) followed by a peak detector, and to another peak detector working on the full bandwidth. The outputs of the two peak detectors represent the amplitudes of band-limited and full bandwidth signals, respectively. The output of the full bandwidth channel can be attenuated as desired by a variable attenuator; this attenuated level is then compared with the output level of the band-limited channel and a decision made as to which clock frequency to supply to the main channel LPF2. When the output of the band-limited detector exceeds the attenuated output of the wide-band detector, the switch sends a low clock frequency to LPF2, and the normal, narrow-band condition results. When the attenuated wide-band detector output exceeds the band-limited detector output, a higher clock frequency is switched to LPF2, and the wide-band condition results - this continues until normal frequency relationships resume.

The above operation is possible because LPF2 has a cut-off frequency determined by the clock frequency supplied to it. The cut-off frequency is the clock frequency divided by 100, so two variable clock generators are used, one for the normal bandpass and another for the wide-band condition. LPF1 is the same type of filter, but it is not necessary to switch its bandwidth, so it

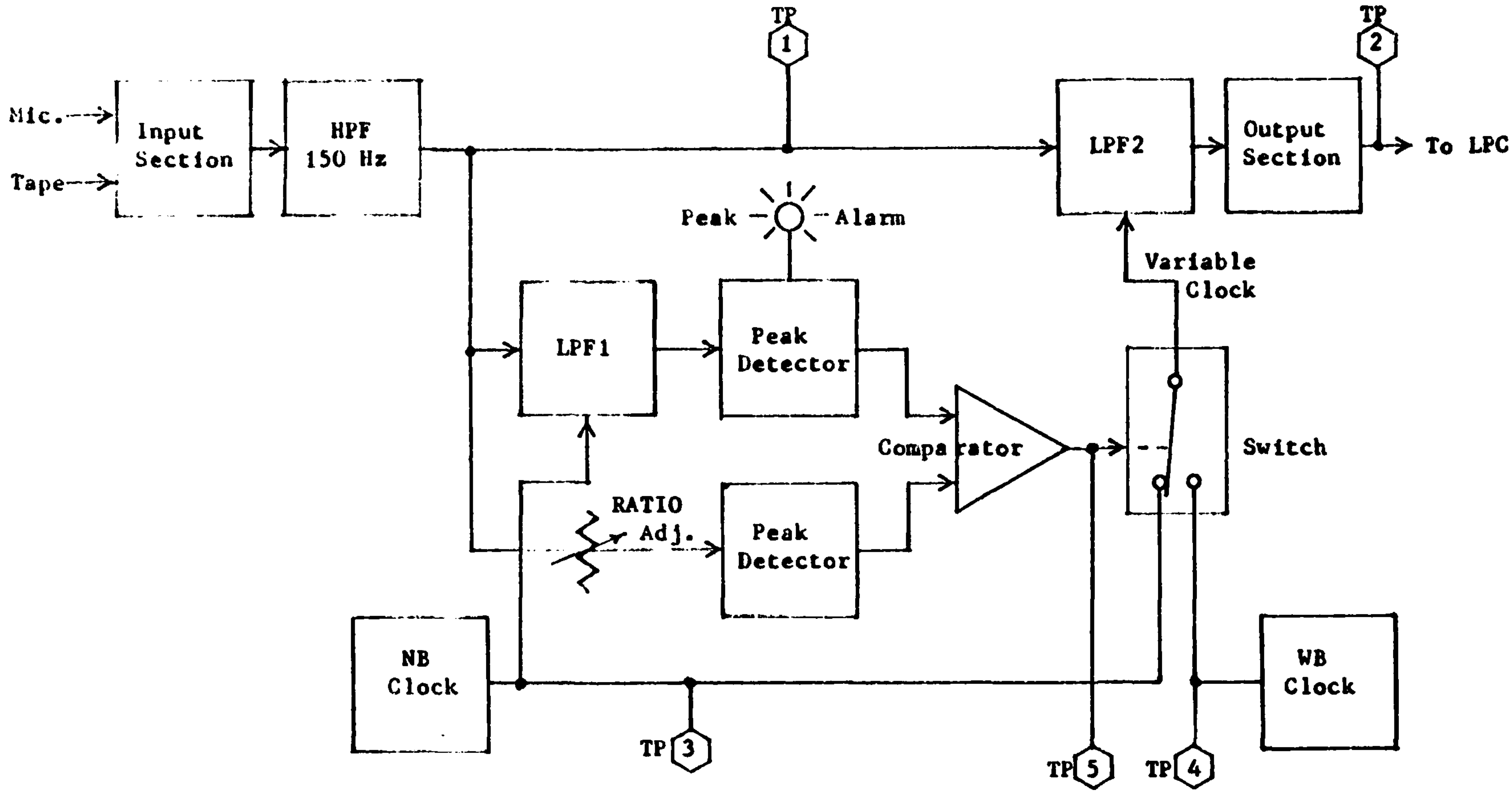


Fig. A1 — Adaptive audio unit block diagram is clocked continuously from the lower frequency clock generator. Both generators are variable, so the two cut-off frequencies can be set independently over a range of values: 3 to 4 kHz for the LF clock, and 3.5 to 7 kHz for the HF clock.

CIRCUIT DESCRIPTION

This description is keyed to the blocks in Figure A1 and the circuit schematic, Figure A2. Circuit references are, generally, to integrated circuit identifiers (U1a, etc.) or control names (Gain, etc.). The block and circuit diagrams have similar signal flow patterns.

INPUT SECTION & HPF. Four sections of a quad op-amp, U1, are used. U1a provides amplification for a microphone output. The gain is variable from 0 dB to 40 dB by changing the setting of the Mic. Gain control. The output of U1a and the signal from the TAPE input go to the Source switch, which routes the selected signal through the Gain control to U1b. U1b is connected for a gain of approximately 20 dB, and drives the HPF. U1c and U1d are connected as voltage followers in an active high-pass filter circuit with a cut-off frequency of 150 Hz. The output of U1d is the audio signal routed to TP1, LPF2, LPF1 and the Ratio control.

