Dolby®

Dolby® Model 422

Reference Encoder/Decoder For Tape Duplication

Users' Manual

Dolby® Model 422

Reference Encoder/Decoder For Tape Duplication

Users' Manual

USERS' MANUAL

FOR

MODEL 422
REFERENCE ENCODER/DECODER
FOR TAPE DUPLICATION

Dolby Laboratories Incorporated

U.S.A. 100 Potrero Avenue, San Francisco, CA 94103
Tel: 415-558-0200; Telex: 34409; Fax: 415-863-1373

U.K. 346 Clapham Road, London SW9 9AP
Tel: 071-720-1111; Telex: 919109; Fax: 071-720-4118

ISSUE 2
S91/8997/9145
Part No. 91197

CONTENTS

SECTION 1
INTRODUCTION AND SPECIFICATIONS

1.1 Introduction

The Dolby Model 422 is a reference encoder/decoder having consumer noise reduction (NR) characteristics, Dolby B-, C- and S-type. It replaces the Model 330 and is designed primarily for use in the production of pre-recorded tapes, either audio cassettes or the soundtracks of video cassettes. All the circuitry is on one mother-board except the S-type circuits which are contained on a separate "piggy-back" board and the input filters which plug on to the mother-board.

The unit contains a pair of channels of processing which can be internally preset as encoders or decoders, plus a further pair of decoders.

The unit is used to prepare "loop-bin" or "running" masters, with two channels encoding the source material to feed the master recorder and the remaining pair decoding the master tape to permit instant quality checking. It may also be used wherever it is necessary to listen to pancakes or loaded duplicated tapes to check their quality.

The Model 422 contains a built-in calibration oscillator giving the Dolby tone characteristic of the type of NR selected.

The unit is normally delivered with a plug-in 16 kHz low-pass filter designed for B- and C-type duplication of Philips compact cassettes. For use with S-type NR this may be bypassed if desired by a link on the rear remote connector. See paragraph 4.2. For duplication of VHS cassettes, this filter should be replaced by the Cat. No. 370/371 filter card which provides a low-pass response at about 14 kHz plus a notch at the television horizontal frequency.

The Model 422 is self-contained in a 1U (44 mm, 1.75 inch) high rack-mounted case.

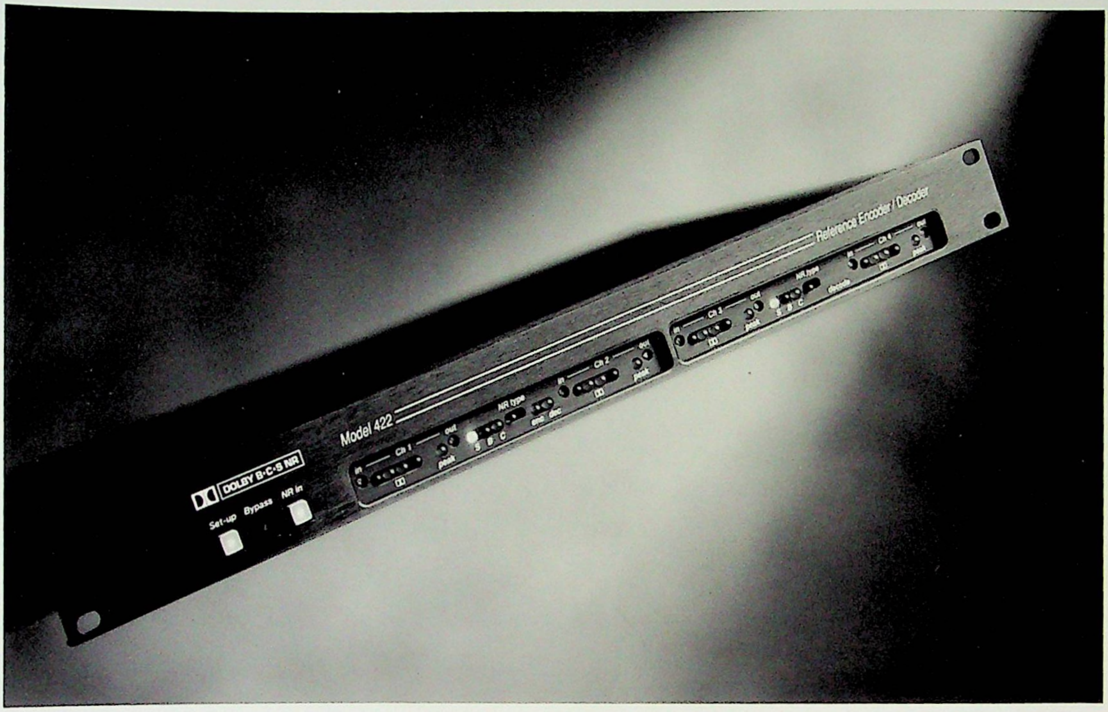
The Model 422 is also available without the S-type NR (designated Model 422B). The S-type NR can be added at any time by plugging in the "piggy-back" board (Cat. No. 252). Owners of Model 422B should disregard references to S-type NR in this manual.

1.2 This manual

This manual is not intended to describe the process of tape duplication in detail, but to show how the Model 422 should be incorporated into the chain. We recommend reading this manual in conjunction with the software information manual published by Dolby Laboratories Licensing Corporation and available from Dolby Laboratories in San Francisco or London.

Dolby® Model 422
Reference Encoder/Decoder

DOLBY B·C·S NR



The Dolby Model 422 is a 1-U high encoder/decoder unit providing four channels of Dolby B-, C-, and S-type noise reduction. It is intended primarily for encoding running masters for audio and video tape duplication, and for decoding recordings for quality assessment.

Channels 1 and 2 can operate either as encoders or decoders depending on the position of an internal jumper, while channels 3 and 4 operate as decoders only. The NR type for each pair is selected separately by front-panel toggle switches and indicated by LEDs. Thus the Model 422 can simultaneously decode four channels; simultaneously encode and decode one pair of channels with the same NR type; or encode one pair of channels with one NR type while simultaneously decoding with another. The latter function, for example, allows checking Dolby S-type encoded material for compatibility with Dolby B-type playback. A front-panel pushbutton switches the processing of all channels on and off.

Like all other professional Dolby NR units, the Model 422 contains a signal generator providing calibration "Dolby tones" with audible characteristics identifying the type of NR processing in use. The generator is activated by a front-panel "set-up" pushbutton, while LED calibration displays allow accurate adjustment of reference levels. In addition, each channel has an overload LED to warn of the onset of distortion in the recording medium. An internal trimmer can be adjusted over a 9 dB range to set the LED's threshold for a particular combination of format, tape formulation, and program material characteristics.

To prevent mistracking of the NR system it is necessary to stop the NR encoder from responding to signals which will not be reproduced into the decoder. The Model 422 normally contains a plug-in 16 kHz low-pass filter suitable for use in duplication of audio cassettes. For video applications this should be replaced by a filter with a notch at television horizontal frequency (Cat. No. 370 for NTSC and Cat. No. 371 for PAL or SECAM).

A rear-mounted, 15-pin D-connector allows most of the Model 422's functions to be remotely controlled. In addition, external filters or processors can be introduced via break points accessible on a 9-pin D-connector on the rear.

All of the Model 422's circuitry except the S-type processors and the input filters is carried on one motherboard, with input and output potentiometers, switches, and indicating LEDs mounted along its front edge; the S-type processors are on a separate plug-in board. The Model 422 is designed to permit continuous operation.

Dolby® Model 422 Specifications

Except where specified, the following apply with the Model 422 adjusted so that Dolby level is +4 dBr[1] at the input and output.

Signal Processing:
Four channels of switchable Dolby B-type, S-type, and C-type NR. NR process for each pair of channels (1 & 2, 3 & 4) is selected by front-panel toggle switches, and is indicated by LEDs (red for B-type, yellow for C-type, green for S-type).

Channel Modes:
Channels 1 and 2: internally selectable as decode or encode pair. Mode indicated by front panel LEDs (red for encode, green for decode).
Channels 3 and 4: decode only.

Signal Connections:
XLR input and output connectors for each channel.
9-pin male D-connector with breakpoints for each channel at a normalized level of –6 dBr.
15-pin female D-connector for remote control.

Operating Controls:
Common controls (all channels):
Bypass (via relays).
Processing in/out.
Set-up switch activates Dolby tone generator for use during alignment.

Controls for each pair of channels:
NR type (S,B,C) with corresponding LED indicators.

Controls for individual channels:
Input multiturn potentiometer.
Output multiturn potentiometer.

Calibration Displays:
4-LED array for each channel permits accurate adjustment of Dolby level by matching the intensity of a pair of green LEDs.

Overload LEDs (one per channel):
Lights when peak level of input signal exceeds threshold (charge time-constant 1 ms, discharge 500 ms approximately). The threshold is user-adjustable over the range +1 to +10 dB with respect to Dolby level.

Dolby Tone Generator:
Delivers appropriate characteristic Dolby tone corresponding to the type of NR selected.

Signal Levels:
Input and output line levels can be adjusted over the range ±10 dBr for Dolby level.

Input Circuits:
Electronically balanced, 20 k ohm substantially resistive. Common-mode rejection, better than 55 dB, 50 Hz to 10 kHz. Maximum input level, +27 dBr balanced, +21 dBr unbalanced.

Output Circuits:
Electronically balanced and floating, output impedance approximately 20 ohms.
Output balance within 1 dB into symmetrical 600 ohm load. Output float[2] better than –40 dB 50 Hz to 1 kHz.
Maximum level into 600 ohm or higher, +26 dBr balanced, +21 dBr unbalanced.
Either leg of the output may be grounded for unbalanced operation with no change in level.

Processor Headroom:
More than 15 dB above Dolby level.

Overall Frequency Response (with input filter bypassed)**:**
20 Hz to 15 kHz ±1 dB, encode-decode (same NR system) at any level.

Overall Harmonic Distortion:
0.2% maximum at Dolby level.

Overall Dynamic Range:
At least 96 dB, clipping level to CCIR/ARM noise.

Matching Between Units:
±1 dB at any level 20 Hz to 15 kHz.

Interchannel Crosstalk:
Channels 1 and 2 as encoders feeding channels 3 and 4 as decoders: better than -50 dB, 20 Hz to 15 kHz.

Power Line Input:
Voltage selector on rear gives nominally 100 V, 120 V, 220 V and 240 V, covering ranges 85-115 V, 102-132 V, 187-242 V and 204-264 V, 50/60 Hz single-phase AC. Total power consumption, approximately 40 VA maximum. The unit is designed to permit continuous operation.

Operating Temperature:
Ambient 0 to 40°C.

Fuses:
500 mA for 100 and 120 V; 250 mA for 220 and 240 V, slow-blow. The unit accepts either 1.25" or 20 mm fuses.

Size:
44 x 483 mm rack mounting (1.75" x 19"). Maximum projection behind mounting surface, 285 mm (10.2"); a further 65 mm (2.5") required for standard XLR connectors.

Weight:
6 kg (13 lb).

Specification subject to change without notice.

1. 0 dBr is 775 mV rms.
2. Output float is the level across a balanced load relative to an interfering signal injected at one end of the load.

Dolby

Dolby Laboratories Inc., 100 Potrero Avenue, San Francisco CA 94103-4813.
Tel: 415-558-0200 • Fax: 415-863-1373 • Tlx: 34409

346 Clapham Road, London SW9 9AP.
Tel: 071-720-1111 • Fax: 071-720-4118 • Tlx: 919109

S91/8869/9147

1.4 Safety Information

WARNING: **Check that the unit has been set to the correct supply voltage and that the correct fuse is installed. To reduce the risk of fire, replace the fuse only with the same type and rating.**

For 100/120 Vac, use 500 mA/250 V—1/4" x 1-1/4" slow-blow fuse.
For 220/240 Vac, use T250 mA/250 V—5 x 20mm time-lag fuse.

The power supply input connector has positions for two fuses and will accept carriers for either 20 mm or 1.25" fuses; only the lower fuse position is electrically connected. Select the appropriate fuse and carrier, and insert the assembly into the lower position with the arrow on the carrier in the same direction (downwards) as the arrows inside the compartment door (**Note:** a spare fuse of the same rating and type can be put in the upper position). When closing compartment door, make sure that it clicks firmly into place.

ADDITIONAL INFORMATION FOR THE SAFE OPERATION OF THE UNIT

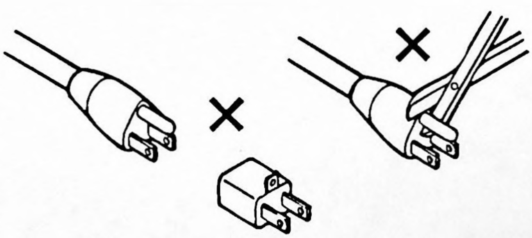
To ensure proper operation and guard against potential shock hazard, the unit must be connected only to a properly wired, grounded (earthed) power outlet. If you are uncertain about the wiring of your outlet **do not use it. Consult a qualified electrician.** The power cable is furnished either with a standard U.S.A. three-prong plug or with unterminated leads for use in other countries. The wires are colored as follows.

	International	U.S.
live or hot	**brown**	**black**
neutral	**blue**	**white**
earth	**green/yellow**	**green or green/yellow**

Before the power cable is connected to the unit, ensure that a qualified electrician has wired it as above.

U.S. Style Plugs

The ground terminal of the plug is connected directly to the chassis of the unit. For continued protection against electric shock, a three-pin power receptacle MUST be used, and the ground wire MUST always be connected. DO NOT use a ground-lifting adaptor and NEVER cut the ground pin on a three-prong plug.



Connections for United Kingdom

WARNING: THIS APPARATUS MUST BE EARTHED.

As the colours of the cores in the mains lead may not correspond with the coloured markings identifying the terminals in your plug, proceed as follows:
- the core which is coloured green and yellow must be connected to the terminal in the plug which is marked with the letter E or by the earth symbol ⏚, or coloured green or green and yellow.
- the core which is coloured blue must be connected to the terminal which is marked with the letter N or coloured black.
- the core which is coloured brown must be connected to the terminal which is marked with the letter L or coloured red.

SECTION 2
DOLBY CONSUMER NOISE REDUCTION SYSTEMS

2.1 Dolby B-type noise reduction

B-type NR is a consumer system intended primarily for use with low-speed tape, especially the Philips compact cassette. It was first introduced in 1969. It reduces audible tape hiss by 10 dB. Its single sliding band provides a fixed encoding boost (or decoding cut) within a band of variable width.

It can be considered as a high frequency emphasis of fixed magnitude whose start and stop frequencies slide upwards along the frequency axis so as not to boost the dominant, high level spectral components of the input while providing a fixed 10 dB of NR at frequencies above those dominant components (see figures 2.1 and 2.2). The fixed magnitude ensures that noise not masked by the input signal has a fixed level and therefore that minimal noise modulation is perceived.

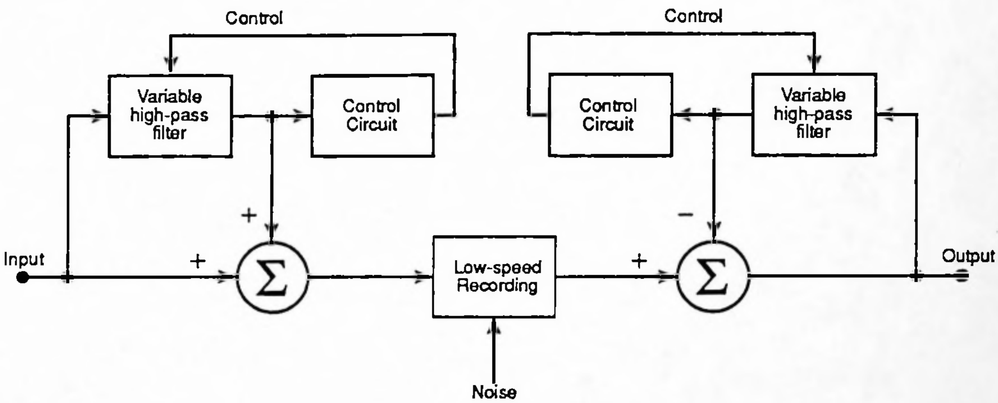


Figure 2.1 Block diagram of B-type noise reduction

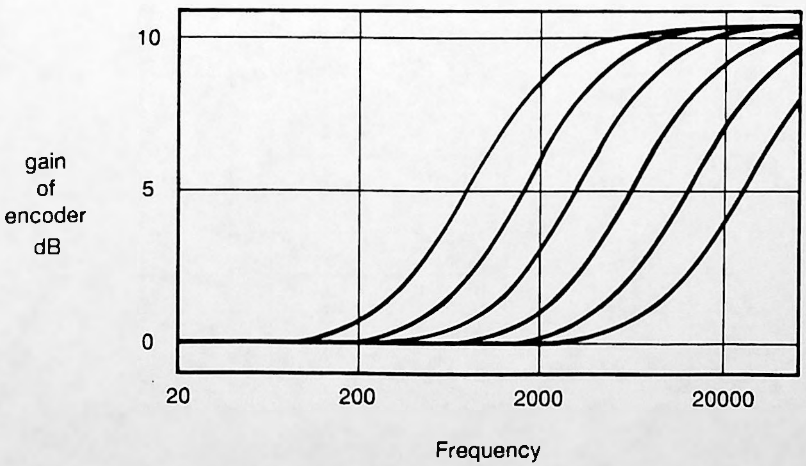


Figure 2.2 Family of response curves for B-type encoder

At any one frequency the output/input characteristic of the encoder shows gentle compression, permitting complementary expansion in the decoder.

The vast majority of B-type circuits are built by over 250 world-wide licensees under license from Dolby Laboratories Licensing Corporation.

For further details refer to appendix A.

2.2 Dolby C-type noise reduction

C-type NR was introduced in 1980 and is used in consumer audio cassette recorders and in the audio channels of professional Betacam*, MII* and U-matic SP* video recorders. It operates in a similar manner to B-type, but offers 20 dB of audible noise reduction. It achieves the steeper filter slopes required to give adequate NR at high frequencies in the presence of lower frequency dominant signals by employing two overlapping processor stages in tandem, operating with offset ("staggered") thresholds and with an action extending two octaves lower than B-type (see figure 2.3).

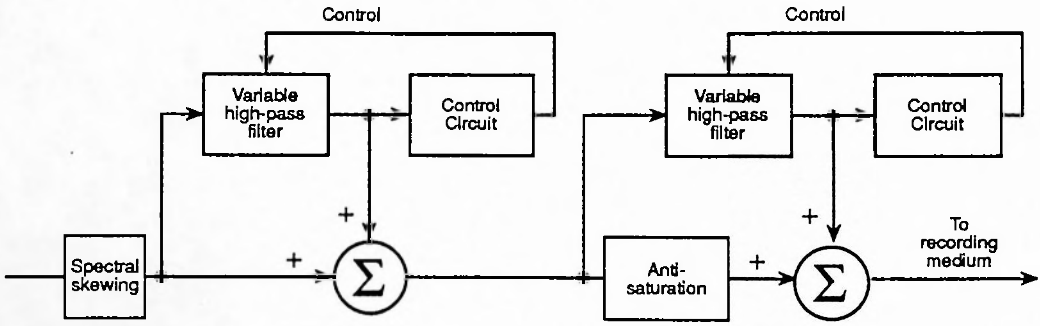


Figure 2.3 Block diagram of C-type encoder

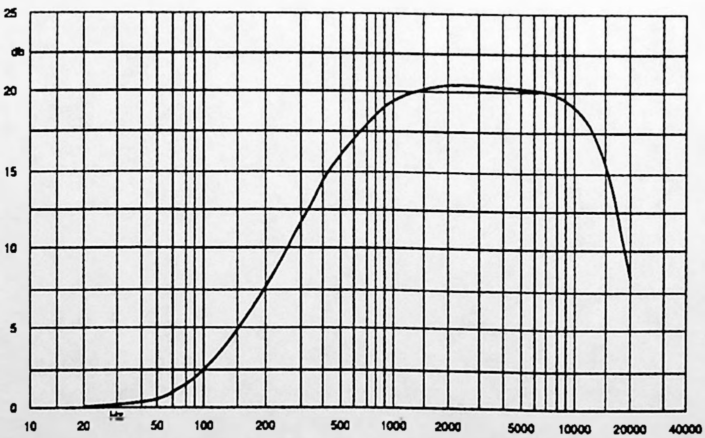


Figure 2.4 Low-level response of C-type encoder

(*Betacam and U-matic are trademarks of Sony Corporation, and MII is a trademark of Matsushita Electric Industrial Co. Ltd.)

Frequency shaping ("spectral skewing") at the input of the encoder desensitizes the processor to the effects of high frequency errors. Additional shaping in the main path ("anti-saturation") lowers the amplitude of high level high frequency signals before they are applied to the tape, reducing high frequency distortion and self-erasure.

Virtually all C-type circuits are built under license.

See appendix B for more information.

2.3 Dolby S-type noise reduction

S-type noise reduction employs the innovative techniques of action substitution and modulation control, first used in the professional Dolby spectral recording (SR) process, to give 24 dB of audible NR when applied to low-speed tape media.

The application of these techniques leads to a better adherence to the principle of least treatment, so that the gain or loss at any one frequency is less dependent on the levels of components at other frequencies. As a result the encoded signal does not have the audible gain pumping which normally accompanies compression, so that even gross mistracking during playback, such as reproducing via a B-type decoder, does not produce unpleasant side-effects.

S-type processors contain five bands, arranged as two stages of pairs of h.f. bands arranged for action substitution, plus one l.f. band whose function is primarily to improve the spectral balance during "compatible" (i.e., B-type) playback.

As in C-type, spectral skewing and anti-saturation are applied to reduce the sensitivity to signals at the top of and beyond the audio range and to improve tape MOL at both extremes of the audio spectrum.