LPF2 & OUTPUT SECTION. LPF2 receives audio from U1d and clock from the Var. Clock line. The audio input level, as monitored at TP1, must not exceed 12 Vpp. The output of LPF2 goes to the Output Level control and to output amplifier U6. The output level, monitored at TP2, must not exceed 10 Vpp. The output of U6 is passed through a low-pass filter consisting of a 50 mH inductor and a 7500 pF capacitor before being sent to the LPC. This final LPF serves to remove high frequency noise due to leakage from the clock circuits.

The control of the signal path bandwidth is done in the block of circuits shown directly below the main signal path in the block diagram. The narrowband control branch consists of LPF1, which is the same as LPF2 except that the clock frequency is constant, followed by a peak detector. The PEAK DETECTOR uses U2a as an amplifier, and U2b as an inverter. The outputs of U2a and U2b are connected through diodes to a filter consisting of a capacitor and resistors to control the charge and discharge currents in the capacitor. The d.c. voltage on the capacitor is proportional to the peak-to-peak audio voltage into the detector circuit. As drawn, the charging current is controlled by the 22k resistor, and the discharge by the series combination of the 22k and 80k resistors. The wideband control branch passes through the Ratio control and into a peak detector using U2c and U2d. This detector is identical to the one described above, except for the filter time constants. The output of the narrowband detector is used to drive a PEAK ALARM. The alarm uses U3d as a comparator to light a LED whenever the detected signal level exceeds a threshold voltage set by the Peak Alarm Adj. control. This control is set so that the LED starts to light when the signal amplitude reaches 10 Vpp, and the LED reaches full brightness for signal amplitudes of 12 Vpp or greater. This alarm is intended to provide warning of signal levels greater than the maximum allowed for the LPF's.