See appendix C for more information.

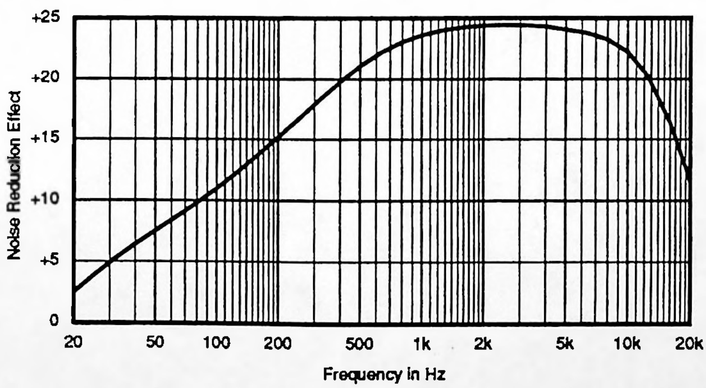


Figure 2.5 Low-level response of S-type encoder

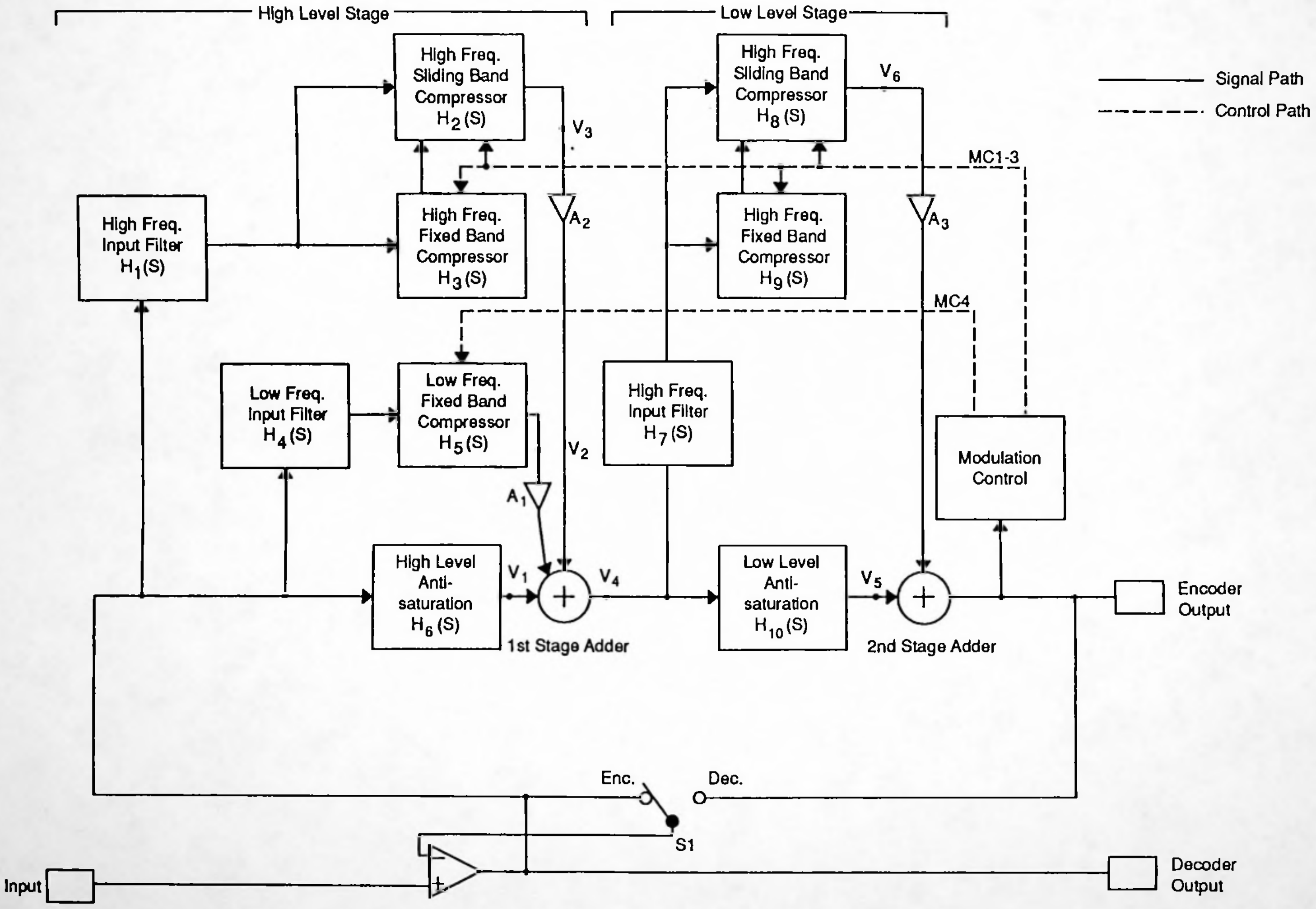


Figure 2.6 Block diagram of S-type encoder

SECTION 3
INSTALLATION

3.1 Planning the installation

The Model 422 is designed for the encoding of duplicating master tapes ("loop-bin" or "running" masters) with any of the consumer Dolby noise reduction systems, and for quality monitoring of these tapes and of subsequent duplicates. It should be considered as part of the recorder or reproducer, not as a studio tool like an equalizer, filter or compressor. For a more detailed discussion of the use of NR in duplicating see the software information manual published by Dolby Laboratories Licensing Corporation.

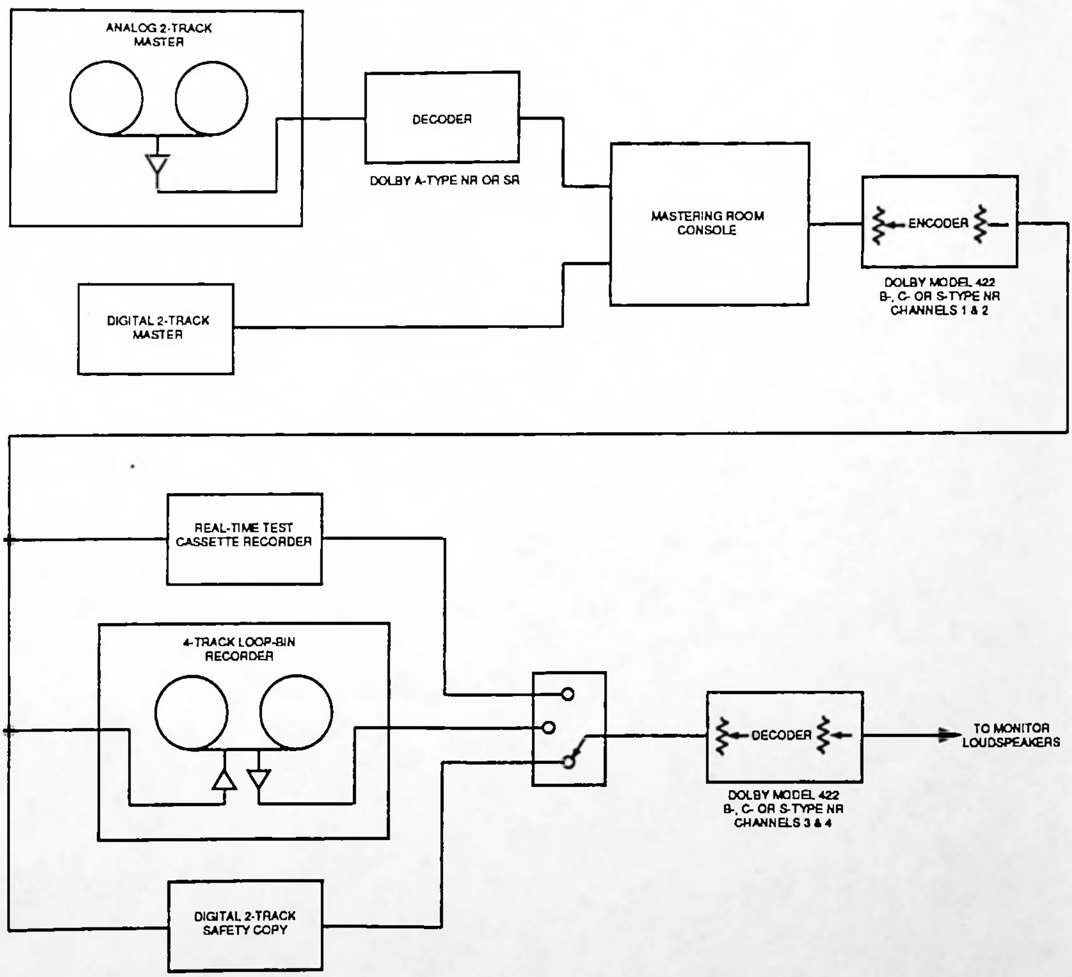


Figure 3.1 Block diagram of loop-bin mastering system

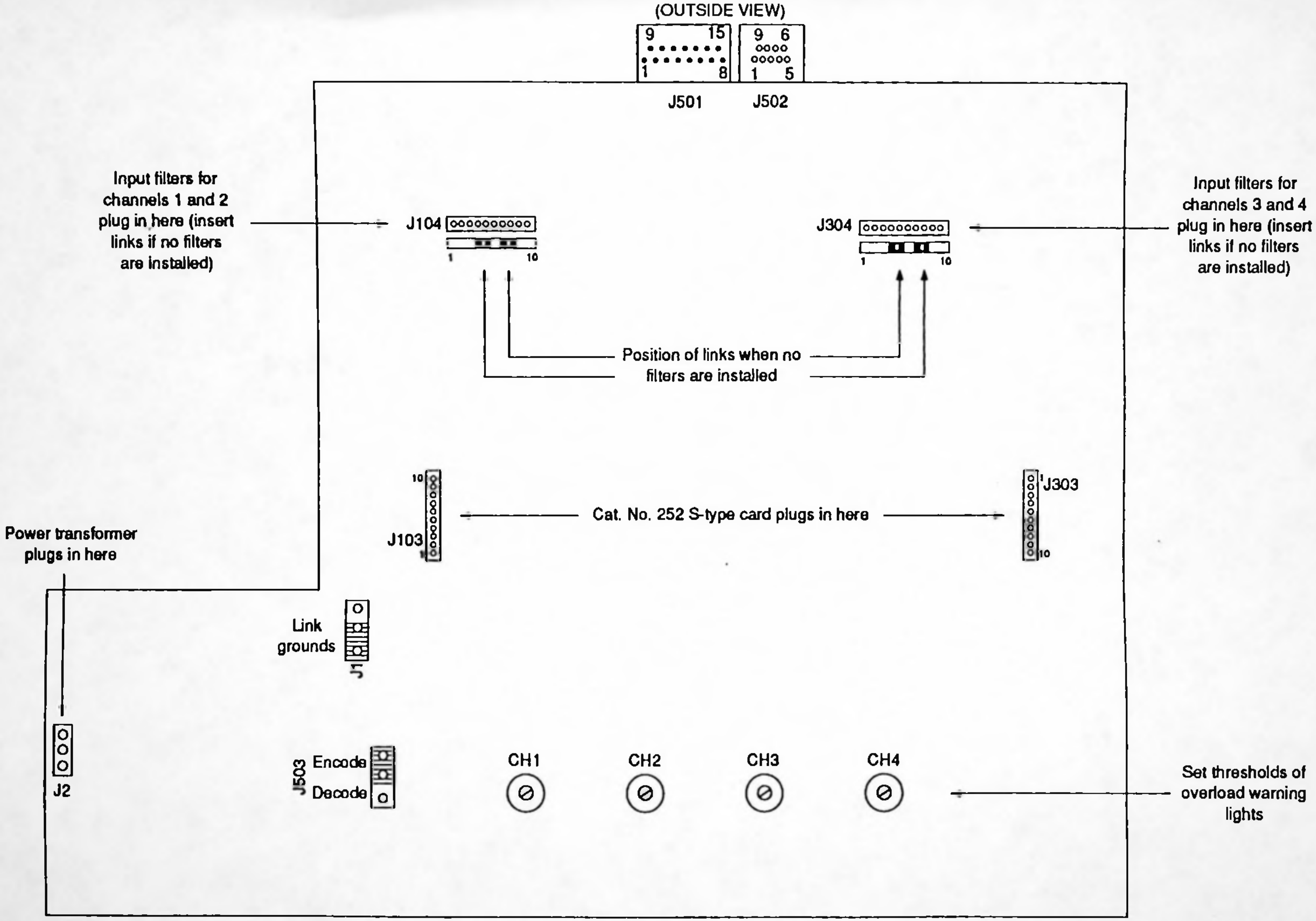


Figure 3.2 Top view of Model 422 mother-board

Figure 3.1 shows how noise reduction fits into the duplicating chain. Channels 1 and 2 of a Model 422 process (encode) the source material prior to recording on a running master, whose quality is monitored in real time by decoding via channels 3 and 4. If the final duplicate is to be an audio cassette, the running master must be turned over to record the B side; Dolby NR encoding cannot be performed backwards.

After duplication, quality assessment of the resulting cassette usually uses a second Model 422, with its channels 1 and 2 set up as decoders. One pair of channels may be used to decode the duplicate and the other pair the running master, permitting immediate comparison of the duplicate with the running master. Alternatively the other pair may decode the duplicate with a different NR system for assessment of compatibility.

3.2 Installation

The Model 422 is designed primarily for 19 inch rack mounting, but may be mounted in any plane and with any orientation.

a) Unpack the Model 422 and check for any damage. Be sure to check the packing material for the power cable and accessories.

b) The unit is shipped with channels 1 and 2 set as encoders. If you need to use them to decode, remove the top of the unit (two black screws each side) and move jumper J503 to the decode position (see figure 3.2).

c) If you need to change the input filters, do so now; they are contained on additional boards which plug on vertically to the main printed circuit board (see figure 3.2 and section 4.2). Replace the top of the unit.

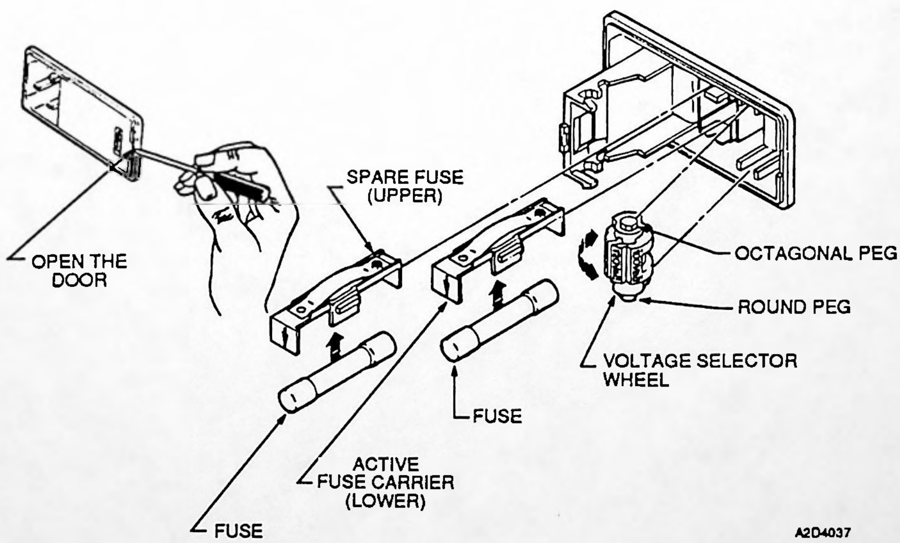


Figure 3.3

d) Open the fuse compartment door in the power supply input connector with a small flat-blade screwdriver (see figure 3.3), and check that the fuse has the correct rating. If necessary, rotate the selector drum until it displays the correct voltage for the installation. (The drum may also be removed and replaced in the desired position; it will only fit one way round.) Close the door.

e) Mount the unit appropriately as desired. Ensure that there is air flow around it, and that it is not mounted directly above other heat-producing equipment.

3.3 Connections

a) Connect audio cables to the inputs and outputs using two-conductor shielded cable. Connect the shields at one end only.

Current IEC convention calls for XLR pin 2 to be "high/hot" and pin 3 "low/cold." In a balanced system, the distinction is arbitrary provided there are no phase inversions through the unit: the Model 422 maintains polarity.

In an installation where the source and/or load are unbalanced, avoid ground loops by using two conductor shielded cable exactly as for balanced circuits; in other words, ensure that the unbalancing occurs only at the end remote from the Model 422. Note that both audio pins of the XLRs (pins 2 and 3) must be connected; neither may be left open. Since the line amplifiers of the Model 422 are floating, it is immaterial which pin (2 or 3) is treated as the "hot" side, provided the input and output are treated consistently. In the interests of maintaining international standardization, we suggest that you follow the IEC recommendation.

No output terminations (600 ohm, etc.) are necessary.

b) If required make the necessary connections to the remote control socket (15-pin D female). See appendix D.

c) Read the safety information in section 1.3. When you are confident that you have observed its provisions, connect the power cable between the Model 422 and a power outlet.

3.4 Grounding

It is normal practice to connect program (audio) ground to power line ground for many reasons, including safety. On occasion (particularly with long signal lines or incorrect wiring of the cable shields) induced hum can sometimes be reduced by separating the two grounds; link J1 provides this feature (see figure 3.2). Note that there is always a 1 kohm resistor across the link so that the audio ground is never totally isolated from the chassis ground.

The chassis is always connected to the ground pin of the power line cord; for safety reasons this ground should never be disconnected.

SECTION 4
OPERATIONAL CONSIDERATIONS

4.1 Reference level ("Dolby level")

All the Dolby noise reduction systems are complementary systems; the processing applied during playback is a mirror image of that applied during recording. In addition, the processing is dependent on both the level and spectrum of the signal being recorded. This gives flexibility to the way the processing can operate (for example, low level signals can be treated differently from high), but it means that for correct operation the levels in the playback processor (decoder) must be the same as those in the record processor (encoder).

Alignment for this equality is made easier by the built-in reference signal generator, which during alignment delivers a tone (Dolby tone) whose level has a known and precise relationship to the compression/ expansion characteristics of the NR systems, and by the 4-led calibration displays which indicate (by equal brightness of green leds) the presence of an input signal at that precise reference level.

For B-, C- and S-type, the tone consists of a 400 Hz sinewave, periodically modulated in frequency; the modulation depends on the NR system in use, so a stretch of Dolby tone at the beginning of a recording can not only permit accurate alignment of the decoder but also identify which system was used to encode the tape.

The actual modulations are as follows:

B-type	The pitch is raised about 10% for 15 ms every 0.5 s.
C-type	The pitch is raised briefly every 1.1 s, giving a click-like sound.
S-type	Like B-type except that the pitch is raised twice every 0.5 s.

The Dolby tone generator is turned on by pressing the front panel push-button switch marked Set-up.

Since there will generally be at least one generation of tape recording between NR encoding and decoding, the magnetic flux level on the duplicate tapes corresponding to this reference level must also be defined. See the table below for the fluxivities on various formats.

Although all three NR systems have nominally unity gain at 400 Hz and Dolby level, this is subject to a tolerance of about ±0.5 dB. All adjustments should be made with NR out, or with both NR and Set-up pressed in (when Set-up is pressed, the processors are disabled as if NR were out). Alignments made with one system will then be correct for all.

Format of duplicated tape	Fluxivity for Dolby level
Compact cassette	200 nWb/m ANSI
VHS (linear tracks)	120 nWb/m DIN
Consumer open reel	185 nWb/m ANSI

Note that the two standardized methods of measuring tape fluxivity, DIN/IEC and ANSI, give results which differ by almost 1 dB; thus for example 200 nWb/m ANSI is equivalent to about 218 nWb/m DIN.

4.2 Input filtering

Both the encoder and decoder of a Dolby noise reduction system generate their control signals from the audio passing through them. Hence over the middle range of levels where compression and expansion occur it is important that the audio signal in the decoder be substantially the same as in the encoder. Otherwise the NR system will "mistrack," and the decoded output will differ from the source.

The most common potential source of error in a correctly adjusted system is loss of high frequencies due to a bandwidth limitation. For example consider the usual case where duplicated audio cassettes are played on a typical consumer home cassette recorder with a combined record/playback head having a wide gap, or on an automobile player using an inexpensive playback head. The playback response will inherently be restricted to 15 kHz or so. If the duplication chain is capable of recording signals up to 20 kHz or more, mistracking will occur whenever the source material contains significant energy above 15 kHz. The problem may be compounded by differences in azimuth setting between recorder and reproducer.

With B- and C-type NR this mistracking may be audible because the effect of mistracking extends to lower frequencies than those which were lost in the bandwidth limiting. S-type is much more resistant to audible mistracking.

The mistracking can be eliminated by filtering the input to the NR encoder so that it does not contain any frequency components which will not be reproduced accurately into the decoder. However, it is obviously impossible at the time of duplication to predict the performance of the equipment on which the duplicate will be played. Low-pass filtering at say 5 kHz would almost certainly prevent any mistracking, but clearly the material would be audibly impaired. At the other extreme, material with energy extending to 30 or 40 kHz (e.g., a live microphone close to a high frequency percussion instrument) recorded without any deliberate filtering would almost certainly lead to audible mistracking when a duplicate was played at home.

Practical experience and common sense indicate that the best compromise is to limit the bandwidth to no more than that which permits reproduction without audible impairment. It has been shown (e.g., "Which bandwidth is necessary for optimal sound transmission?", Plenge et al., JAES March 1980) that a filter flat to 15 kHz (within say 0.5 dB) and falling rapidly thereafter does not introduce audible impairment (while one at lower frequencies is audible). Fortunately this frequency corresponds closely to the bandwidth of typical consumer cassette equipment, and such a filter is therefore usually sufficient to remove audible mistracking.

Units to be used for duplication of Philips compact cassettes with B- or C-type NR should contain Dolby Laboratories' plug-in 6-pole low-pass filter (Cat. No. 435), which is substantially flat to 15 kHz but falls at an ultimate slope of 36 dB/octave at higher frequencies; the Model 422 is normally delivered with this filter in place. This filter is also generally adequate to remove the ultrasonic spurious signals which emerge from some digital audio equipment and which can disturb the operation of the NR encoder. Note that use of a filter with a higher cut-off frequency will not improve the audible quality but will increase the probability of mistracking and hence of complaints from customers. It is permissible to bypass this filter (see below and appendix D) when encoding with S-type NR.

Most applications involving television (e.g., duplication of VHS cassettes with B-type NR on the stereo linear tracks) yield audio corrupted by small amounts of horizontal (line) frequency. For such applications we strongly recommend replacing the plug-in low-pass filter in the encoder by a notch filter at the horizontal frequency. Dolby Laboratories' Cat. No. 370/371 provides a suitable notch plus low-pass filtering designed for this purpose.

All these filter cards contain two channels of filtering and plug into the mother-board of the Model 422 (see figure 3.2 in section 3). If you need the video filter you should order it with the appropriate notch frequency; the Cat. No. 370 is at 15734 Hz for 30 frame/s sytems and the Cat. No. 371 at 15625 for 25 frame/s.

The filtering can be bypassed for test purposes by applying a ground to pin 4 of the remote control connector J501 (see appendix D).

It is possible to connect external filters and/or metering to the auxiliary connector on the rear of the Model 422 (9-pin D-connector); if you want to use this facility, contact Dolby Laboratories.

SECTION 5
ALIGNMENT OF MODEL 422

5.1 Alignment for preparation of running masters

See figure 3.1 in section 3. For a detailed discussion of the requirements see sections IV and V of Dolby Laboratories' software information manual.

a) Preliminaries

You must know in advance the fluxivity on the running master required to give a tone on the final duplicate at the level shown in the table above (paragraph 4.1). This may involve some calculation; see section IV paragraph 2.2 of the software information manual.

Correct operation of the noise reduction requires also that you know the optimum program level on the duplicate for a particular item of speech or music before the running master is prepared. It is a good idea to evaluate the audibility of noise or distortion on the duplicate by using a real-time recorder in the desired format (e.g., audio or video cassette recorder, with the appropriate formulation of tape) in place of the master recorder, as shown in figure 3.1. This optimum recording level must be arrived at by adjusting the console output, not the output of the Model 422 nor the input gain control of the real-time recorder. When the running master is recorded, the optimum program level on the final duplicate will then be obtained automatically.

In a permanent installation mastering for a known formulation, the overload point of the tape with respect to Dolby level will be known (this is part of the information required to determine fluxivity on the running master, per section IV of the software manual); the thresholds of the overload leds of the Model 422 may be adjusted by means of potentiometers RV103, etc. (see figure 3.2) so that the leds come on at this level.

If a change of formulation for the duplicates demands a different program level, a new running master should be made.

b) Before adjusting the Model 422

Switch the Model 422 into Bypass (middle push-button switch, red light will flash). Line up all the other equipment carefully. Experience has shown that whenever Dolby noise reduction is not working satisfactorily, the problem is usually the result of inaccurate alignment of other links in the chain.

When you are satisfied with the performance, release the Bypass button.

c) Adjustment of encoders

Press Set-up (left-hand push-button, yellow light will flash) to turn on the Dolby tone generator.

Switch to the desired NR system (toggle switches for each pair of channels), and press the NR switch in (right-hand push-button which controls all channels). This feeds out the Dolby tone appropriate to the selected NR system.

Adjust the outputs of the Model 422 encoders (channels 1 and 2) to get the predetermined fluxivity on the running master, and record a few seconds of Dolby tone.

Release both the Set-up and the NR buttons.

Feed in a 1 kHz tone from the console at any convenient level.

Adjust the inputs of the encoders to give unity gain within the Model 422. (If the impedance of the source is low, 50 ohm or less, and that of the load is high, 10 kohm or more, this can be done most easily by switching in and out of bypass with the front panel push-button, and adjusting for no change in output level.)

Finally press NR in.

d) Adjustment of decoders

Channels 3 and 4 are used to decode the reproduced running master while you are recording it, so that you can listen to it and compare it with the source, or can assess compatibility of the material when reproduced with a different NR system.

Ensure that both the Set-up and NR buttons are released (out).

Play the part of the tape containing Dolby tone (recorded as in paragraph c) above) and adjust the inputs of channels 3 and 4 to make the green leds on the calibration displays equally bright.

Adjust the decoder outputs for unity gain through the Model 422, or for the desired output line level.

(Alternatively, if you are aligning channels 3 and 4 while recording and simultaneously playing back Dolby tone from channels 1 and 2 as in c) above, leave both Set-up pressed and NR in while you carry out these adjustments.)

Select the desired type of NR and press the NR button in.

5.2 Alignment of decoders for quality assessment of duplicated tapes

Switch NR out (or press both Set-up and NR in).

Play a reference level tape (i.e., a tape known to be recorded at the appropriate fluxivity listed in the table above) and adjust the inputs of the Model 422 decoders for "equal greens." Adjust the outputs either for unity gain or for the desired output line level.

Select the desired type of NR and press the NR button in.

SECTION 6
CIRCUIT DETAILS

6.1 General

The S-type processors are contained on a plug-in printed circuit board, Dolby Laboratories' Cat. No. 252. Various filters (Cat. Nos. 370, 371 or 435) also plug in. All the rest of the circuitry is on an easily removed mother-board, with the input and output connectors along its rear and switches and indicator leds along the front.

Much of the electronics can be serviced by the customer or distributor; however if a fault develops in the immediate region of the Dolby B/C processors or on the Cat. No. 252 S-type processor board, the unit and/ or Cat. No. 252 as appropriate must be returned to Dolby Laboratories for repair.

Notes on dismantling the unit can be found in section 7.

The following descriptions apply to channel 1 or where appropriate to channels 1 and 2 where they are treated as a pair. Channels 3 and 4 are substantially identical except that they are permanently in the decode mode and do not deliver Dolby tone when setup is pressed. Refer to figures 6.1 to 6.4 for circuit diagrams.

6.2 Input and output line amplifiers

Each input amplifier has a precision differential configuration of three operational amplifiers (IC101A, B and C) using 0.1% tolerance resistors in the critical positions. Fets QF101/102 are normally held off by a -22 V line. When the unit is not powered, these fets turn on and act as a short-circuit to prevent the non-linear distortion which would arise if large signals were applied to the input operational amplifiers in their unpowered state; the input impedance then drops to about 10 kohm. Diodes D101-D104 protect against damage due to excessive input swings, while diodes D105 and 106 stop the outputs of IC101A and B from going rapidly to the positive rail if their non-inverting inputs are driven beyond the negative rail.

The input signal then passes via the input gain control RV101 and a three-pole low-pass filter at about 35 kHz (IC101D, etc.) to the sockets for extra input filters, internal (J104) or external (9-pin D-connector J502). If no extra filters are used, jumpers (JMP101, etc.) should be placed on J104 as shown on the silkscreen of the mother-board. (The unit will work without these jumpers, but the crosstalk will be degraded.)

A "hard" bypass is provided by relay RL101 which passes the input signal directly to the output XLR connector when either the bypass button is pressed or the power is removed. Transistor Q1 and associated parts ensure that the relay is not energized unless the main power rails are high enough for the output amplifiers to work; this prevents excessive clicks at the output as the unit is turned on or off.

The output amplifier uses four operational amplifiers IC108 and 109 in a balanced floating configuration to simulate a transformer. The differential output resistance is 22 ohm, but the output impedance of each leg to ground is many kohm. When the external load is unbalanced, either leg may be grounded (and one must be). IC110B acts as a servo to define the dc potentials at the outputs. The amplifier is preceded by an output gain variable resistor RV102.

6.3 B- and C-type processors

Channels 1 and 2 use a dual integrated circuit (Sony CXA1330, IC106) as a unity-gain B- or C-type encoder. Three-state logic applied to pin 26 determines the mode of operation: ground = NR out, open = B-type and high = C-type. The chip is powered from the +15 V rail only, so coupling capacitors are needed at the input and output (C117 and C120). In the encode mode, IC105A acts as a unity-gain buffer, but in the decode mode the processor is connected as a negative feedback path round this amplifier, thereby converting an encoder into a decoder. IC104A and B perform the necessary switching to rearrange the feedback and to select the appropriate output point.

The calibration display and overload led (see paragraph 6.4) are fed from the output of IC105A. To ensure that the display does not light up in the bypass condition, the input to this amplifier is muted by IC104C when the bypass button is pressed.

When the NR type switch is in the S-type position and NR is in, the signal path via IC305A and IC106 is broken and the output of the S-type printed circuit board plugged into J103 is selected by IC107A. In the combination encode mode and NR in and setup pressed, IC107A selects instead the output of the Dolby tone generator.

If S-type NR is not installed (no Cat. No. 252), but S-type is selected with the toggle switch, the Model 422 defaults to B-type, with the B-type led illuminated.

When power is first applied, IC107 is held off (i.e., the output is muted) for roughly a second by C427 and R483 to allow the coupling capacitors around the B/C NR processor to charge.

6.4 Calibration display

IC111B and C form a full-wave averaging rectifier which feeds an array of four window detectors (IC112) driving a string of four leds (DS102A-D). Resistor R179 is selected at the factory so that when the signal at the processors is at Dolby level, -6 dBr or 388 mV rms, the two green leds are equally bright. The lower red led lights if the signal is within about 10 dB below Dolby level and the upper lights if the signal is above Dolby level.

IC111B and A form a peak hold circuit whose output is applied to comparator IC111D; the reference for this comparator can be adjusted over the range 2 to 10 dB above Dolby level by means of internal potentiometer RV103. When the peak signal exceeds the reference, led DS101 lights.

6.5 Dolby tone generator

The required 400 Hz sinewave is generated by a phase-shift oscillator. The approximately sinusoidal output of the phase-shift network, R606-608 and C603-605, is amplified by IC603C and applied to comparator IC603D, whose output controls CMOS switch IC602A. The input of the phase-shift network therefore receives a square-wave with a defined amplitude of 15 V peak to peak symmetrical with respect to ground. Resistor R610 is selected to give precisely 388 mV rms at the output.

C602 and R627 alter the frequency at which the phase-shift network has a 180 degree shift. These components are switched in and out of circuit by Q601 to give the required frequency modulation (with no amplitude change).

The modulating waveforms for B- and C-type NR are generated by two conventional astable multivibrators, IC601D and IC601C. The S-type waveform comes from monostable multivibrator IC601A which is triggered by both the positive and negative edges from astable multi IC601B.

CMOS gates IC602B and C select the appropriate waveform corresponding with the position of the toggle switch controlling the type of NR in channels 1 and 2.

6.6 Power supply, etc.

The power transformer has two 120 V primaries, one of them tapped at 100 V. The voltage selector connects them as follows:

100 V	120 V windings in parallel, power fed between 0 and 100 V
120 V	120 V windings in parallel, power fed between 0 and 120 V
220 V	120 V winding in series with 100 V winding
240 V	120 V windings in series

The power supply uses a conventional center-tapped secondary feeding full-wave rectifiers and 15 V linear regulators. The bypass relays are powered from an unregulated nominal 22 V which is turned off by Q1 if the potential on C1 is less than about 11 V. The unregulated negative supply is smoothed by R6 and C13 and used to hold off the fets shunting the line input amplifiers.

Reference diode DZ601 and IC603A and B generate temperature-independent nominal ±7.5 V rails for some of the logic and for the Dolby tone oscillator.

Astable multivibrator IC604A delivers a square-wave at between 2 and 3 Hz to make leds flash as required.

6.7 S-type board, Cat. No. 252 (figure 6.5)

The S-type processors are in the form of small surface-mount circuit boards, one for each of the four processors, carried on a larger card which plugs into the Model 422 mother-board at J103 (channels 1 and 2) and J303 (channels 3 and 4). Each processor is fed via an inverting unity gain buffer (IC101A, etc.) and the record or playback output is selected as required by IC102A, etc. The mode of operation of the processor itself is determined by the voltage on pin 24 of the processor board, where -6 V gives play and +6 V gives record. All ICs except the input buffer op amps run on + and -6 V supplies generated by regulators IC1 and IC2.

For fault-finding, the whole Cat. No. 252 assembly can be turned 180 degrees, thereby interchanging pairs of channels. If a fault appears to be on one of the little S-type processor boards, the whole Cat. No. 252 must be returned to Dolby Laboratories.

6.8 Plug-in filters (figures 6.6 and 6.7)

Each plug-in filter card carries two channels of filtering. The filters are normally used at the input of encoders (i.e., for channels 1 and 2 only), when a card is plugged into J104 of the Model 422 mother-board (see figure 3.2). However there is provision for inserting filters into decoding channels 3 and 4 at J304. Note that when no filter is inserted, jumpers should be plugged on to J104 and J304, as shown in the silkscreen of the mother-board.

The filters can be switched out of circuit by applying a ground to pin 4 of the remote control connector J501.

a) Standard filter (Cat. No. 435)

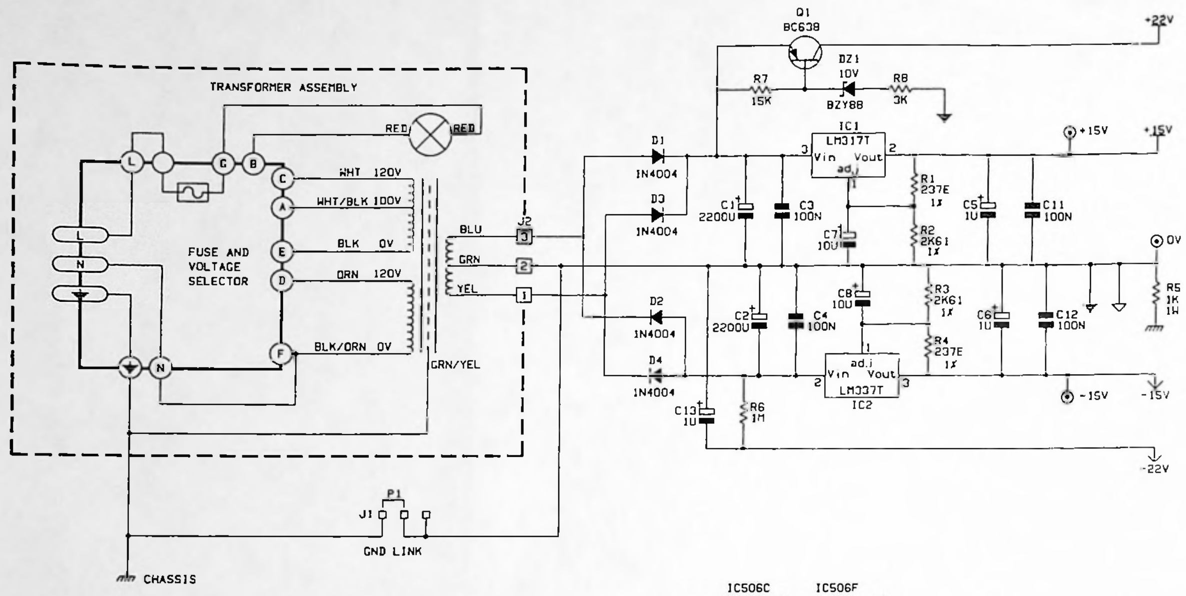
The model 422 is normally delivered with a 6-pole low-pass filter installed on J104. This filter is substantially flat to 14 kHz and 3 dB down at 16 kHz, matching the bandwidth of typical consumer cassette playback systems, and is appropriate for use in the duplication of compact cassettes. See figure 6.6 for a circuit diagram.

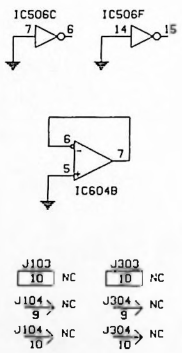
b) Cat. No. 370/371

For applications involving television (notably duplication of VHS cassettes), the standard filter should be replaced by a Cat. No. 370 or 371. These consist of a three-pole low-pass filter plus a deep notch at television horizontal (line) frequency; the precise frequency of the notch is adjustable over a small range by means of a multi-turn pot. The frequency will be set in the factory in accordance with the order from the customer. The filter is designated Cat. No. 370 if set to 15734 Hz (30 frame/s NTSC) and Cat. No. 371 if set to 15625 Hz (25 frame/s).

The response of this filter is ±1 dB up to 13.5 kHz, -3 dB at 15 kHz and falling rapidly thereafter, with a notch at least 50 dB deep at the horizontal frequency.

See figure 6.7 for a circuit diagram.



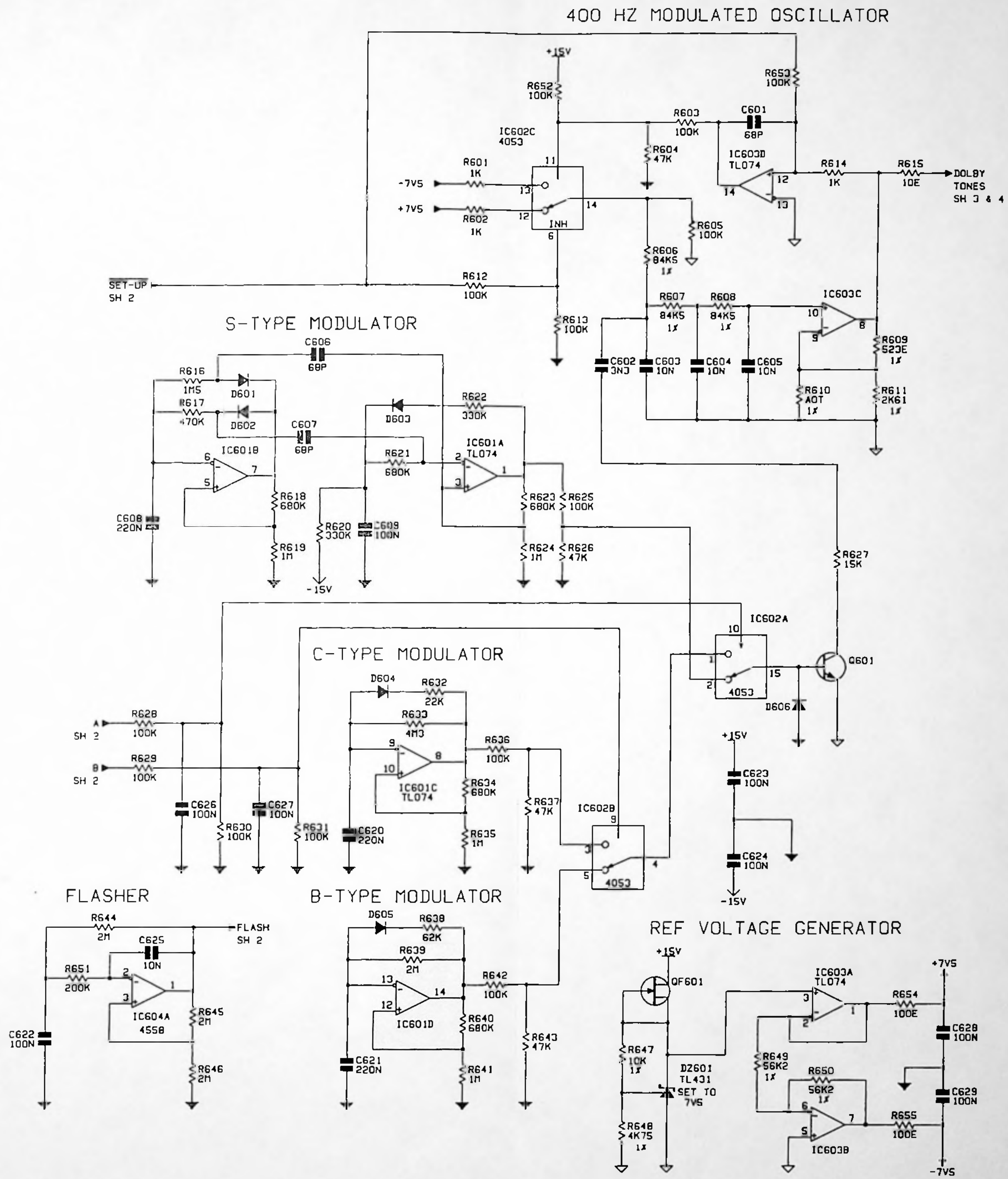


SUPPLY TABLE

REF DESIGNATOR	DEVICE TYPE	LGND	+7VS	-7VS	+15V	-15V
IC101,111,201, 211,301,311, 401,411,601, 603	TL074				4	11
IC102,103,105, 110,202,205, 302,303,305, 310,402,405	TL072				8	4
IC104,204,304, 404,602	4053	8	16	7		
IC107,307	4052	8	16	7		
IC108,208,308, 408,604	4558				8	4
IC109,209,309, 409	4556				8	4
IC112,212,312, 412	LM324				4	11
IC501,506	4049	8			1	
IC502,503	4052	7,8			16	
IC504	4053	7,8			16	
IC505	4011	7			14	

NOTES: (UNLESS OTHERWISE SPECIFIED)

1. RESISTOR VALUES ARE IN OHMS, 5%.
2. CAPACITOR VALUES ARE IN FARADS.
3. DIODES ARE 1N4148.
4. BIPOLAR TRANSISTORS ARE DOLBY STANDARD DEVICES:
   PNP = BC416, 2SA970 OR SIMILAR
   NPN = BC414, 2SC2240 OR SIMILAR
5. FIELD EFFECT TRANSISTORS ARE DOLBY STANDARD DEVICES:
   N-CH = 2N5458 OR SIMILAR
6. [ ] DENOTES COMPONENTS EXTERNAL TO MAIN BOARD.
6. [ ] DENOTES COMPONENTS NOT INSTALLED.
7. FP DENOTES FLAMEPROOF.