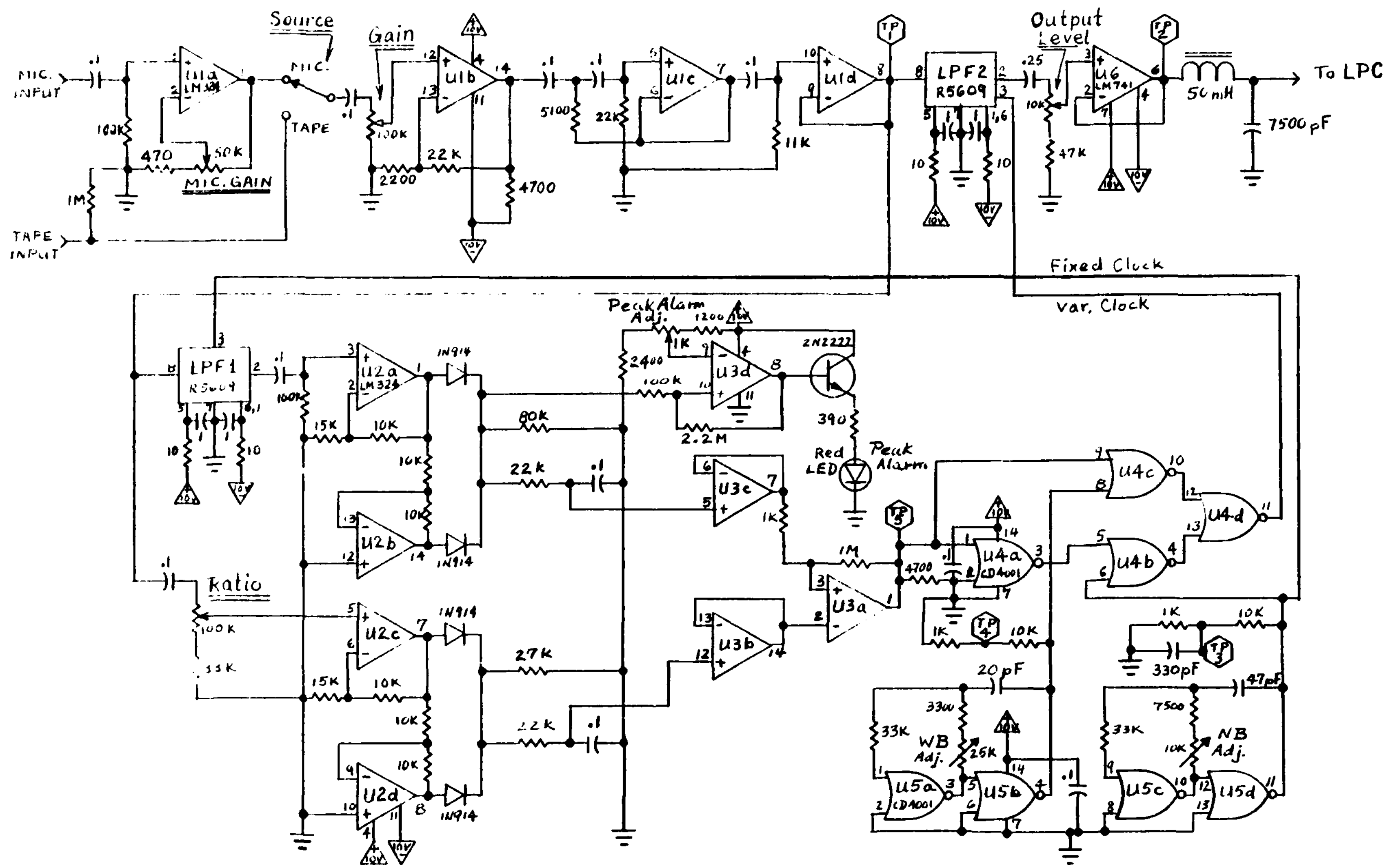


Fig. A2 — Circuit diagram

COMPARATOR & CLOCK SWITCH. The voltages from the capacitors at the outputs of the peak detectors are isolated from the circuits following by voltage followers U3c and U3b. The outputs of the voltage followers are connected to the inputs of comparator U3a. When the narrowband output is greater than the wideband output, the comparator output is high. When the wideband output is greater, the comparator output is low. The comparator output controls a switch composed of the four NOR gates of U4. These switch the appropriate clock frequency to LPF2. TP5 is provided so that the comparator action can be monitored.