A1C3913 REV 1

Figure 6.1 Power supply, Dolby tone oscillators

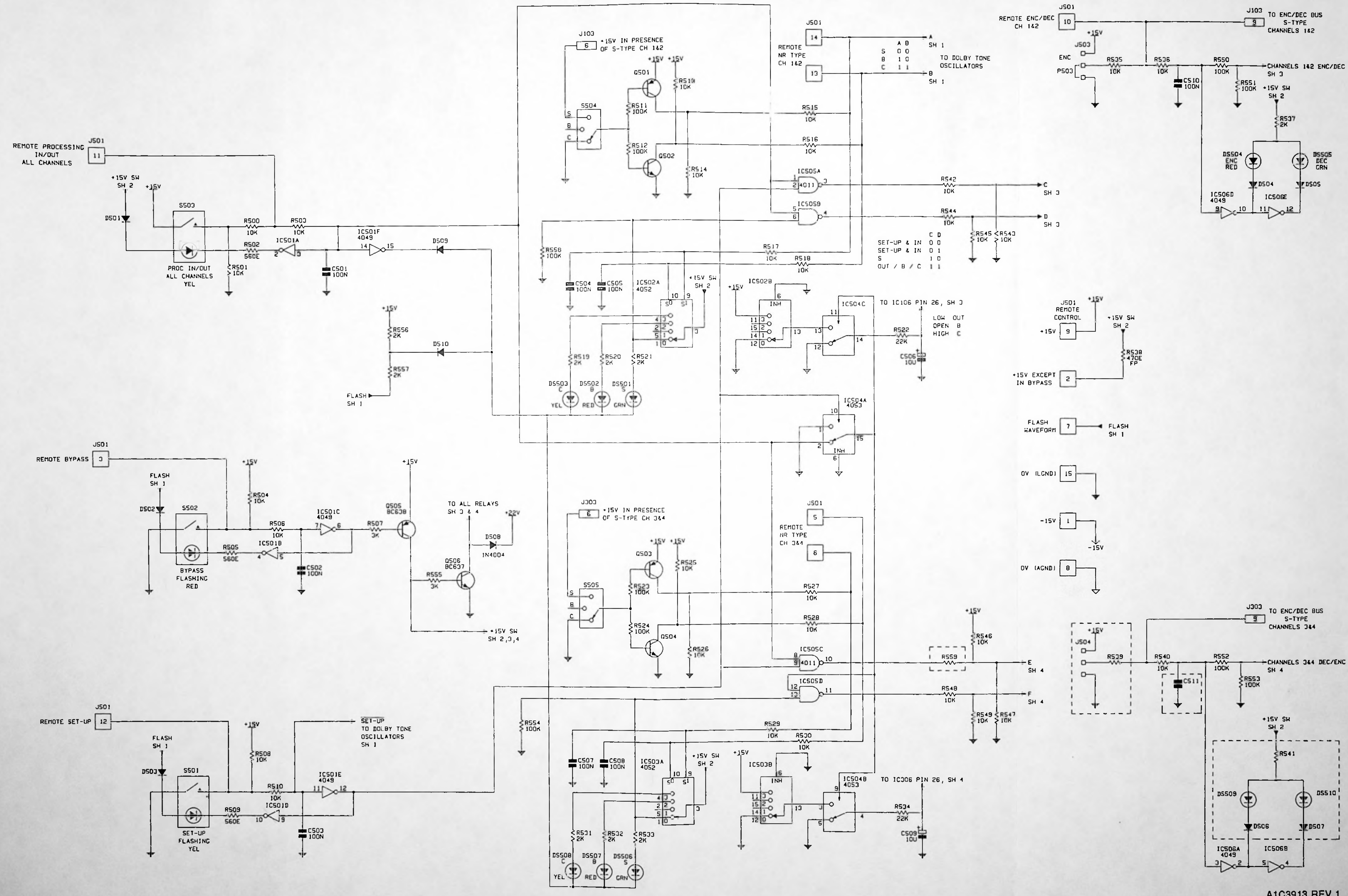
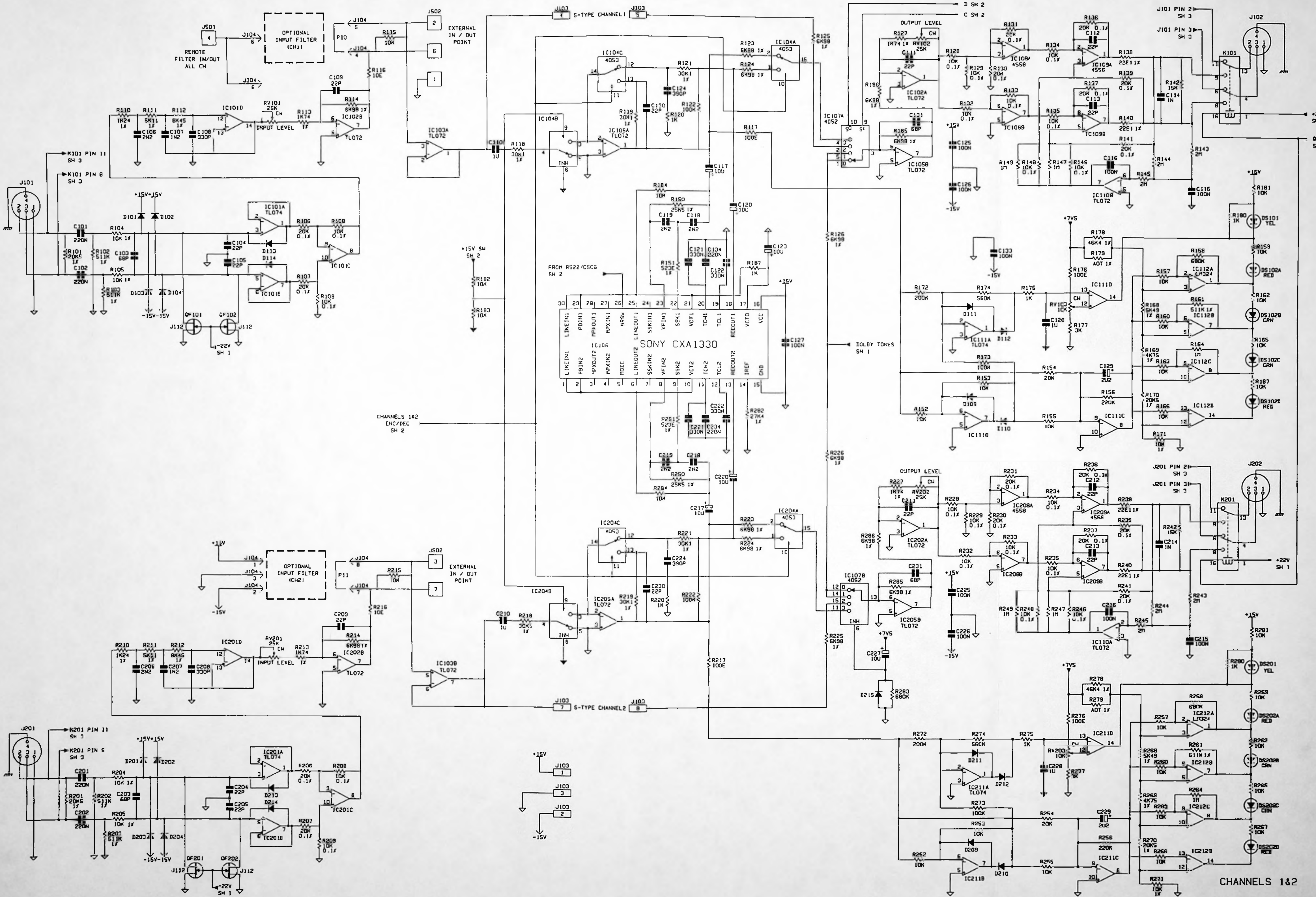
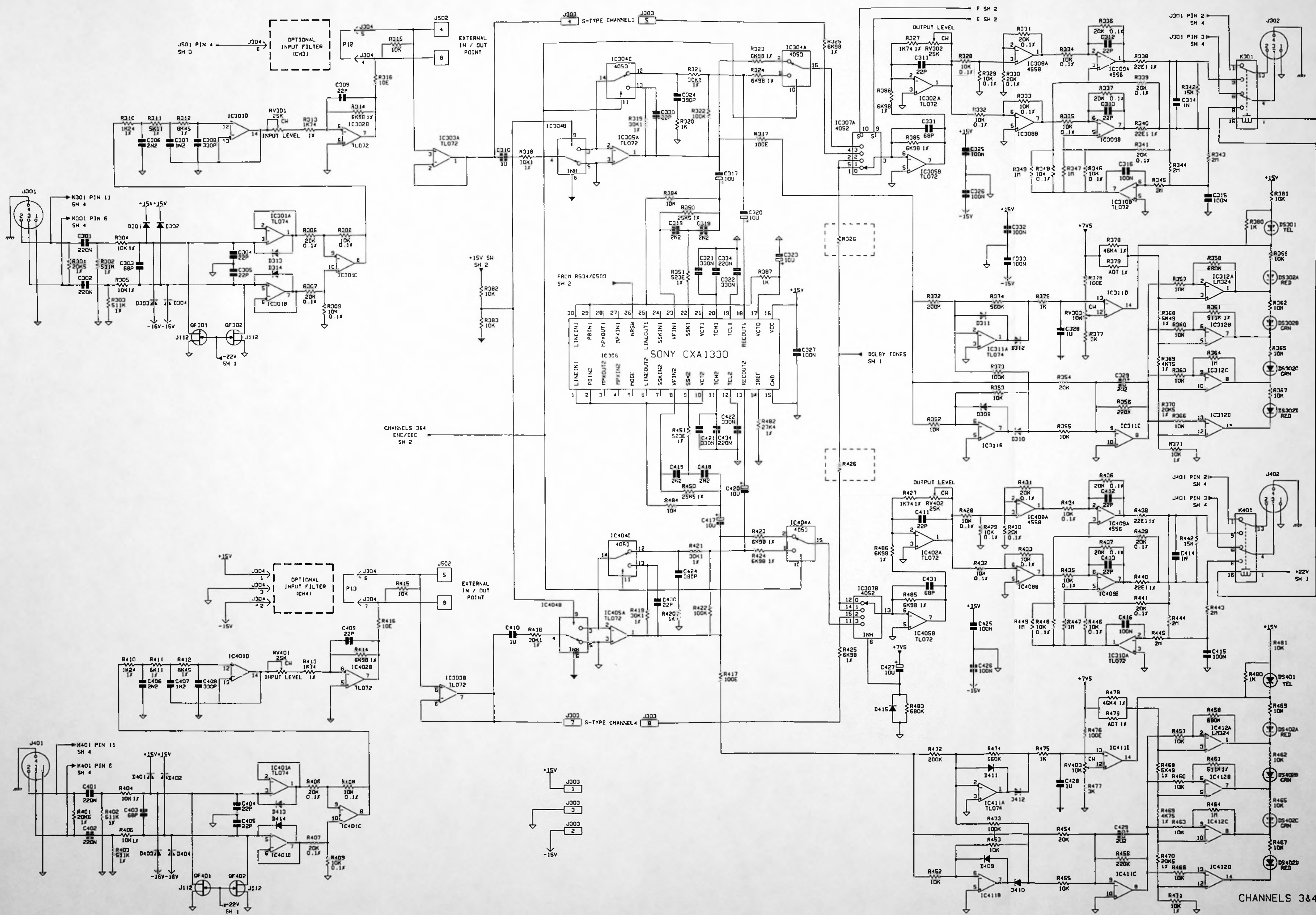


Figure 6.2 Control logic



A1C3913 REV 1

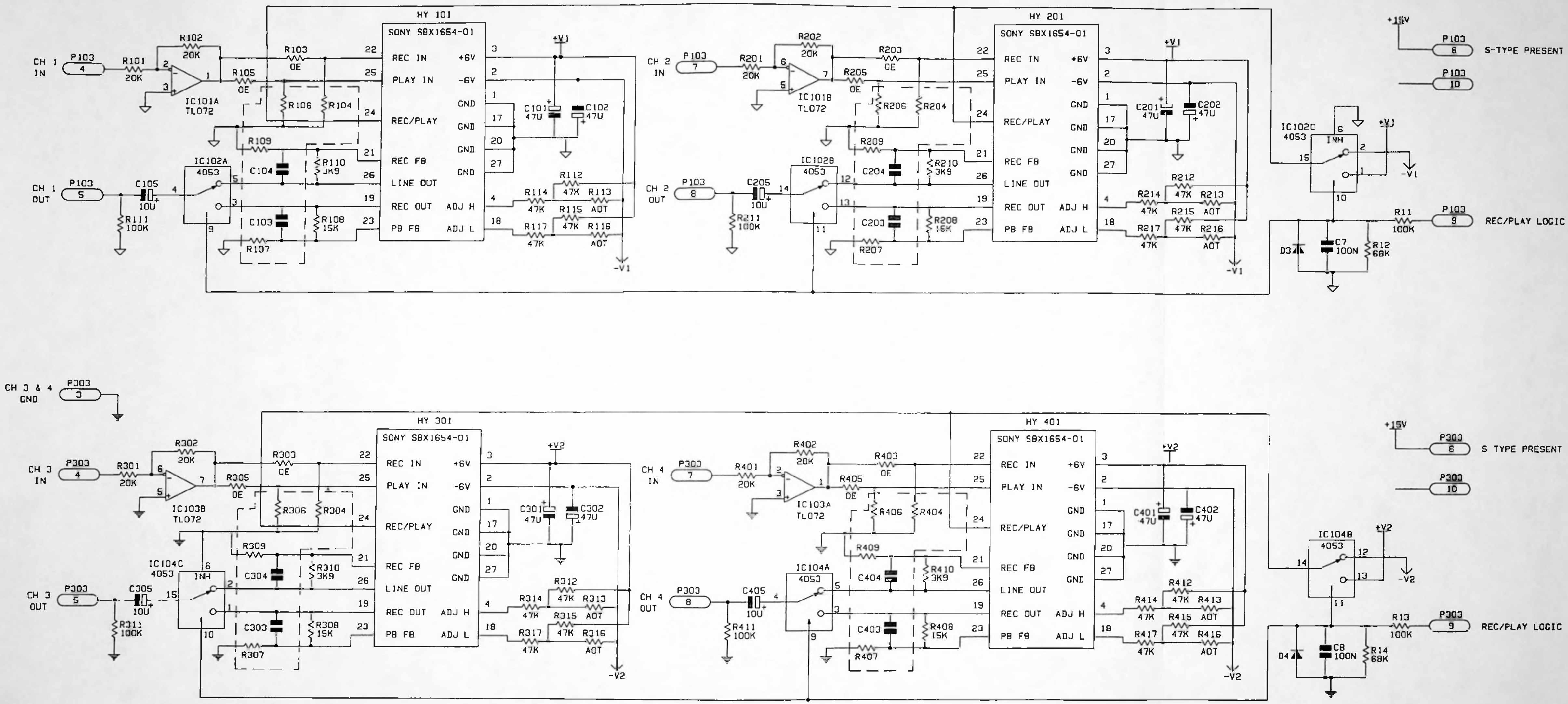
Figure 6.3 Channels 1 and 2



CHANNELS 3&4

A1C3913 REV 1

Figure 6.4 Channels 3 and 4

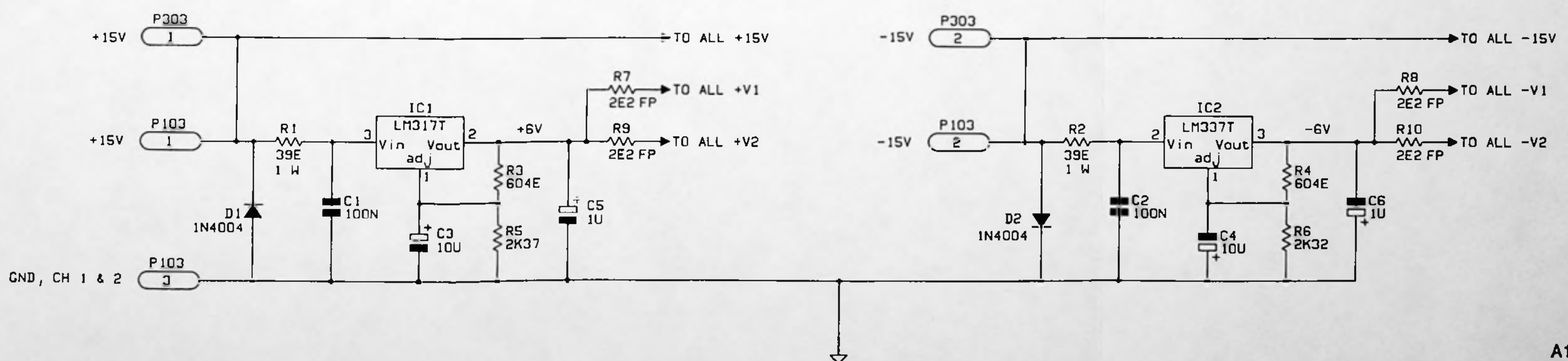


NOTES: (UNLESS OTHERWISE SPECIFIED)

1. RESISTOR VALUES ARE IN OHMS.
2. CAPACITOR VALUES ARE IN FARADS.
3. DIODES ARE 1N4148.
4. FP DENOTES FLAME PROOF
5. DENOTES COMPONENTS NOT INSTALLED.

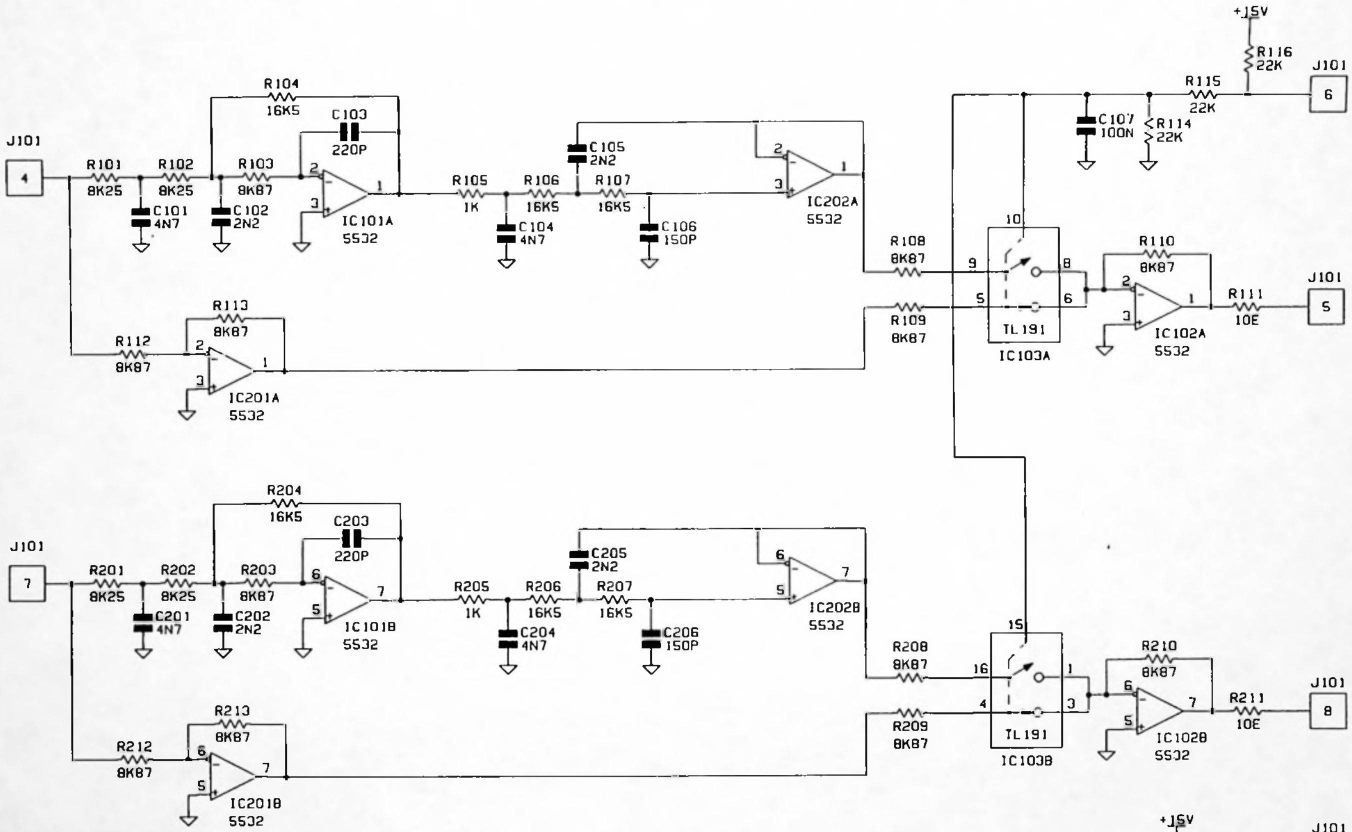
SUPPLY TABLE

REF DESIGNATOR	DEVICE TYPE	A GND	D GND	+15V	-15V	+V1	-V1	+V2	-V2
IC101, 103	TL072	—	—	8	4	—	—	—	—
IC102	4053	6,8	—	—	—	16	7	—	—
IC104	4053	—	6,8	—	—	—	—	16	7



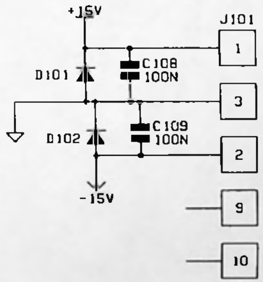
A1C4075 REV 1

Figure 6.5 Cat. No. 252—S-type board



NOTES: (UNLESS OTHERWISE SPECIFIED)

1. RESISTOR VALUES ARE IN OHMS.
2. CAPACITOR VALUES ARE IN FARADS.
3. DIODES ARE 1N4003



SUPPLY TABLE

REF DESIGNATOR	DEVICE TYPE	GND	+15V	-15V
IC101,102,201,202	5532	-	8	4
IC103	TL191	13	11,12	14

A2C4159 REV 1

Figure 6.6 Cat. No. 435—Standard filter

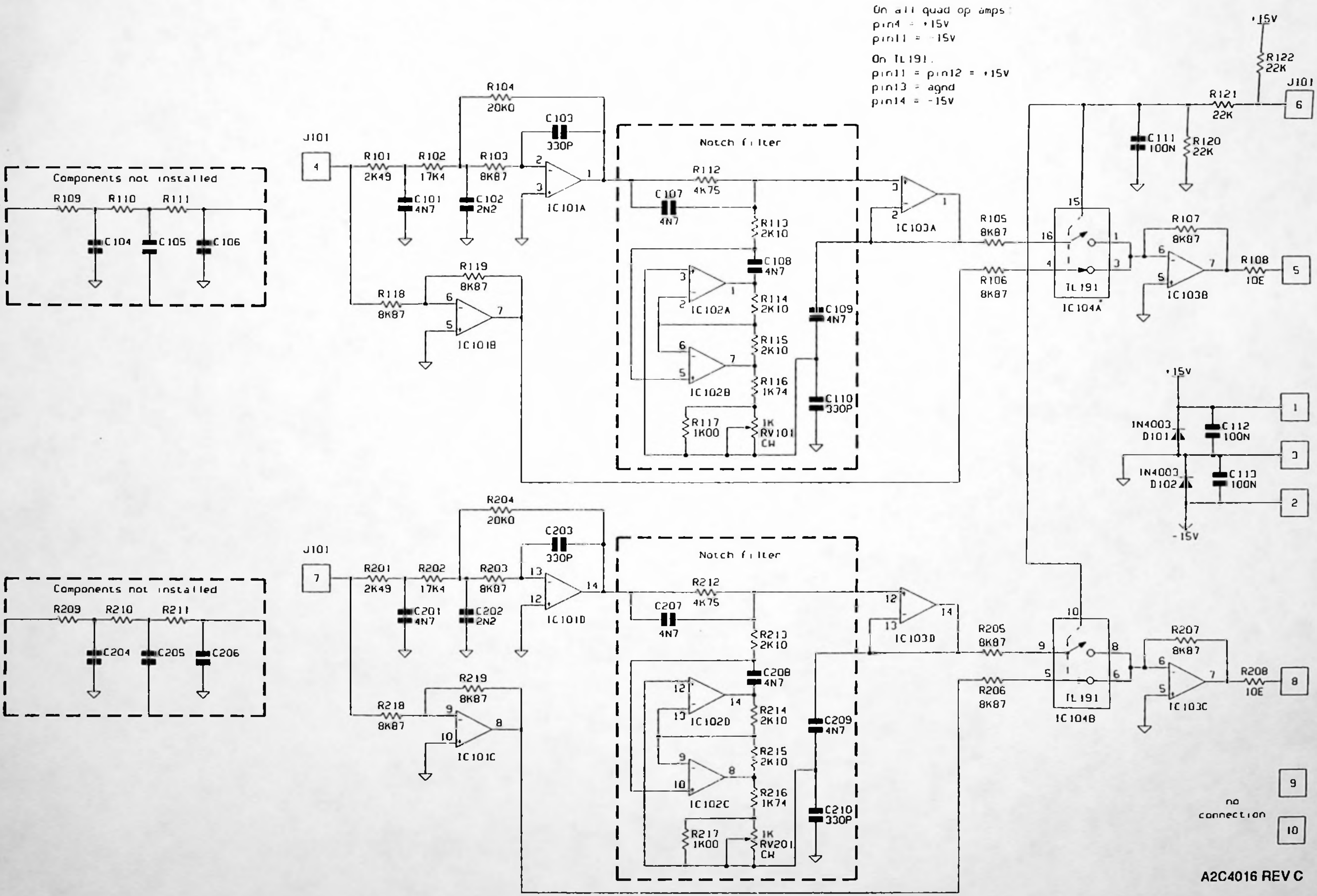


Figure 6.7 Cat. No. 370/371—Filter for television

SECTION 7
SERVICING

7.1 Disassembly (see Figure 7.1)

1. Disconnect power from the unit and remove the top cover (2 screws and washers on each side, 10 and 16).

2. Unplug the three-pin connector P2 linking the power transformer to the mother-board.

3. Turn the unit over and remove the two screws holding the heat-sink to the chassis (13). Do not remove the mounting screw for the transformer.

4. Remove the three screws (13) holding the front panel to the main chassis.

5. Turn the unit right-side up and remove the two screws (14) and crinkle washers (17) holding the extreme left and right sides of the front panel to the chassis sides (counter-bored holes).

6. Remove the front panel from the chassis.

7. Unlock each XLR body from its shell. One of three types may be used. You can deduce which you have from figure 7.2. Note the direction of turning to unlock. For the Neutrik connectors a very narrow screwdriver is needed, about half the diameter of the hole.

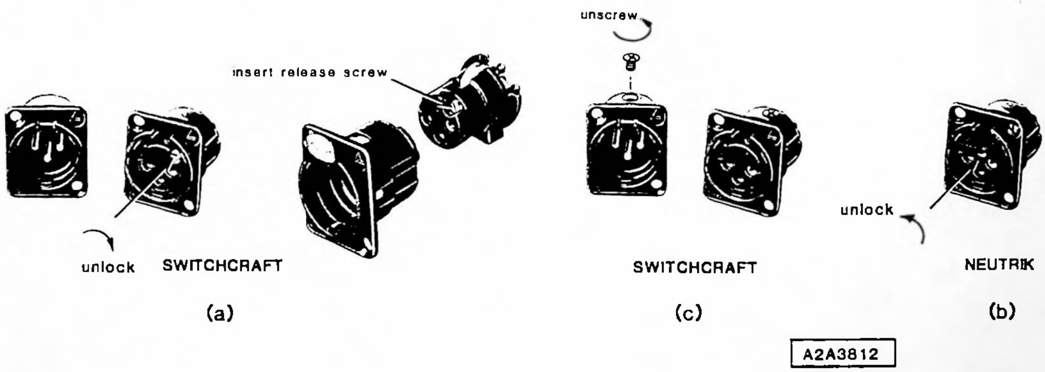


Figure 7.2

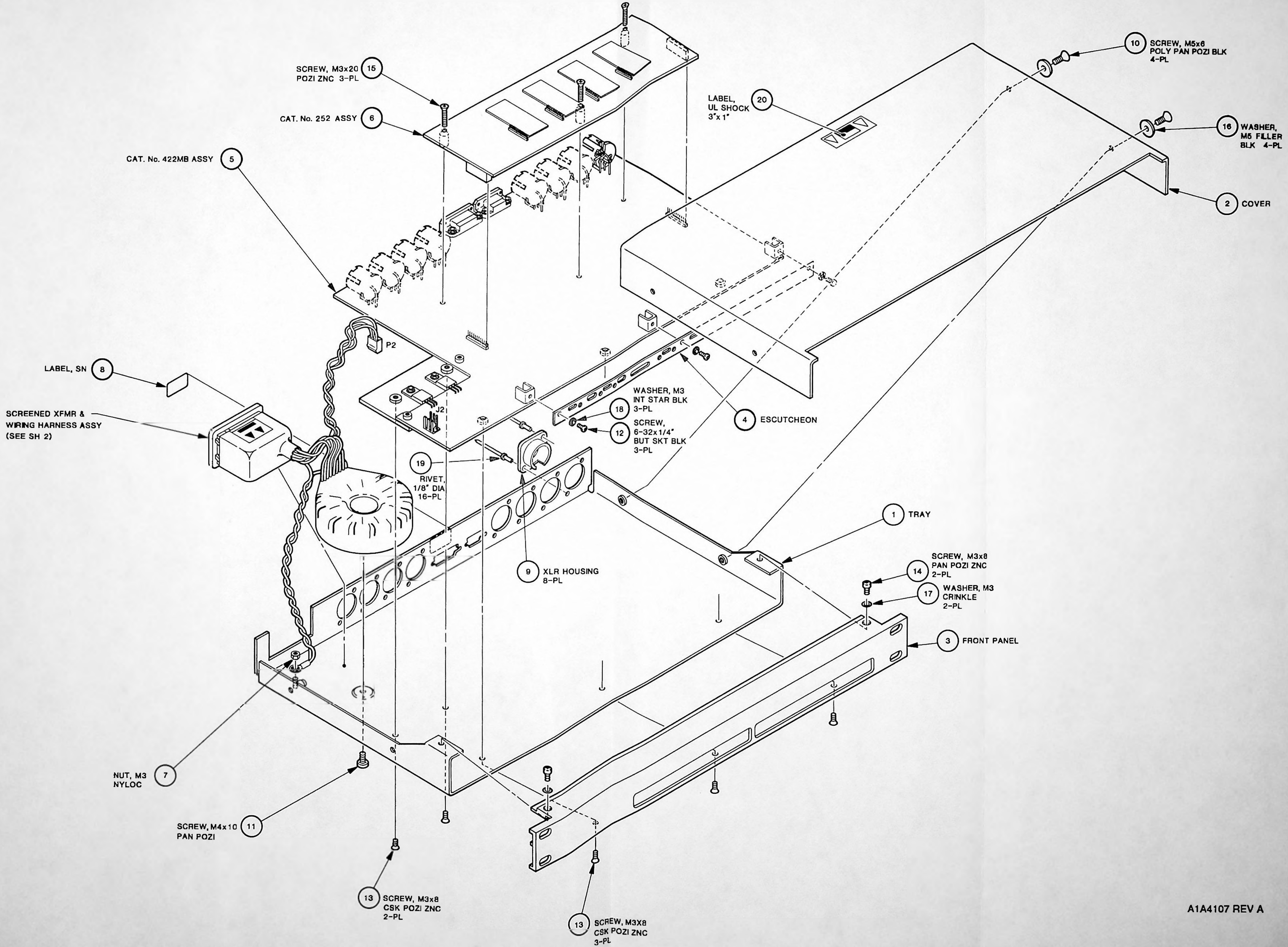
8. The mother-board can then be withdrawn through the front of the tray (do not try to lift it out vertically). Once out, it can be reconnected to the transformer.

   CAUTION: Do not leave the mother-board powered by the transformer outside the chassis for long periods; the heatsink will become very hot and the regulators may shut down, preventing further troubleshooting.

   To avoid any danger of over-heating during long-term fault-finding, feed in ±18 V from a bench power supply in place of the transformer at J2 (the middle pin is ground and the polarity is unimportant because the rectifiers will route the power appropriately).

9. If necessary the complete transformer assembly can removed by unplugging the secondary (P2), removing the lock-nut (7) holding the grounding lug to the chassis and the screw (11) mounting the transformer itself, and sliding the power inlet module upwards out of its slot in the rear of the tray.

10. Re-assemble the unit by reversing the above steps.



A1A4107 REV A

Figure 7.1 Model 422 tray assembly

SECTION 8
APPENDICES

8.1 Appendix A: B-type noise reduction

Dolby B-type noise reduction system/ Robert Berkovitz and Kenneth Gundry/ Audio Magazine, September 1973.

8.2 Appendix B: C-type noise reduction

A 20 dB audio noise reduction system for consumer applications/Ray Dolby/JAES, March 1983.

8.3 Appendix C: S-type noise reduction

A new analog recording process for use with consumer recording formats/Stan Cossette.

8.4 Appendix D: Remote control of Model 422

The 15-pin D-connector on the rear of the Model 422 (J501) permits remote control of most of the functions. The logic levels are nominally +15 V ("1") and ground ("0"), and the +15 V supply is available on pin 9, permitting the powering of 74C- or 4000-series CMOS where necessary; there is ample current available to operate led indicators if desired.

A "flash" waveform is available on pin 7 so that indicating lights can be made to flash like those on the unit itself; note that this swings nominally between +15 V and -15 V, and does not have enough current available to drive external leds, so buffering is necessary.

The functions of the front panel push-button switches, NR in/out, Bypass and Set-up, can be replicated remotely with indicators using mechanical latching switches as shown in figure 8.1. With this arrangement, the local and remote switches are in effect in parallel, and the "in" position will dominate.

If remote selection of NR type is needed, J501 pins 14 and 13 control channels 1 and 2 (and pins 6 and 5 channels 3 and 4), with the following truth table:

	pin14(6)	pin13(5)
S-type	0	0
B-type	1	0
C-type	1	1
invalid state	0	1

Here, 0 represents ground and 1 represents +15 V, both from a low impedance so as to override the internal drive. 74C- or 4000- series CMOS gates will serve for this purpose. These pins may also be used to drive indicator lights, via buffers.

For remote control of the mode of operation of channels 1 and 2, tie pin 10 to +15 V for encode and to ground for decode; this will override the internal link.

The plug-in input filters may be bypassed from outside the unit by connecting pin 4 of J501 to ground (pin 15).

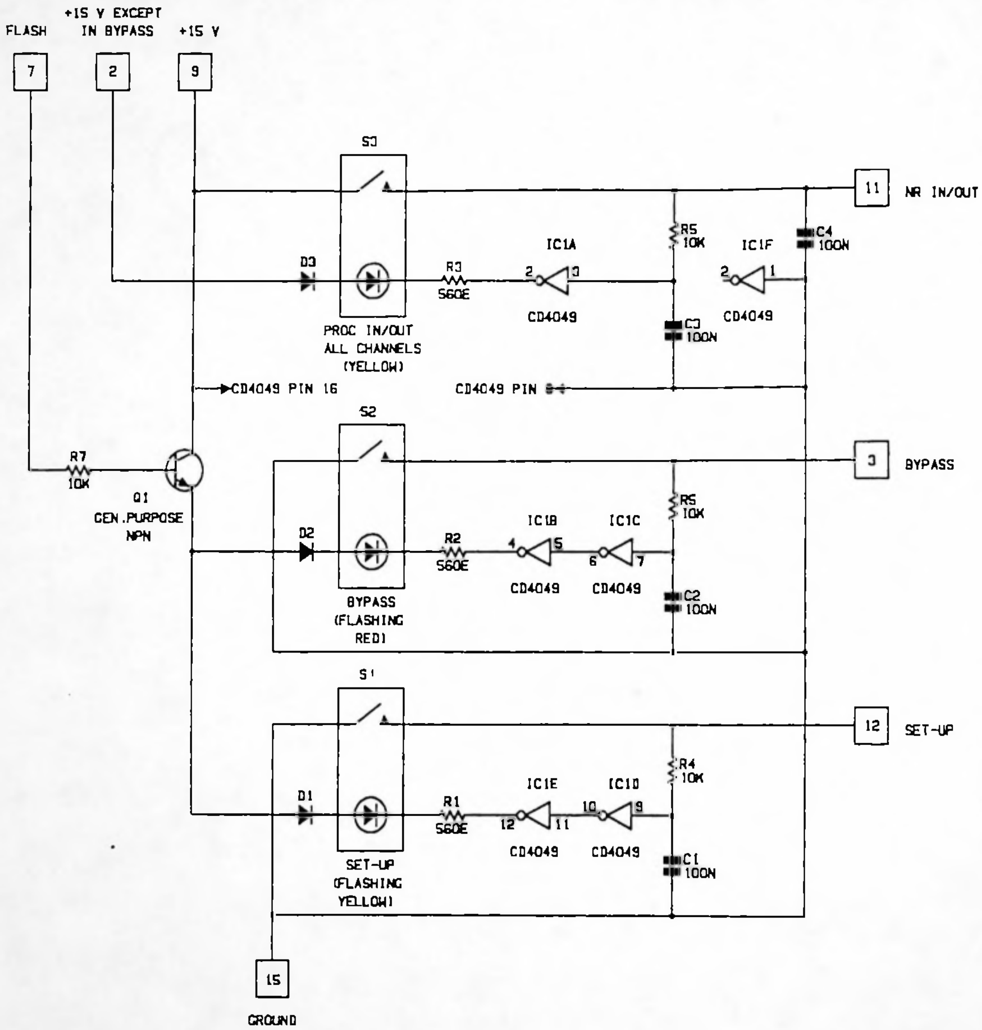


Figure 8.1 Simple remote control of Model 422
(connector pin numbers are those on remote control connector J501).

8.5 Appendix E: Patents

The consumer Dolby noise reduction systems are covered by the following:

U.S patents: 3,846,719 4,490,691 4,498,055 4,736,433
U.K.patents: 2,079,112 2,079,113 2,079,114 2,111,355
Other U.S., U.K. and worldwide patents granted and pending.

Dolby B-Type Noise Reduction System

Robert Berkovitz and Kenneth Gundry*

An unsatisfactory signal-to-noise ratio has remained the major obstacle to attaining an adequate level of performance from consumer media for music reproduction. This is especially true of the music-cassette, because of its slow tape speed and narrow track width, but it is also true of stereo FM broadcasting and the phonograph record. Although hopes were raised in recent years that further development of magnetic tape would eliminate its inherent noise as a problem, these hopes have been frustrated by the relatively modest gains achieved and by studies which indicate that the available signal-to-noise ratio of present-day tapes is very near the maximum value imposed by theory.

It is therefore not surprising that numerous attempts have been made to devise methods of noise reduction satisfactory for professional and consumer use. However, almost all of the methods proposed have had unacceptable drawbacks.

The effectiveness of single-ended (non-complementary) systems, for example, which are designed to be used only during playback, extends only as far as the listener's willingness to sacrifice musical information. In principle, all playback-only systems depend upon the idea that the signal and the objectionable noise occupy separate domains; if this is correct, then the problem of noise reduction is one of defining the boundary between the domains, in terms of frequency and/or level, and designing a circuit to suppress everything on the "noise" side of the boundary. However, if the noise spectrum of ferric oxide cassette tape is taken as an example (see Fig. 1), it is seen that the noise, when passed through a standard DIN weighting network simulating the ear's sensitivity, remains considerable in the 1-4 kHz range. Since this range includes many of the lower harmonics and upper fundamental tones in music, it is not possible to suppress it, even at low listening levels, without serious loss of information. On the other hand, the noise within this range is so disturbing that if it is not reduced by such a circuit, the amount of subjective improvement obtained is minimal.

Complementary methods, i.e., those which require some signal processing or encoding during both recording and playback, offer greater promise, but can also present difficulties when put into practice. Pre- and de-emphasis schemes, for example, in which high frequencies are increased during recording and decreased by the same amount during playback, are only of limited value. Even in FM broadcasting, where such standardized pre-emphasis has been employed for many years, the usefulness of its continued application is in doubt. The primary problem is that modern microphones and recording equipment now routinely reproduce high frequencies at amplitudes so high that they were considered unlikely when current FM standards were set. Broadcasters are now forced to use limiters to prevent overmodulation, if they also wish to maintain reasonable levels at middle and low frequencies. In magnetic tape recording, pre-emphasis is difficult to use because tape saturation occurs at lower levels at high frequencies. Since high-frequency signals already present problems in cassette recording because of their short wavelength, added pre-emphasis would complicate a task which is already difficult.

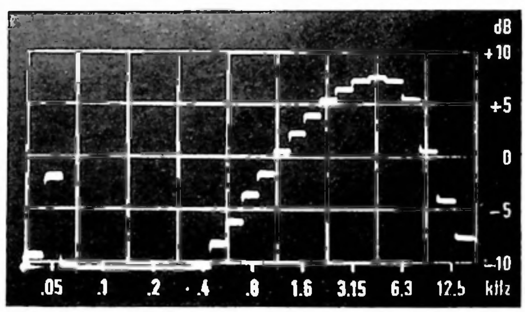


Fig. 1—DIN-weighted noise spectrum of low-noise ferric oxide cassette tape displayed on screen of real-time analyzer. Vertical scale is relative only. DIN noise measurement of sample shown was −47 dB with respect to 200 nWb/m. A substantial amount of the noise present falls well within the range of musical fundamental tones.

*Dolby Laboratories, Inc.

Reprinted by courtesy of North American Publishing Co.

AUDIO · SEPTEMBER 1973

The compandor type of noise reduction system, in which the dynamic range of the signal is compressed during recording and expanded again during playback, offers more promise. However, a simple compandor, even if precise in its action, also presents problems. In recordings and broadcasting, one of the most serious drawbacks of the compandor is the danger of signal overshoot, which can result in distortion or overmodulation of a transmitter.

An even more serious problem of compandors, from the listener's point of view, is noise modulation. When a conventional full-band compandor is used, low-level passages are recorded at a level higher than normal. They are then played back at reduced level, restoring correct signal dynamics and reducing noise at the same time (see Fig. 2). There can be no noise reduction effect during high-level passages, because this would require increasing the level of such passages during recording, resulting in overload. The simple compandor therefore requires that one assume that noise is not objectionable when the signal level is high. However, this is not always the case. A high-level bass drum beat, for example, does not mask high-frequency tape hiss; as a result the drum and other instruments introduce noise modulation during playback—each note is accompanied by a "swish" as the noise level rises for the duration of the note. While it is not audible with all types of program material, noise modulation limits the usefulness of the compandor considerably.

The extreme diversity of available source material and the high quality of present-day master recordings are the factors which really determine the conditions to be met by a satisfactory noise reduction system for home use. It must be remembered that many home listeners own playback equipment with very low distortion and wide frequency range, disclosing audible effects which might have passed unnoticed in earlier times. Therefore, it is especially important that the program be recovered accurately after noise reduction, without addition of any audible sound. For the listener's sake accuracy of recovery and effectiveness of the system should not require adjustment of system parameters to match various kinds of program material. At the same time, the size and cost of the system should introduce no obstacle to its use. Furthermore, as a practical matter for the industry, it is clear that the system should require no modification of present professional practice in master recording, duplicating, or broadcasting.