CLOCK GENERATORS. There are two clock generators, one for LPF1 and the narrowband mode of LPF2, and another for the wideband mode of LPF2. They are free-running multivibrators, using the NOR gates of U5 as inverting amplifiers. The high-frequency clock is produced by U5a and U5b, with an adjustment range by the WB Adj. control of 350 kHz to 700 kHz. The low-frequency clock, controlled by NB Adj., has a range of 300 kHz to 400 kHz. TP3 and TP4 are provided for monitoring the clock frequencies. These test points are isolated from the clock generators by attenuators to minimize any shift of frequency that might be caused by a frequency measuring device.

POWER SUPPLY. No power supply is shown, but all circuits operate from supplies of +10 V and -10 V. The maximum loading is about 35 mA for the positive voltage and 5 mA for the negative voltage. Separate, variable power supplies have been used to date, but small, fixed supplies have been ordered. When they arrive, they will be installed in the chassis on which the circuit card is mounted.

The integrated circuits in the audio unit are listed below:

Identifier	Part No.	Description
U1	LM324	Quad Op-Amp
U2	LM324	Quad Op-Amp
U3	LM324	Quad Op-Amp
U4	CD4001BE	Quad NOR Gate
U5	CD4001BE	Quad NOR Gate
U6	LM741CN	Op-Amp
LPF1	R5609	Reticon LPF
LPF2	R5609	Reticon LPF

HARDWARE DESCRIPTION

All the circuits except the power supply and some of the controls are located on a Douglas 11-DE-3 plug-in circuit card, which fits a 22-pin socket. Included on the board are color-coded test points and screwdriver adjustments for the narrowband and wideband clock frequencies and for the peak alarm threshold level. Figure A3 is a sketch of the component side of the circuit board, showing component and test point locations. Figure A4 shows the wiring of the card.

The circuit board plugs into a chassis (which will contain the power supplies) having input and output connectors on its upper surface and operating controls on its front surface. The controls, from left to right, are