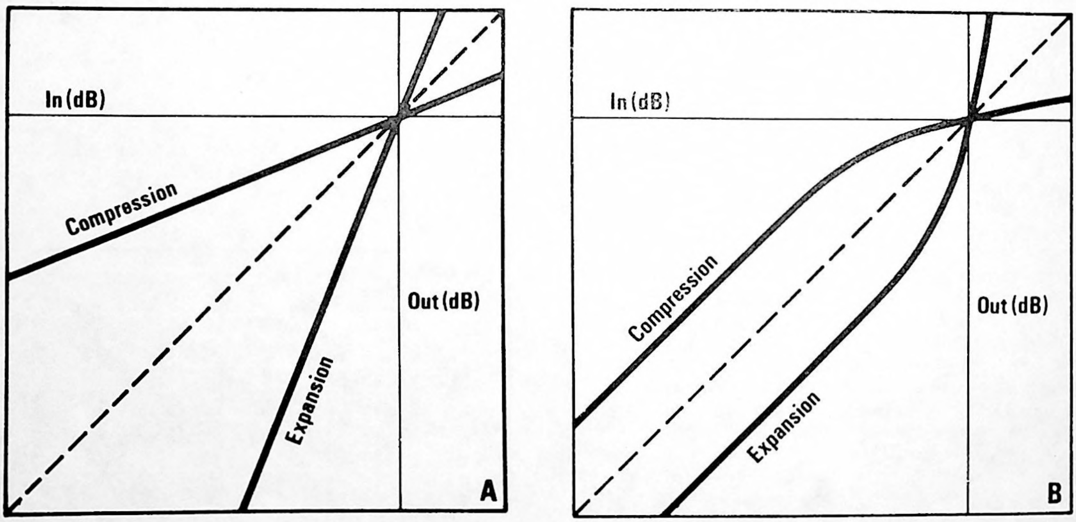


Fig. 2—Transfer characteristics of two conventional compandors (solid lines). A, constant-slope type; B, high-level type. Since compression and expansion are functions of signal amplitude only, in a single frequency band, such compandors fail to suppress noise whenever natural masking fails (see text).

Dolby B-Type Noise Reduction System - Part 2

Robert Berkovitz and Kenneth Gundry*

The Dolby B-Type Noise Reduction System

The Dolby B-Type circuit is a specialized form of compandor which avoids the usual deficiencies of compandors. The operational principle of the B-Type system is complementary low-level compression and expansion in a frequency range which varies in bandwidth as the signal changes.

Most objectionable noise encountered in home listening is at middle and high frequencies, from about 500 Hz to the upper limit of audibility. In the interest of circuit economy, the action of the B-Type circuit has therefore been limited to this range. A feedback control circuit adjusts system parameters automatically as a function of signal level and spectrum, so that the system's action complements the psychoacoustic masking of noise which occurs naturally in the course of the program. A block diagram of a Dolby type of noise reduction system is shown in Fig. 3. The circuits used for encoding (during recording or transmission) and decoding (during playback or reception) are quite similar and can be considered as the same circuit, switched to operate in either mode.

The compression and expansion characteristics of the Dolby B-System are fixed and are referred to Dolby Level, a specific internationally standardized reference level. In the case of cassette tape, Dolby Level is a flux of 200 nWb/m; in FM broadcasting, Dolby Level is ± 37.5 kHz deviation.

Figure 4 is a block diagram of a switchable (encode-decode) B-Type circuit. There are two paths which the input signal follows: a *main path* (at the lower part of the figure) in which no change other than linear amplification occurs, and a *secondary path*, a variable filter through which only low-level, high frequency components of the input signal are allowed to pass. To encode the signal, the output of the secondary path is combined with the signal in the main path *additively;* this boosts low-level, high frequency portions of the signal. Decoding is accomplished by feeding the secondary path from the circuit output, which is opposite in phase to the input (note phase inverter in Fig. 4); the secondary path is then part of an a.c. negative feedback loop which reduces output, i.e., the output of the secondary path is combined with the main path *subtractively*. In the decode mode, therefore, the circuit reduces the level of precisely the same information which was increased in level during encoding.

*Dolby Laboratories, Inc.

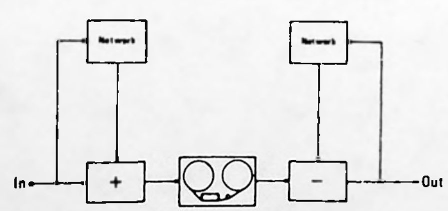


Fig. 3—Block diagram of Dolby type noise reduction circuit as used in typical record-reproduce chain.

As Fig. 4 indicates, the action of the B-Type circuit is controlled by the output of the filter in the secondary path. Above a fixed threshold level, the bandpass of the filter, in turn, is modified by the d.c. feedback loop.

At very low levels, i.e., below the threshold, which at high frequencies is about 40 dB below Dolby Level, the output of the filter is not sufficient to generate d.c. feedback; consequently, the output of the secondary path is simply proportional to signal level within the filter pass band. The output of the circuit is then essentially as shown in Fig. 5.

As signal level rises above the threshold level, the rectified filter output is returned to the FET gate where it is applied as negative feedback, raising the filter cutoff frequency so that the output of the secondary path, while still increasing, no longer does so in proportion to the change in signal level. As signal level becomes even larger, the increasing d.c.

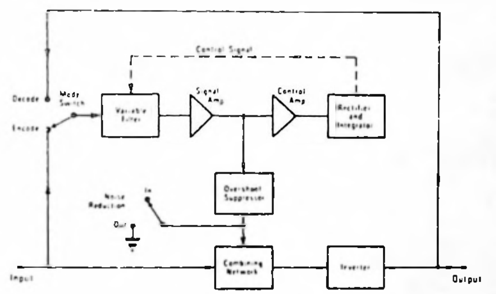


Fig. 4—Block diagram of Dolby B-Type noise reduction circuit. The configuration shown can be switched to encode or decode the signal.

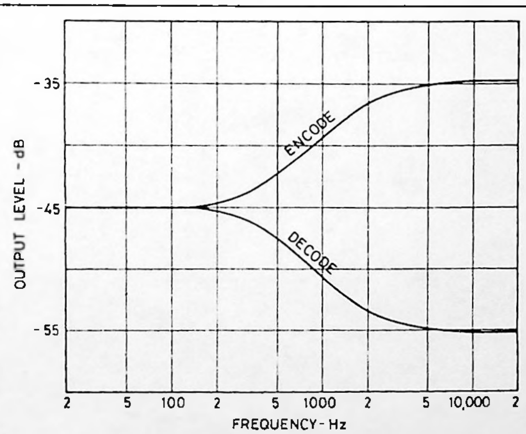


Fig. 5—Output of B-Type encoder and decoder circuits under low-level input signal conditions. The two operations are symmetrical and the result is an overall frequency response which is level.

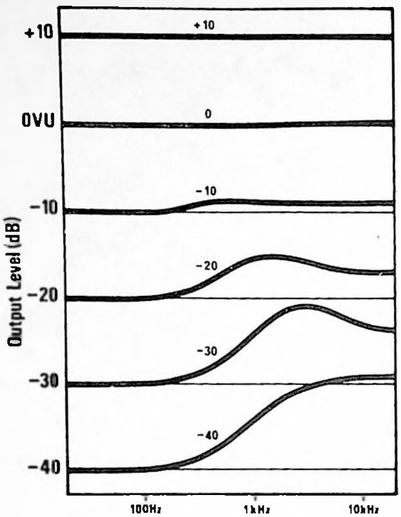


Fig. 6—Characteristics of encoding processor at several levels. The gradual reduction in boost with increasing level avoids possible tape overload.

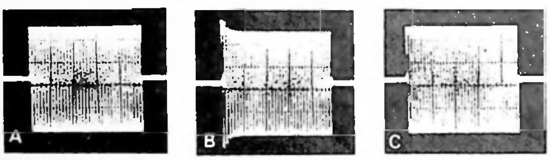


Fig. 7—Effect of the B-Type circuit on a tone burst; frequency = 3 kHz; burst duration, 12 milliseconds; low level = 40 dB; high level = +6 dB; (A) Input to system; (B) Encoded; (C) Encoded and Decoded.

feedback generated restricts the filter bandwidth further, and near Dolby Level the output of the secondary path remains relatively constant. The net effect is that the secondary path has no audible effect on output at low frequencies, and increasing effect with increasing frequency and decreasing level to about 40 dB below Dolby Level. At high levels, the effect of the extra signal is so small as to have no significance; at low levels, in the spectral region in which noise reduction is required, the increase during encoding is as much as 10 dB, and is of considerable importance.

The manner in which the secondary path changes from constant-gain to constant-output is determined by the adjustment of gain within the feedback loop. In addition, the exact variation in filter bandpass with changing level is set optimally by making the control amplifier frequency-dependent. The overall frequency response of a B-Type encoder circuit for different input levels is shown in Fig. 6.

A compandor operating over a wide frequency range must be designed to take into account the problem of noise modulation discussed above. If some high-level passages in the program differ sufficiently in frequency content from the noise components present, the latter will remain audible during the program in many cases. However, these passages cannot be increased in level when encoded, because of the danger of overmodulation. Under these conditions, compression may be applied intermittently, and high-frequency noise modulated audibly by mid-frequency components of the signal. The B-Type circuit overcomes this problem because it continues to function when a high-level signal occurs within its operating range; instead the feedback control shifts the range upward in frequency. This avoids the danger of overmodulation, but retains full noise reduction at frequencies higher than those masked by the signal.

The attack time of the B-Type circuit is dependent on the amount and rapidity of the signal change, due to the nonlinear design of the integrator, varying from about 100 milliseconds to as little as 1 millisecond. The recovery time of the rectifier-integrator is shorter than that of the human hearing system, about 100 milliseconds.

All compressors exhibit overshoot, including the B-Type

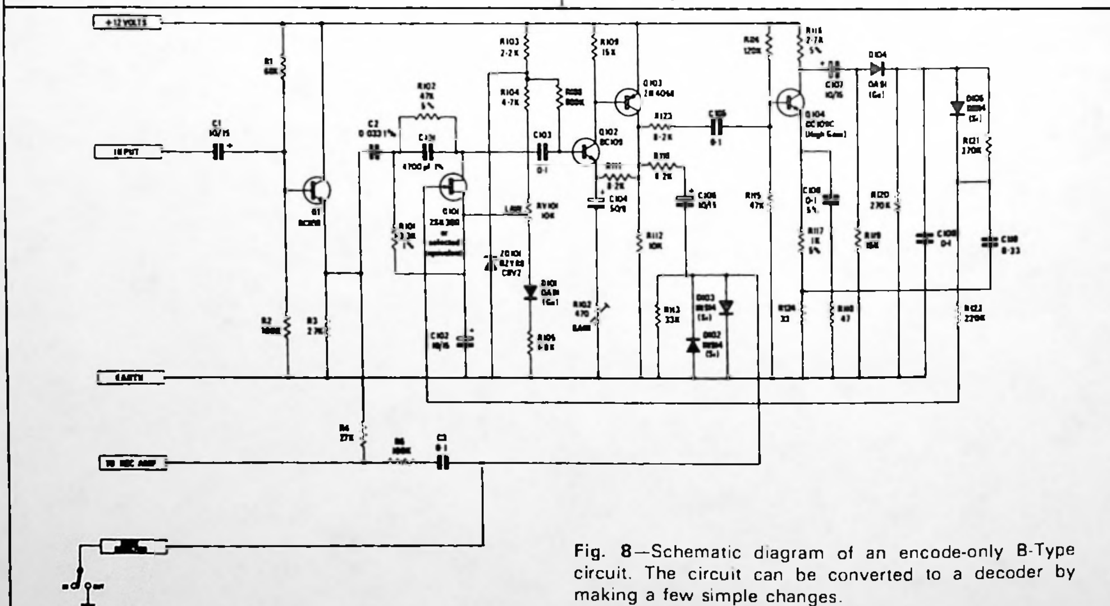


Fig. 8—Schematic diagram of an encode-only B-Type circuit. The circuit can be converted to a decoder by making a few simple changes.

circuit. However, the dual-path approach used makes it possible to reduce the amplitude of overshoots significantly. Overshoot, which can occur only in the secondary path (where it can be suppressed without affecting the main signal) is comparatively small, and essentially disappears when the signal is decoded again. When signal levels are low, or when changes in signal level take place slowly, there is no overshoot problem: when signal changes are large and rapid, diodes in the overshoot suppressor stage limit the peaks of the overshoot. Since this takes place in the secondary path, the result of the suppressor action is to limit overshoot to a relatively small fraction of the full-level main path signal. Further, by restricting overshoot suppression to the secondary path, it is possible to avoid introducing audible distortion to the encoded signal. Because a complementary action takes place during decoding, the small remaining overshoot in the encoded signal is eliminated, and as with other effects produced during encoding, the original signal is restored. Figure 7 shows the result of encoding and decoding a short burst of 3 kHz, which changes in level from −40 dB to +6 dB.

Figure 8 is a typical schematic diagram of an encode-only B-Type circuit; the circuit for decoding-only is similar. As can be seen, only five transistors plus an FET are required; the parts cost of the circuit is approximately $2.40.

Figure 9 is the schematic diagram of a B-Type processor which has been designed to integrate noise reduction with other tape recorder electronics requirements as much as possible. The resulting circuit provides 26 dB of gain, whether or not noise reduction is in use, bias and multiplex filtering, and meter and monitor amplifiers. In fact, the only additional electronics needed to complete the recorder are a bias oscillator, recording amplifier (one transistor) and a microphone and head amplifier (two transistors). With the active elements used in the record/play switchable processor shown (eight transistors and one FET), the total used in the recorder, for two channels, is 22 transistors and two FET's. The cost to a manufacturer of the components shown in Fig. 9 is about $3.20, excluding the bias and multiplex filter components, which are, of course, necessary in the circuits of any properly designed tuner and recorder.

Dolby Laboratories and Signetics have collaborated in the development of an integrated-circuit version of the B-Type circuit. The IC is expected to offer manufacturers economy of assembly, elimination of adjustments, and somewhat smaller space requirement than the discrete-component version.

The characteristics of Dolby B-Type noise reduction can be summarized as follows:

1. Program recovery characteristics, with regard to frequency response, phase response, transients, and signal dynamics, are theoretically perfect; in practice, this ideal is attainable to any desired accuracy. Distortion in practical B-Type circuitry is considerably lower than that of the tape recorders or tuners with which it is used. Any type of program material can be encoded and decoded without audible loss.

2. The circuit is simple, inexpensive, and small in size, either in discrete-component or IC form.

3. The circuit is easy to manufacture and use because of the absence of critical components or adjustments. The circuit can be quickly and easily calibrated during manufacture, after which further calibration is not required. In use, only a simple level adjustment is necessary if tape of significantly different sensitivity is substituted for that formerly used.

4. No modification of broadcasting or duplicating practice is required to incorporate B-Type encoding. The use of the noise reduction system often makes worthwhile other im-

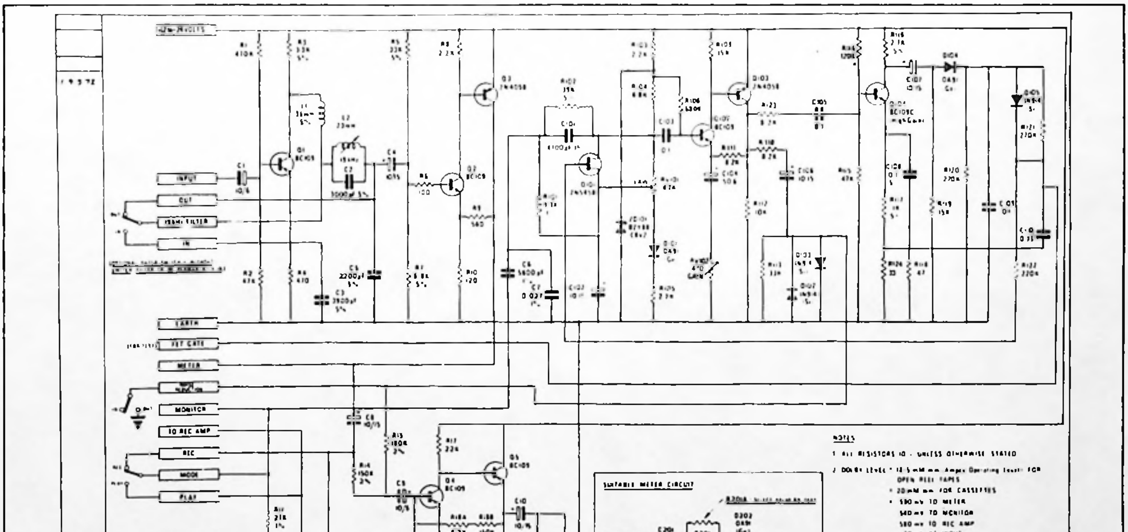


Fig. 9—Schematic diagram of complete switchable encode/decode circuit for use in tape recorder, including HF filtering, meter drive, monitor output and 26 dB of gain. Only a few more parts need be added to make complete record/play electronics for one channel of the recorder.

provements, however, such as extension of frequency response and dynamic range, or reduction of distortion by use of lower modulation levels, or some combination of these.

Effects Upon Noise Spectra

Figure 10 is a multiple exposure of the screen of a 1/3-octave real-time analyzer, allowing a direct comparison of the noise spectra at the output of a high-quality cassette recorder when different kinds of tape were used with and without the Dolby B-Type noise reduction circuit. Curve 1 is that produced by C90 ferric oxide tape; curve 2 is that of C90 chromium dioxide tape; curve 3 is produced by the same tape used for curve 1, but the B-Type circuit is switched "in," and curve 4 represents the noise spectrum of the chromium dioxide tape with the circuit in. The tapes shown were biased before the measurements were made; no changes in gain or other control settings were made during the tests, other than to set equalization differently for the chromium dioxide tape from (70 microsecond). In fact, most of the improvement in noise level obtained when chromium dioxide tape is used appears to be due to the change in equalization; if this change is not made, there is little advantage in chromium dioxide tape from a noise point of view. On the other hand, the combination of chromium dioxide tape, 70 microsecond equalization, and B-Type noise reduction results in an excellent noise figure, 57 dB below Dolby Level in the example in the photograph (DIN 45405).

The advantages of B-Type noise reduction are also obtained when the system is used for FM broadcast transmission and reception, i.e., the improvement in signal-to-noise ratio obtained by use of the B-Type circuit is approximately the same as that produced by a 10 dB increase in field strength. The significance of this improvement can be appreciated when it is realized that such an increase would usually require an increase in transmitter power by a factor of ten. Considerable experimentation and broadcast experience in the USA have demonstrated, as one would expect, that the area in which listening is satisfactory is greatly extended by use of the B-Type noise reduction system. Several American classical music FM stations are already broadcasting full-time using Dolby B-encoding.

Compatibility

When any improvement is made in a system as widely used as the compact cassette system, it is highly important that the new development should be fully compatible with existing equipment. Improved cassettes must be playable on any machine which can play old-type cassettes, and fortunately this is true of Dolby B-Type cassettes. Such cassettes are subjectively compatible (i.e., generally pleasing to the listener) when played without decoding circuitry, to a great extent because of the unique approach taken in the B-Type circuit. Because most low-cost cassette machines are deficient in high-frequency response, the increase in low-level high frequency content in a B-Type cassette is usually welcomed by listeners with such equipment. Cassette recorders of higher quality, or the associated equipment with which they are used, contain tone controls which permit the balance to be adjusted to suit the taste of the listener. It is quite likely that many of the millions of B-Type encoded cassettes which have been made commercially are owned and played by listeners who are unaware of the special nature of the program material they hear. In any case, the subjective difference between encoded and other cassettes is sufficiently unobtrusive that none of the recording companies offering "Dolbyized" cassettes have found it necessary to offer old-type cassettes as an alternative.

It is worth noting that almost all pre-recorded cassettes are already compressed, for only in this way can the audibility of low-level passages be preserved in programs of wide dynamic range. B-Type cassettes differ mainly in that the listener now is able to remove the compression by pushing a button on his cassette machine restoring program dynamics and reducing noise. This is only possible because B-Type compression is standardized, while other types of compression vary considerably.

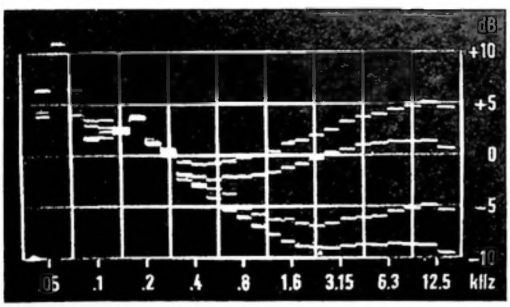


Fig. 10—Noise spectra at output of cassette recorder during playback of four kinds of tape. All tape was biased and control settings were the same for the four tests. (1) Ferric oxide C90 low-noise, —45.7 dB DIN 45405; (2) Chromium dioxide C90, —48 dB; (3) Same as (1) with B-Type noise reduction, —55 dB; (4) Same as (2) with B-Type noise reduction, —57.1 dB. All noise measurements referred to Dolby Level = 200 nWb/m.

Commercial Use

Within a few years of its introduction, the Dolby B-Type noise reduction system has been licensed to most major manufacturers of consumer tape recorders. At the present time there are more than 40 licensees manufacturing over 100 different B-Type products. Licensee payments for use of the circuit are on a sliding scale, based on quantity, from a maximum of 50¢ (U.S.) to 10¢ per circuit. Royalty charges are typically 60¢ per stereo unit for a major manufacturer.