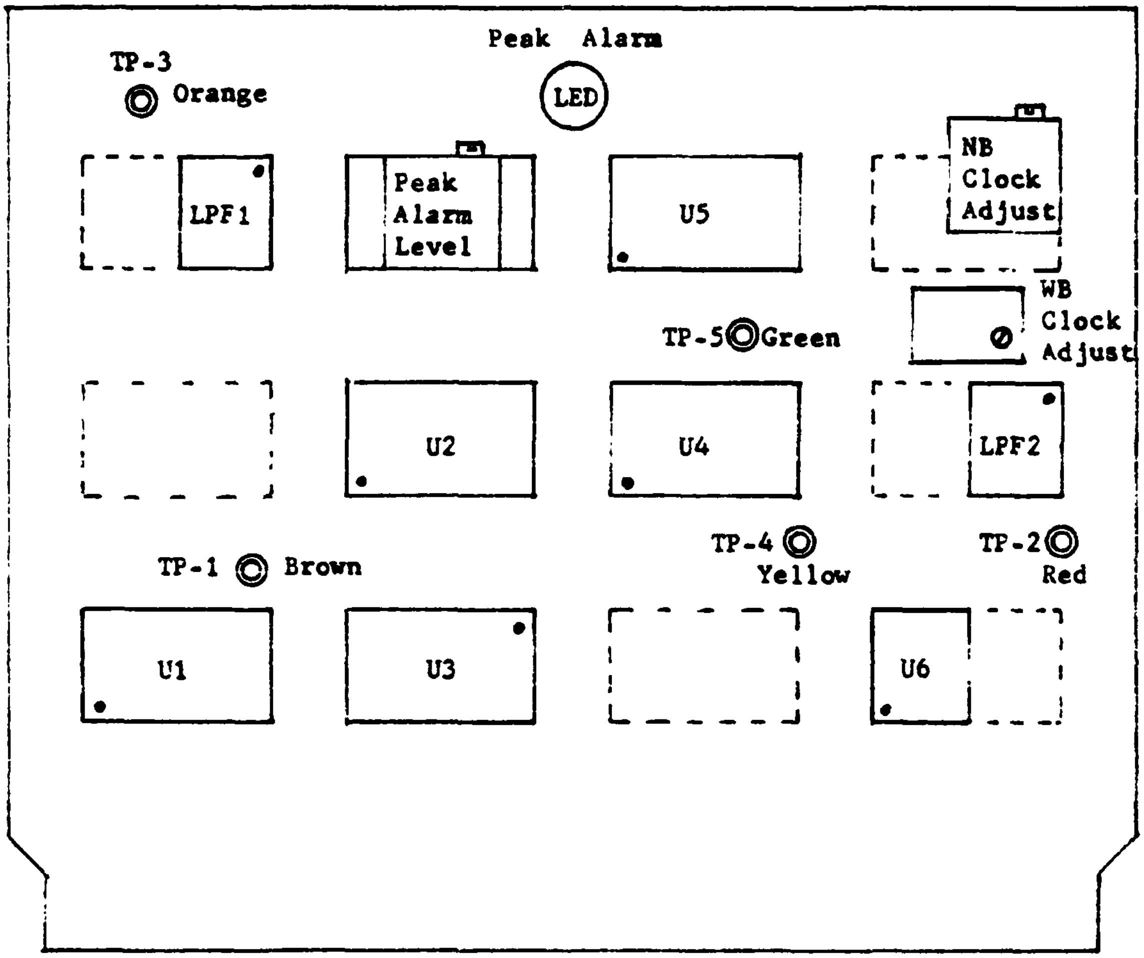


Fig. A3 — Component side of circuit board

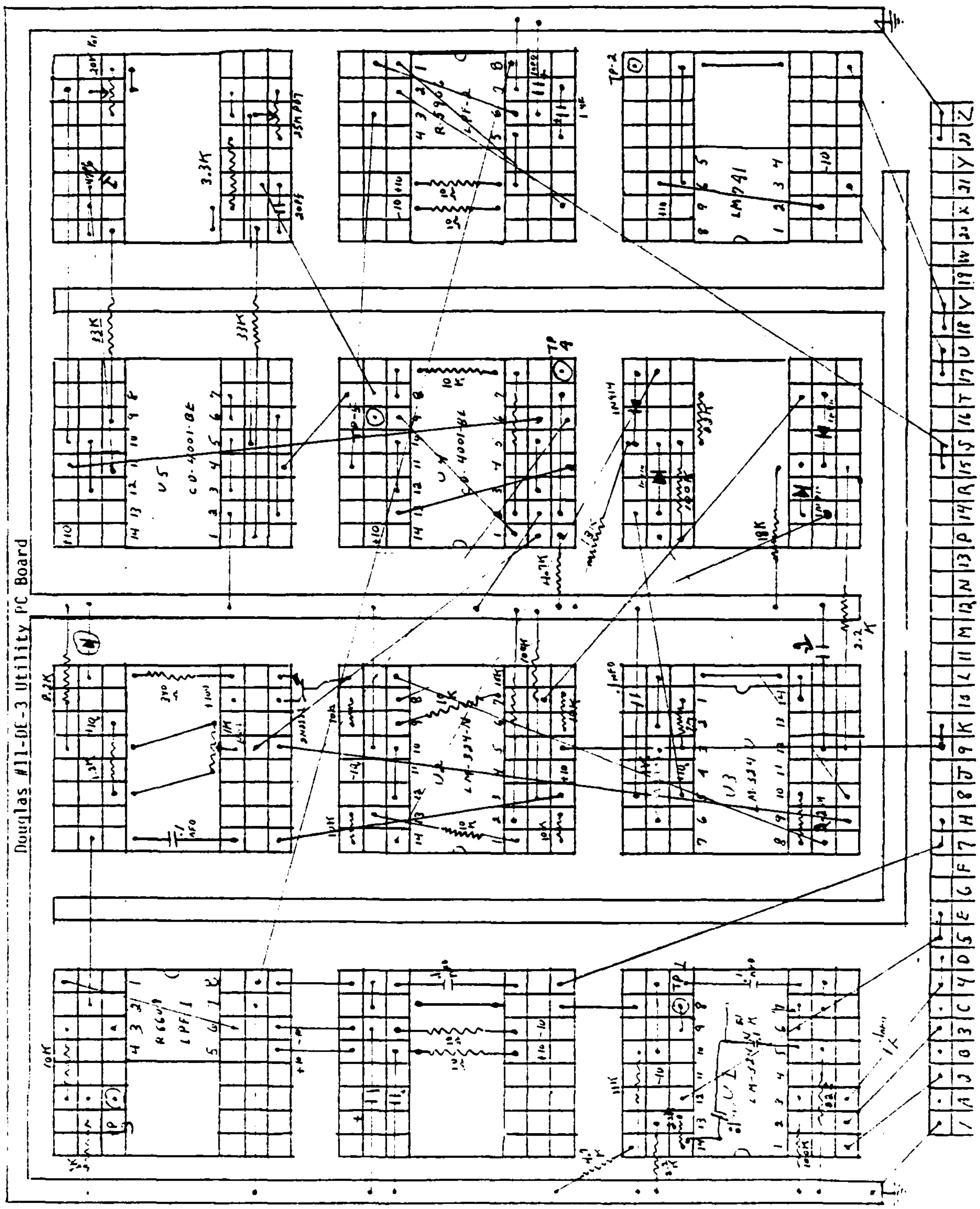


Fig. A4 — Circuit board wiring

MIC. GAIN, MIC/LINE switch, LINE (master) GAIN, RATIO and OUTPUT LEVEL.

Figure A5 is a frequency response plot for the circuit. The effect of the 150 Hz high-pass filter is seen at the left of the plot, and the effect of the switchable low-pass filter (for two cutoff frequencies) is seen at the right. The two LPF clock frequencies used were approximately 360 kHz and 703 kHz. A line having a slope of 100 dB/octave is shown to help gauge the slopes of the LPF responses.

CIRCUIT ADJUSTMENTS

The MIC. GAIN and LINE GAIN controls will be set during operation for a satisfactory output signal level.

The RATIO control can be set by observing the relative amplitudes of a constant level test signal at U2-1 and U2-7 while adjusting the control. A roughly calibrated scale is provided for the RATIO control.

The peak alarm threshold adjustment determines the peak-to-peak signal level at which the LED indicator starts to light. It is adjusted by supplying a steady test signal within the filter passbands (say, 1 kHz) of the desired amplitude, measured at TP1, and adjusting the potentiometer until the LED just lights. Signal levels above the set level will cause brighter indications, and lower levels will not light the LED. The indicator is now set to light for levels of 10 Vpp or higher.

The narrowband and wideband clock frequency adjustments are made while monitoring the frequencies at TP3 and TP4, respectively. The frequencies of the clocks are to be 100 times the desired cut-off frequencies of the filter. The clocks have been set initially to approximately 400 kHz and 600 kHz.

The OUTPUT LEVEL control is provided so that the output signal to the LPC can have the correct value (10 Vpp) while the circuits within the Adaptive Audio Unit operate at optimum levels. This control has been set so that 10 Vpp at TP1 yields 10 Vpp at the output (TP2).

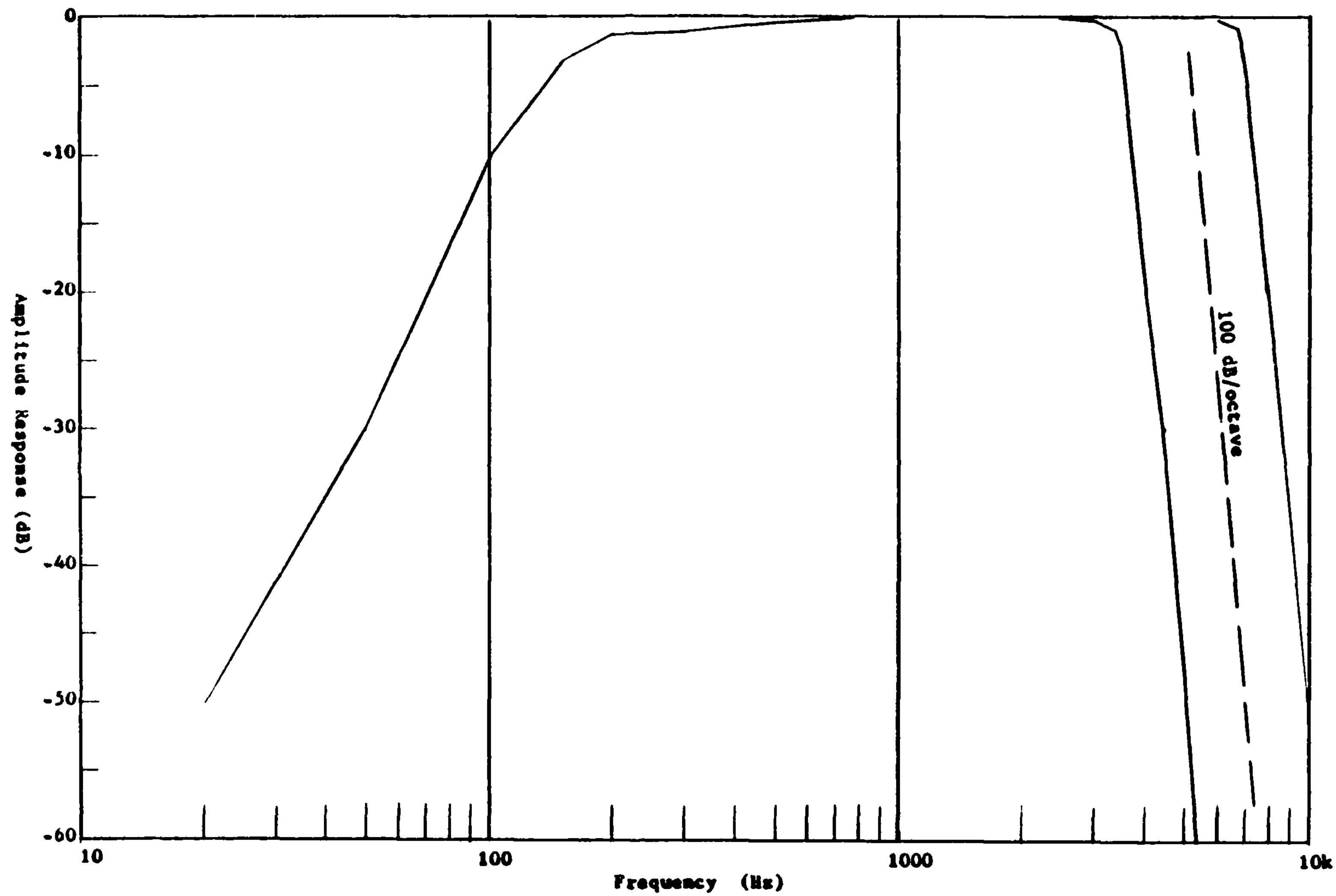


Fig. A5 — Frequency response plot for the circuit

NOTES:

(1) U1, U2 and U3 are LM324 quad op-amps. These amplifiers introduce noticeable cross-over distortion in the signal. The distortion percentage is small for large signals, but still undesirable. Only U1 is in the signal path from input to output, so it is the only amplifier likely to have a direct effect on signal quality. It is recommended that future versions of the circuit use different amplifiers; LM1558 dual op-amps or TL084C quad op-amps would be suitable. The TL084C has the same pin connections as the LM324 (except for reversed power supply polarity) and could be used without circuit changes.

(2) When LPF1 and LPF2 are overdriven (>12 Vpp), their outputs show distortion in the form of output peaks exceeding the input peaks. In future versions of the circuit, it would seem prudent to put a limiter (clipper) ahead of the LPF's to prevent overdriving. With the present circuit, the output level should be set equal to the input level, and the peak alarm should start to light at a signal level of 10 Vpp.

(3) A low-pass filter has been added in series with the output of the circuit card to reduce the high frequency noise output. This noise consists of feed-through of the clock signals to the two LPF's and coupling of transients from the two clock generators (through stray capacitances and through the power lines - bypass capacitors have been placed across the clock chips). The added filter consists of a 50 mH inductance and a 7500 pF capacitor, as shown at the right on drawing 4. The inductance has a shunt capacitance of approximately 42 pF. The filter has 3 dB attenuation at a little above 8 kHz, 20 dB at 28 kHz and more than 45 dB above 80 kHz.

4-8

DTI