In addition, most of the pre-recorded cassettes now made in the United States, the United Kingdom and Japan are "Dolbyized," and many of the largest recording companies issue their cassette output in this form, among them Ampex and CBS in the United States, Decca and RCA in England, and CBS-Sony, Nippon Columbia, King, and Apollon in Japan. Pre-recorded open-reel tapes and 8-track cartridges are also becoming available. In the United States, a number of FM stations have already started to broadcast regularly in B-Type encoded form, and this procedure is under study in other countries as well. There is no royalty payable for encoding cassettes or other tape recordings, or broadcasts.

Conclusions

The reduction of background noise by the Dolby B-Type noise reduction system has contributed importantly to the improvement in quality of home tape recording and playback. It has helped to make the extension of frequency response, the reduction of wow and flutter, and other improvements worthwhile, particularly in cassettes. The unique characteristics of the B-Type system permit excellent noise reduction without program losses, noise modulation and other drawbacks which have afflicted earlier attempts to solve the noise problem. The simplicity and economy of the B-Type circuit facilitate its use in consumer products at all price levels. Æ

A 20 dB Audio Noise Reduction System for Consumer Applications*

RAY DOLBY

Dolby Laboratories Inc., San Francisco and London

A 20 dB noise reduction system, designated C-type, for use in cassette tape recording and similar applications is described. An arrangement of two compressors and two expanders in cascade has been developed in which the signal-to-noise ratio improvement is compounded without significant accompanying increases of the overall maximum compression and expansion ratio. Overshoots, modulation distortion, and noise modulation are well controlled. The maximum demands made on transmission channel uniformity are generally unchanged from those associated with the B-type system, although the uniformity requirements extend over a greater range of signal frequencies and levels. An improvement has been made in one condition of compressor/expander mistracking, namely, low-level mid-frequency signals in combination with dominant signals in the region above 10 kHz and incorrect channel response at such frequencies. A further development reduces the tendency of highly equalized channels to saturate, thereby increasing the useful signal levels which can be handled.

0 INTRODUCTION

The B-type noise reduction system [1], [2] was developed in 1967–1968 and first applied to open-reel recording (KLH models 40 and 41). However, by this time there was a general feeling that a more convenient tape format was required for widespread use. In late 1968 we therefore began experimenting with 8-track cartridges and the B-type system. The results were encouraging if suitable tape formulations and oxide thicknesses were used, but ergonomic and aesthetic considerations persuaded us that the 8-track cartridge would not be a success as a quality recording medium; anyone willing to tolerate an endless-loop format would be unlikely to be very interested in sound quality.

In early 1969 we turned to the Philips Compact Cassette, which was another of several tape formats competing for the popular market at that time. The Compact Cassette offered the advantage of rapid access, which appeared to be a requirement for acceptance by critical listeners. The disadvantage of very low tape speed (1⅞ in/s, 4.75 cm/s) was to some extent offset by the special tape formulations and oxide thicknesses that had to be created, since there was no practical possibility for the industry to use standard thick-oxide ¼-in (6.3 mm) tape in cassettes (as there had been in 8-track cartridges).

Throughout 1969 we researched the properties of available cassette recorders and cassette tapes, and by the end of that year we had adapted and improved several cassette decks to provide wide frequency response, low distortion performance with high stability. Using the B-type noise reduction system, we demonstrated our results to cassette deck manufacturers and cassette duplication firms. There was general agreement that this was a promising development; under good conditions it was possible to produce overall results that were comparable with the best disks. The technology has since been adopted and widely used, and by this time it is possible to draw up a comparative list of the main technical defects of disks and B-type encoded cassettes when produced and reproduced under the best conditions:

Disks	Cassettes
Mold-grain noise	Hiss
Hiss	High-frequency overload distortion
Ticks and pops	

* Presented at the 70th Convention of the Audio Engineering Society, New York, 1981 October 31–November 2.

The low-frequency mold-grain noises (rumbling and rushing sounds) produced by disks are evidently unnoticed by most listeners; perhaps these noises are masked by the ambient noises of typical listening environments. Disk processing hiss is variable but usually not too obtrusive. The main audible defect of most disks is low-level ticks and pops. In contrast, cassette tapes have no audible rumble or other low-frequency noises, and, of course, there are no ticks and pops. However, the hiss level is audibly greater than that of disks; the continuing presence of this hiss has evidently been the main factor in causing the cassette to fall just short of disks in the estimation of quality conscious listeners. A further element is that cassette tapes, especially the ordinary formulations used in mass duplication, do not have the high-level high-frequency recording capability of disks. For economic reasons, most duplicators are reluctant to use tapes that might overcome this problem.

In 1978 we developed a system, HX (headroom extension), to improve the high-level high-frequency performance of normal cassette tapes [3]. This system was introduced in consumer cassette recorders in 1979–1980. While this development was welcomed by the technical community, there was still a feeling that the basic noise performance of the cassette, using the B-type noise reduction system, was inadequate. Several different noise reduction systems offering more than 10 dB of noise reduction became available, and many cassette deck manufacturers requested a response from us to this activity.

Until early 1980 the author remained unconvinced that the underlying demand for an improved (and more costly) system would be sufficient to justify the industry infrastructure required to support a new high-performance standard. However, performance expectations do not appear to diminish. Thus a new noise reduction system called C-type has been developed, which, it is hoped, will meet a reasonable proportion of these expectations. The author, as well as many others, will be waiting with interest to see whether the long-term demand is broadly based enough to result in a significant change in usage patterns.

This paper describes the new system, which utilizes two series-connected sliding-band compressor and expander stages, operating at different levels, to solve the problem of increasing the overall compression, expansion, and noise reduction without introducing side effects. Further developments reduce high-frequency tape saturation and improve the tolerance of the system to irregular response of the recorder at very high frequencies. Good frequency response and level reliability are nonetheless required at lower frequencies.

1 STAGGERED ACTION DUAL-LEVEL FORMAT

In the development of the A-type noise reduction system in 1965–1966 [4] the author found that a two-path configuration and a maximum dynamic action of the order of 10 dB, placed some 30 dB below the nominal maximum level, provided a good margin of safety in solving the problem of suppressing compressor overshoots without introducing audible distortions caused by rapid modulation of the signal. In the development of the B-type system in 1967–1969 these facts, coupled with tests to determine the maximum dynamic action likely to be allowable for reasonable compatibility when encoded recordings were reproduced without decoding, established the maximum noise reduction at 10 dB.

In the development of the C-type system in 1980, the compressor overshoot and modulation distortion consideration pointed strongly toward the retention of the dual-path 10 dB low-level format which had proved to be successful in the A-type and B-type systems. While it was tempting to contemplate stretching the capability of the basic 10 dB circuit to performance levels in the 15–20 dB region, only a few experiments were enough to reconfirm that such an approach would be hazardous at best; it would be better to accept the cost penalty of a more complex method and to be safe. A two-band configuration would not be much help, since each band would still be required to operate with the full dynamic effect. However, if two stages could be cascaded, then the stage gains and resultant compression and expansion would be multiplied (or added on a dB basis) to yield an overall noise reduction of, say, 20 dB. While early tests indicated that this was an attractive method under ideal conditions, the resulting high compression ratios (up to 4:1) would clearly be a problem with the production and operating tolerances of practical cassette recorders. A method was therefore devised whereby the dynamic actions of the two stages could be spread out or staggered into different level regions. Such dynamic action staggering, in which one stage operates at levels comparable with those of the B-type circuit, and the second stage treats signals some 20 dB lower in level, is possible with compressor and expander stages having a certain type of transfer characteristic which will be discussed. This staggering technique proved to be a key element in the development of the C-type system.

Referring to Fig. 1, the A-type and B-type noise reduction systems employ a level transfer characteristic which at any particular frequency comprises the following elements:

1) A low-level linear portion up to a threshold (where "linear" in this context denotes constant gain with changing input level).

2) An intermediate-level non-linear portion (changing gain with changing input level) above the threshold and up to a finishing point, providing a certain maximum compression or expansion ratio.

3) A high-level linear portion having a gain different from the gain of the low-level portion.

This type of characteristic can be designated a bilinear characteristic because there are two portions of substantially constant gain. Such characteristics may be distinguished from other types of characteristics, namely:

1) A logarithmic or nonlinear characteristic with

either a fixed or a changing slope and with no linear portion: the gain changes over the whole dynamic range.

2) A characteristic having two or more portions of which only one portion is linear.

An advantage of a bilinear characteristic is that the threshold can be set above the input noise level or transmission-channel noise level in order to exclude the possibility of control of the circuit by noise; the low-level region is a reliable "gain floor," which contributes to overall stability of the signal. The high-level portion of substantially constant gain avoids the nonlinear treatment of high-level signals which would otherwise introduce distortion, either by rapid modulation of the signal or by overshoots and subsequent clipping. In the region of dynamic action, at intermediate levels, relatively long attack and recovery times are used in order to reduce modulation distortion. The attack and recovery times are progressively reduced with increasing amplitude steps, the high-level portion providing a region within which to deal with the overshoots, which in a dual-path system are suppressed by clipping diodes acting upon the noise reduction signal only.

Thus with 10 dB of dynamic action spread over an input signal level range of about 20–25 dB, so that the maximum compression ratio does not substantially exceed 2:1, it is possible to set the threshold at a level high enough to be well clear of input signal noise and recorder noise, that is, in the region of 40 dB below the nominal peak level. This leaves a high-level linear region of some 20 dB for the suppression of overshoots.

Note that bilinear compressors and expanders determine the two end regions of constant gain by means of fixed, preset circuit elements, such as resistors and capacitors, which are inherently stable and cannot cause dynamic errors, waveform distortions, and the like. Only in the transitional area can any dynamically active portions of the circuits introduce signal errors.

In contemplating the possibility of a multistage circuit, it should be noted that prior attempts have resulted in a multiplication of the maximum compression ratios of the individual stages with the consequence of an overall high compression ratio, which is not very useful in a practical noise reduction system (for example, one circuit with a compression ratio of 2:1 and the other with a compression ratio of 3:1 will yield an overall ratio of 6:1). Other cascaded approaches have utilized compressor stages operating in mutually exclusive frequency ranges. While such an arrangement may not necessarily result in any increase in the maximum compression ratio over that of a single stage, it cannot provide an overall increase of noise reduction at a particular frequency.

Experience has shown that with a compression ratio of much more than 2:1 it becomes increasingly difficult to ensure complementarity between the compressor and the expander; in particular, level errors or errors in the frequency response of the recorder lead to correspondingly multiplied errors at the output of the expander.

An examination of bilinear circuits used in a series connection shows that they not only have the previously discussed advantages but further ones as well, namely, a way of solving the high compression ratio problem and a way of dealing with the larger overshoots which accompany greater overall compression.

Note that the superposition of the high- and low-level linear regions does not increase the compression ratio in these regions (since by definition the compression ratio is 1). The compression ratio is increased only in the limited region in which dynamic action takes place. Therefore it becomes possible to separate the areas of dynamic action of the two stages in such a way as to obtain the required overall increase in compression without altering the overall maximum compression or expansion ratio significantly. A further feature is that the overall result is bilinear, with all of the attendant advantages. Thus the action staggering possibility of bilinear compressors and expanders represents a further advantage of this type of device.

At any given frequency, the thresholds and dynamic regions of the compressor or expander stages are set to different values so as to stagger the intermediate-level portions of the characteristics of the stages. This results in a change of gain over a wider range of intermediate input levels than for each of the stages individually, an increased difference between the gains at low and high input levels, and a maximum compression or expansion ratio which is substantially no greater than the maximum compression ratio of any single stage.

The thresholds of the overshoot suppressors are also staggered along with the stagger of the syllabic thresholds. The overshoots of the low-level stage are correspondingly reduced.

Fig. 2 shows the basic block diagram of the staggered action method. A high-level bilinear compressor feeds

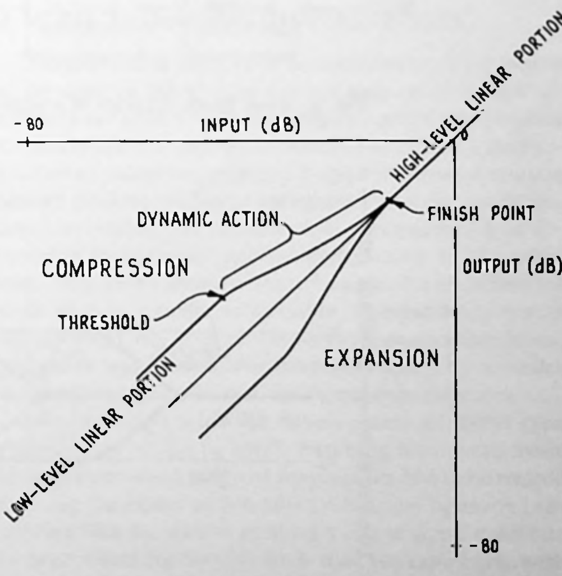


Fig. 1. Bilinear compression and expansion characteristics.

the low-level bilinear compressor connected in series. During playback a pair of series-connected bilinear expanders receives the input from the signal channel and provides an overall noise reduction system output at the output of the high-level expander.

For overall complementarity of the system, the order of the stages in the compressor is reversed in the expander. Thus the last stage of the expander is complementary to the first stage of the compressor (and likewise the first stage of the expander to the last stage of the compressor) in all respects, both steady state and time dependent.

The separation or staggering of the high- and low-level stages is depicted in Fig. 3, which plots compression ratio versus input amplitude (horizontal axis) for the compressor or expander stages operating at a particular frequency. The top curves are those of compressors, the bottom curves those of expanders. In this example the areas of action as a function of input level are separated such that the product of the two curves results in an overall characteristic having a compression ratio or expansion ratio which does not exceed 2:1 (1:2) between the two maximum compression points 1a and 2a (1b and 2b) of the two devices. For clarity, the curves are shown in idealized form; as a practical matter the curves may be somewhat asymmetrical. The compressor portion of curve 2 represents the variations of the compression ratio of the high-level stage as a function of the input level to the high-level stage, while the compressor portion of curve 1 is the variation of the compression ratio of the low-level stage as a function of the input level to the high-level stage, as if the high-level stage had a constant gain. In practice, the high-level stage modifies the input signal to the low-level stage as a function of signal level. The overall characteristic produced is the left-hand portion of curve 1, the section from 1a to 2a, and the right-hand portion of curve 2. Analogous considerations apply in the case of the expanders depicted on the lower half of the figure.

Thus even with two compressors or expanders in series, the end regions of operation still remain fixed, the maximum compression and the maximum expansion ratios are not increased beyond those of single devices, and the advantages of single bilinear devices are retained. Consequently, the maximum error in level occurring within the range of dynamic action caused by the devices in series should not substantially exceed the maximum error of a single device. With the continually changing levels of real signals, however, the time-probability of a level error is increased because of the greater range of dynamic action of the cascaded devices over that of a single device.

Note that in the representation of Fig. 3 the dynamic action of a logarithmic compressor or expander becomes a horizontal line; line 3, for example, is the characteristic of a 2:1 compressor, and line 4 is that of a 1:2 expander. It is clear that there is no opportunity for separating or staggering the actions of such devices.

To obtain a first-order approximation of the parameter relationships in action staggering, it is useful to idealize Fig. 3 even further. Assume that each compressor (and expander) immediately reaches its maximum compression ratio at a threshold level and holds that ratio until it reaches a finishing point at a higher level where its dynamic action abruptly stops.

Based on observations of the resulting transfer char-

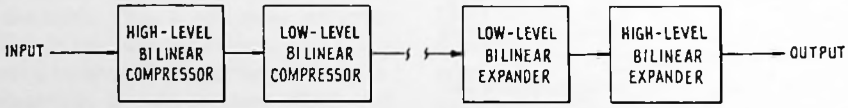


Fig. 2. Basic block diagram of staggered-action bilinear noise reduction system.

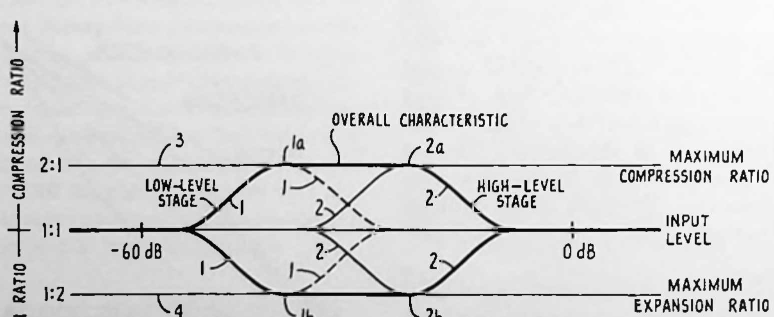


Fig. 3. Action-staggering principle.

acteristics (Fig. 4), the following equation sets forth the relationship between threshold level T, finishing point F, compression ratio C, and gain G of the stages:

$$T = F - \frac{CG}{C - 1} . \quad (1)$$

Using this equation is straightforward for the first stage. For the second stage, the first-stage threshold becomes the second-stage finishing point. However, the calculated threshold is the overall threshold, referred to the first-stage input. To obtain the threshold of the second stage referred to its own input, the low-level signal gain of the first stage is taken into account. The equation can also be arranged to give the finishing point F, the compression ratio C, or the gain G.

Consideration of the above equation and Fig. 4 shows that for the case of a 2:1 compression ratio, half of the threshold staggering is provided by the signal gain of the first stage and the other half must be provided by the control circuitry of the second stage.

As previously mentioned, a 2:1 compression ratio appears to be about the maximum that can be used in cassette recording systems, because of error amplification effects during decoding. A lower compression ratio (such as 1.5:1) would permit an expander to track the compressor more easily, but on the other hand, the dynamic action would have to extend down to lower levels, resulting in greater susceptibility to noise modulation for a given maximum amount of noise reduction. Hence there is a trade-off between undesirable effects caused by both high and low compression ratios.

Similar considerations apply in arranging the staggering of a dual-stage system. Once the maximum allowable compression ratio has been decided, then it is best to employ the minimum amount of staggering consistent with keeping the overall ratio within the design goal. Squeezing the area of dynamic action of the low-level stage up close to that of the high-level stage results in improved noise modulation performance of the low-level stage; there is little virtue in keeping the two areas well separated.

The two stages of the C-type system are each of the sliding band type, similar to that of the B-type circuit [1], [2]. The first stage of the compressor is set for operation at levels similar to that of the B-type circuit, and the second stage is set for operation at lower levels. In this order there is a useful interaction between the stage gains and the areas of dynamic action; the area of action of the downstream stage is partly determined by the signal gain of the preceding stage. Thus with 10 dB of low-level gain per stage, the control amplifier gain requirement of the second stage is reduced by 10 dB. When a high-level signal appears, the 10 dB gain of the first stage is eliminated from the overall effective amplification used to derive the control signal of the second stage. This improves the noise modulation performance of the second-stage sliding band.

If the arrangement were reversed, with the low-level stage first, there would be reduced interaction. The control amplifier of the first stage would need a high gain in order to achieve the required low threshold. This high gain and low threshold would then apply even in the presence of high-level signals, which in the case of a sliding band system would result in poorer noise modulation performance. Thus the arrangement actually used takes best advantage of the prevailing signal gains of the individual stages, namely:

1) Under very low level (sub-threshold) signal conditions the control amplifier gain requirement of the second stage is reduced by 10 dB over what would otherwise be required to achieve the desired staggering.

2) A signal-dependent variable threshold effect is achieved, which with sliding band stages reduces noise modulation effects.

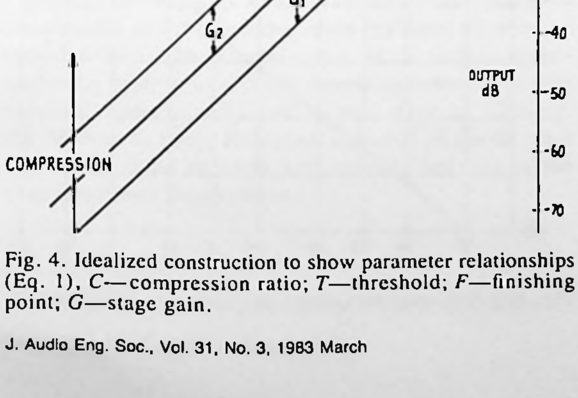


Fig. 4. Idealized construction to show parameter relationships (Eq. 1), C—compression ratio; T—threshold; F—finishing point; G—stage gain.

2 NOISE-REDUCTION CHARACTERISTIC

The maximum amount of compression and expansion to be used in the C-type system was more or less arbitrarily set at 20 dB. This seemed a natural goal, neither too little nor too ambitious, moreover offering the possibility of adapting existing B-type integrated circuits before dedicated C-type integrated circuits would become available. Nevertheless, it was necessary to determine an optimal spectral distribution of the noise reduction. If the frequencies to be treated were restricted to as narrow a range as possible, compatibility would be improved, noise-modulation performance enhanced, and there would be fewer troubles caused by recorder response irregularities at the frequency extremes.

In connection with the development of the B-type system, beginning in 1967, listening tests were made to determine what range of frequencies had to be treated to bring the noise of 3¾-in/s (9.5-cm/s) open-reel recording into subjective spectral balance using moderate to high listening levels, such that the tape hiss, with noise reduction, was discernible but not excessive. Thus the high-pass filter cutoff frequency used in the B-type circuit was set at 1.5 kHz. This cutoff frequency was retained in adapting the circuit for use with cassettes in 1969, although the filter configuration was changed to provide more noise reduction in the 300 Hz to 1.5 kHz range, as well as improved noise modulation performance.

The same kinds of listening tests were made in the development of the C-type system, using high-quality Type II cassette tape, 70 μs equalization, a quiet residential listening environment, and volume settings corresponding to rather loud listening conditions; that is, as in the B-type tests, the volume was set such that tape noise with noise reduction was perceptible but not annoying. Many filter configurations and combinations were tried; in the early stages of the development it had been hoped that one of the two stages could be left as a standard B-type circuit, for easy switchable compatibility between B-type and C-type operation. However, the spectral distribution tests eventually proved that this placed too heavy a burden on the second stage; it was required to produce substantially more than 10 dB of noise reduction in the several hundred hertz to 2 kHz region. The solution to this problem was to abandon the attempt to retain one stage as a standard B-type circuit; both circuits had to be nonstandard. Unfortunately, this approach increased the switching and component complexity, but it yielded a system in which the dynamic action burdens are more evenly shared between the two stages. The listening tests ultimately set the filter cutoff frequencies of the two circuits equal, at two octaves below that of the B-type circuit, namely at 375 Hz.

The use of equal cutoff frequencies yields a full compounding of the frequency discriminations of the two circuits, giving a steeply rising overall characteristic. This results in leaving the low-frequency region, in which little treatment is necessary, essentially untouched while providing substantially the full amount of noise reduction above about 500 Hz. The resulting noise reduction begins at about 100 Hz (3 dB), is about 8 dB at 200 Hz, 16 dB at 500 Hz, and is essentially 20 dB above about 1 kHz. This characteristic was determined using several types and qualities of loudspeakers and headphones, with both daytime listening and late night listening, when the ambient noise level (in San Francisco) is significantly reduced.

The cassette recorders used in the tests were standard production models selected for low hum levels. The selection process revealed such wide variations in hum level and character that only the best recorders were used in the final tests, so that hum reduction would not be a factor in determining the noise reduction characteristic. It was abundantly clear that it is possible to design recorders which are free of audible hum; a specification simply had to be established for allowable levels for the power line fundamental and each of its harmonics. Thus the shape of the low-level noise reduction characteristic was set only on tape noise considerations; with good head preamplifiers, tape noise predominates down to about 200 Hz. Below this frequency the audible noise, relative to noise at higher frequencies, is negligible with C-type noise reduction switched in, from either the amplifier or the tape.

3 SPECTRAL CHARACTERISTIC—HIGH FREQUENCIES

During the tests to determine the low-frequency characteristics of the system, attention was also directed to the high-frequency end of the spectrum. Consideration of the shape of the CCIR noise-weighting characteristic (Fig. 5), which was established for wideband, relatively low noise audio systems, shows that there is a significantly reduced need for noise reduction at extremely high frequencies (above about 10 kHz). Cassette tape recording has problems with record/playback frequency-response reliability and tape saturation in this frequency region. Moreover, with certain kinds of signals, compressor/expander tracking accuracy is affected. It seemed that the introduction of a new noise reduction system could be used as an opportunity to optimize the overall performance of the cassette medium, including the noise reduction system, using the above facts.

3.1 The Midband Modulation Effect

Compressor/expander complementarity requires not only that the expander have the inverse characteristics of the compressor, but also that the transmission channel between the compressor and the expander preserve relative signal amplitudes, and preferably also phases, at all frequencies within the bandwidth of the signals compressed. As received by the expander, changes in level caused by the transmission channel are indistinguishable from signal processing by the compressor. The resulting errors in the expanded signals can be significant and audible, depending on the spectral content of the signals. With the sliding band B-type system, the most audible error is not the direct effect on very-high-frequency signals themselves, but rather the modulation effect on mid-frequency signals, such as in the several hundred hertz region. For discussion purposes, this effect will be referred to as the midband modulation effect.

In wideband companders an amplitude error at the controlling frequency will manifest itself to the same

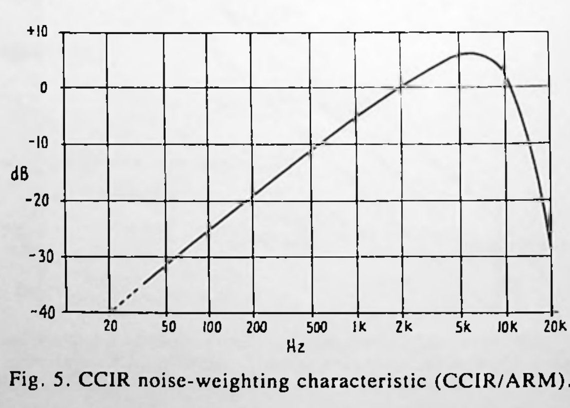


Fig. 5. CCIR noise-weighting characteristic (CCIR/ARM).

degree in all other portions of the spectrum; this may or may not be acceptable. In sliding band companders (B-type and C-type) an error at a dominant high frequency is substantially multiplied at mid-frequencies. (Conversely, if the controlling frequencies are at mid-frequencies, as they usually are, then any errors at the high-frequency extreme are reduced; this is an advantage of sliding band companders.) The midband modulation effect is rare with normal music sources; it may, for example, be audible with intermittent high-level high-frequency signals such as brushed cymbals in combination with a more or less continuous low-level mid-frequency sound, such as background violins. In such a case, the violins may be modulated in amplitude, even after decoding, because the cymbals cause the encoder band to slide without a complementary sliding of the decoder band. This effect is basically a frequency response error effect, as opposed to a tape saturation effect; it might be caused by inaccurate biasing and equalization or by gap loss, poor azimuth, and the like. However, the effect will be worse if there is also saturation in the controlling frequency region.

Reduction of the midband modulation effect is one reason for the incorporation of sharp low-pass filters, popularly known as multiplex (MPX) filters, in audio products using the B-type noise reduction system. Such band-limitation filters have corner frequencies at the edge of the useful bandpass of the system (about 16 kHz) in order to avoid limiting the system bandwidth unduly. Such filters have several functions:

1) Attenuation of subcarrier components and the 19 kHz pilot tone used in FM broadcasting, in order to avoid bias "birdie" beats (whistles), impairment of the noise-reduction action, and encoder/decoder mistracking.

2) Attenuation of tape recorder bias which may leak into the signal circuits, in order to avoid encoder/decoder mistracking.

3) Attenuation of supersonic signal components or of spurious radio frequency components in the encoder input signal which may otherwise result in audible intermodulation products and/or bias birdies.

4) Attenuation of supersonic tape noise or other transmission channel noise at the decoder input, in order to avoid encoder/decoder mistracking.

5) A signal bandwidth definition to promote complementarity of the encoder/decoder—that is, to reduce the midband modulation effect.

Strictly speaking, if an ideal channel exists between the encoder and the decoder, then the input filter to the decoder should be disconnected, as its inclusion theoretically results in a slight noncomplementarity (the encoder signal is subjected to one stage of filtering, the decoder to two). However, removal of the decoder input filter must be done with caution because of the considerations listed above.

3.2 Spectral Skewing

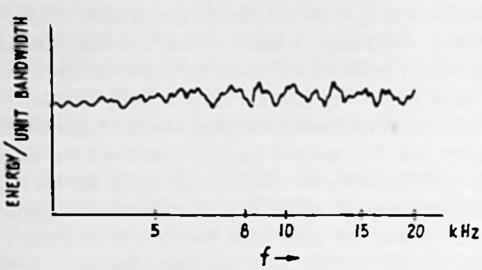
The solution to the midband modulation problem is rather surprising in its simplicity, and is termed spectral skewing. Advantage is taken of the fact that the high-frequency signals which cause the problem are usually complex in nature; that is, they occupy a relatively broad band of frequencies and are not at single discrete frequencies. The method used is to subject the signals to be processed by the compressor to an abrupt high-frequency drop-off which is within the useful bandpass of the system but somewhat below the frequency at which the record/playback response becomes highly unreliable. A corner frequency of 10–12 kHz fulfills these conditions. In this way the distributions of the signals processed by the compressor are altered or skewed such that the compressor action is significantly less susceptible to the influence of signals beyond the abrupt roll-off frequency. Signals processed by the expander are subjected to a complementary boost so that an overall flat frequency response is maintained. The spectral skewing network is situated at the compressor input; the de-skewing network, with complementary characteristics, is located at the expander output.

The spectral skewing principle as applied to sliding band companders can best be understood by reference to Fig. 6. Fig. 6(a) shows the spectrum of a signal that might provoke the midband modulation effect (such a signal might be generated by a wideband percussive sound). The compressor control circuit preemphasis results in an energy spectrum as shown in Fig. 6(b). After rectification, the peak in the preemphasized ac control signal spectrum provides the dc signal that controls the sliding band action of the compressor.

Fig. 6(c) illustrates the different frequency responses of four tape recorder channels, a, b, c, and d. The effect on the spectrum of Fig. 6(a) is to cause four different spectra [Fig. 6(d)] to be present in the control circuit of the expander, resulting in the four dc control signals shown; clearly, errors in decoding will result.

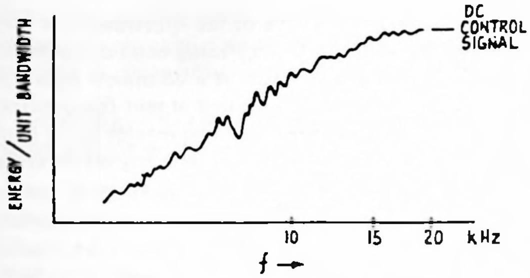
An idealized spectral skewing characteristic [Fig. 6(e)] causes the compressor and expander to generate the same dc control signal in each case, as shown in Fig. 6(f), which results in accurate decoding not only of the high-frequency signals, but also of any other signals at lower frequencies. Note that the network does not eliminate the sliding of the frequency band. Indeed, it may be only slightly reduced. However, the sliding now becomes recoverable during playback.

The spectral skewing characteristic used in the C-type system has a simpler and more economical form than the idealized curve of Fig. 6(e). A 12 dB notch characteristic is formed by combining the input and output signals of a resonant notch filter with a center frequency of 20 kHz and a Q of 1. The resultant characteristic within the audio band can be seen in Fig. 7. Compare this with Fig. 8, which shows representative measured high-frequency response curves for several typical cassette recorders. These curves show that, for levels below saturation, the typical recorder in good adjustment has little deficiency in response below 10–20 kHz. Hence, at most levels the spectral skewing network will ensure that there will be a significantly



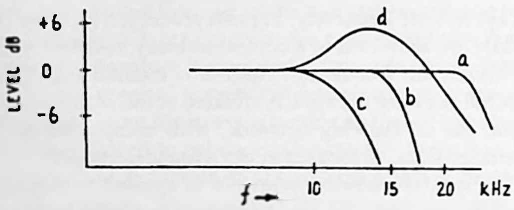
(a)

Fig. 6(a). Representation of spectral distribution of signal having a significant wideband component.



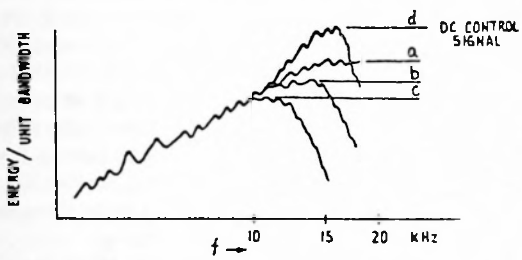
(b)

Fig. 6(b). The signal of (a) after control amplifier preemphasis in the compressor.



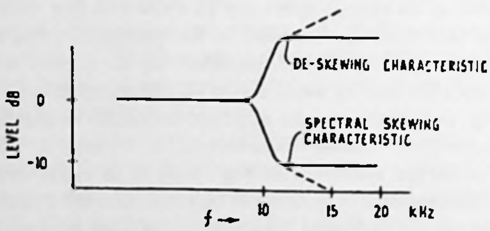
(c)

Fig. 6(c). Four different tape recorder frequency responses (a, b, c, d).



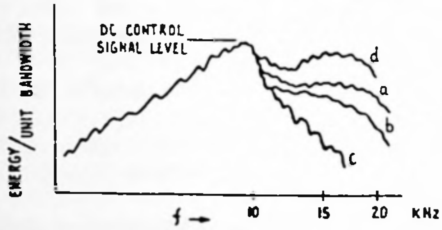
(d)

Fig. 6(d). The signal of (a) compressed and then sent through recorder channels a, b, c, d, resulting in the above signal distributions at the point of rectification in the expander; different dc control signals a, b, c, d are thereby produced.



(e)

Fig. 6(e). Idealized spectral skewing characteristic.



(f)

Fig. 6(f). As (d), but with spectral skewing treatment at the input of the compressor; the same dc control signal is produced in the expander with the four different recorder responses a, b, c, d.

Fig. 6. Example showing how spectral skewing tends to desensitize the noise reduction system to recorder errors at very high frequencies.

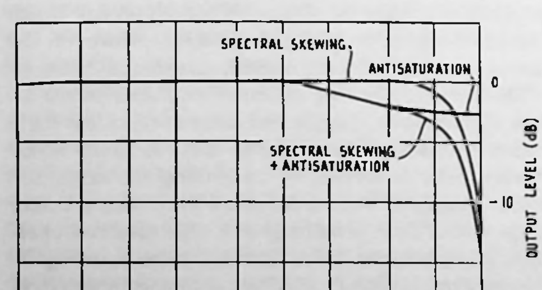


Fig. 7. Spectral skewing and antisaturation characteristics used in C-type system. Overall antisaturation effect is produced by combination of the two characteristics.

reduced discrepancy in the decoder control signal caused by uncertainties in response at extremely high frequencies.

The spectral de-skewing network used during decoding results in about a 12 dB loss of noise reduction in the 20 kHz region, leaving only about 8 dB of noise reduction. However, reference to the CCIR weighting curve (Fig. 5) shows that the frequencies above 10 kHz are on the steeply declining portion of the curve. In the 20 kHz region the ear is some 30 dB less sensitive to noise than in the 5 kHz region; this fact makes the spectral skewing technique possible.

The reduced psychoacoustic need for maintaining substantial noise reduction at frequencies above 10 kHz is the high-frequency counterpart of the ordinarily observed ability of the B-type noise reduction system to provide a subjectively useful amount of noise reduction even though the low frequencies are not treated at all. Good engineering can eliminate hum, which, as mentioned previously, is the only low-frequency noise which is subjectively troublesome in cassette tape recording.

Note should be taken that the use of a spectral skewing network does not obviate or replace an overall band-limitation filter (MPX). As discussed, band-limitation filters used in both recording and playback have several functions in addition to reducing the midband modulation effect. Therefore, even in the case of the highest quality recorders, it is essential to have band-limitation filters and to use them. Cleaner, more accurate recordings will be the result. It may be noted, however, that when spectral skewing and de-skewing are employed, as in the C-type system, then the band-limitation filters may have comparatively high cutoff frequencies (such as 20 kHz), without provoking the midband modulation phenomenon (but a switchable 19 kHz notch should be provided for recording FM broadcasts).

4 SATURATION REDUCTION

Inspection of Fig. 8(d) shows that high-frequency saturation is a serious problem in cassette recording. Usable peak levels at lower frequencies are some 8–10 dB higher than shown in this particular graph, with an even further deterioration of performance at high frequencies.

A useful by-product of the use of the spectral skewing network is the reduction of very high frequency saturation. Thus the network not only desensitizes the compressor to the frequency components likely to cause trouble during decoding, but it also reduces the chance for recording deficiencies at those frequencies, compounding the advantage. The significant improvement observable in single-tone frequency response curves is likely to be interpreted as the advantage of the technique; the improvement is easy to demonstrate graphically. The spectral skewing network, by itself, improves the high-frequency saturation performance of cassette tapes by several decibels in the 10–20 kHz region. However, cassette tapes suffer from saturation problems down to frequencies as low as 2 kHz. To accommodate this it is not possible to increase the bandwidth of the spectral skewing notch, or to extend the effective cutoff frequency downwards significantly, for two reasons: 1) the noise reduction effect would be audibly impaired, and 2) the effectiveness of the spectral skewing network in treating the midband modulation effect would be reduced. The CCIR weighting curve (Fig. 5) shows that full noise reduction action must be maintained up to about 10 kHz. Moreover, the efficiency of the spectral skewing action is dependent upon a relatively abrupt characteristic [Fig. 6(e)]. On the other hand, at least in approximately the 2–8 kHz frequency range, the saturation characteristics for typical cassette tapes are comparatively gradual, as can be seen in Fig. 8(d) (0 dB curve).

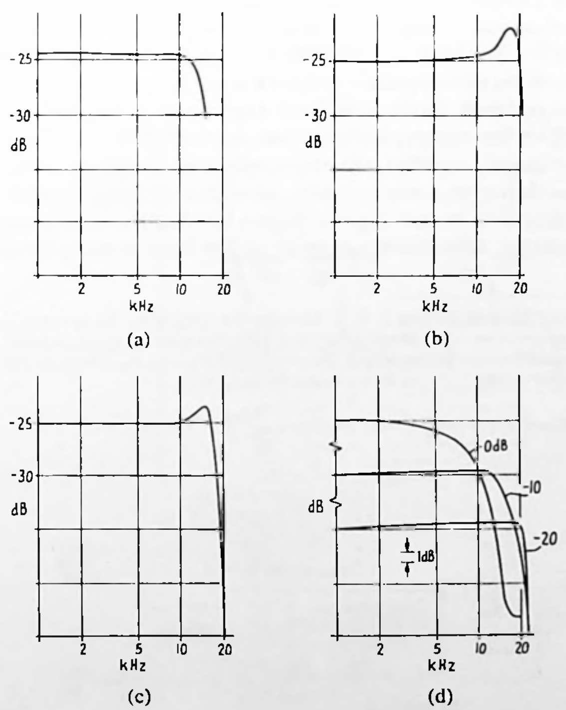


Fig. 8. Measured high-frequency responses of four typical cassette recorders.

The above considerations point to the need for a different method of solving the saturation problem in the mid-high-frequency area. A changed tape equalization characteristic could be used, but this would have a direct bearing on the overall noise level. An equalized high-frequency limiter could be employed, but this would be costly and in addition require complementary treatment during decoding. Headroom extension [3] is relevant but also has the drawback of cost.

The following antisaturation method, which is both simple and effective, is incorporated into the C-type noise reduction system. Note that in a dual-path compressor or expander circuit the output at very low signal levels is provided mostly by the noise reduction path. For 10 dB of dynamic action, the contributions of the main and the noise reduction paths are in the ratios of 1 and 2.16, respectively. At high signal levels the roles of the two paths are reversed: the main path provides the predominant signal component, and the further path contribution is negligible.

The saturation-reduction method is based on the above observations; namely, an equalizer providing the required attenuation of high-frequency drive is placed in the main path of the compressor, as shown in Fig. 9. At high signal levels essentially the full effect of the equalization is obtained, with a consequent reduction in high-frequency saturation. However, at low levels the equalization effect is reduced, since the contribution of the noise reduction path becomes significant. If, for example, the antisaturation network provides for a 3 dB attenuation at a particular frequency, then, ignoring phase considerations, the low-level effect will be:

$$0.71 \times 1 + 2.16 = 2.87 \quad (9.2\,\text{dB}) .$$

That is, a 3 dB reduction in high level recording drive is obtained for a 0.8 dB loss in noise reduction effect.

It is necessary that a complementary correction be provided on the playback side, so that the signal is restored. The type of correction required can be deduced from Fig. 9, which shows a symmetrical compressor and expander configuration, including the placement of networks in the main signal path. Let the input signal to the compressor be x, the signal in the recorder channel be y, and the output signal of the expander be z. Let F_1 and F_2 be the transfer characteristics of the noise reduction path of the compressor and expander, respectively, and F_{AS} the transfer characteristic of the antisaturation network. Let F'_{AS} be the required compensating characteristic in the decoder.

$$y = (F_{AS} + F_1)x \tag{2}$$

and

$$z = yF'_{AS} - zF_2F'_{AS} . \tag{3}$$

Thus

$$z = \frac{F'_{AS}F_{AS} + F_1F'_{AS}}{1 + F_2F'_{AS}}\, x . \tag{4}$$

Inspection shows that $z = x$ if $F_1 = F_2$ and if $F'_{AS} = 1/F_{AS}$.

The above shows not only that the two noise-reduction networks should be identical, as is known in the A-type and B-type systems, but also that the antisaturation compensation network in the decoder should have an inverse characteristic to that of the network employed in the encoder.

The antisaturation network used in the C-type system is a simple shelf network (two resistances and one reactance) with time constants of 70 and 50 μs, corresponding to turnover frequencies of about 2.3 and 3.2 kHz, respectively. High-frequency attenuation is provided in the encoder, with a corresponding boost in the decoder. Referring to Fig. 7, this results in a saturation reduction of about 1 dB at 2 kHz, 2.3 dB at 5 kHz, and 2.8 dB at 15 kHz.[1] At frequencies above 10 kHz, the spectral skewing network augments the overall antisaturation effect, as shown in Fig. 7.

5 BLOCK DIAGRAM—C-TYPE CIRCUIT

Based on the principles discussed, Fig. 10 shows the basic block diagram of the C-type compressor and expander. The networks N_1 and N_2 are the noise reduction side chains. The spectral skewing network is placed at the input of the high-level stage of the compressor, thereby affecting the operation of both the high- and the low-level compressor stages. Note that the de-skewing network, being situated at the output of the whole system, has no effect on the operation of either of the expander stages; its only function is to restore an overall flat frequency response. For simplicity and economy the antisaturation network is placed only in the low-level stage.

Fig. 11 includes more complete diagrams of the individual stages and also shows the distinctions between the B-type and C-type systems. The figure shows the function changes necessary to provide a switchable record/play circuit with either B-type or C-type capability. If desired, further switching can be provided so that one spectral skewing network and one antisaturation network can serve in both the record and the play modes with the required complementary characteristics.

For B-type operation the low-level stage and spectral skewing and de-skewing networks are switched out of the circuit; the filter frequencies, overshoot suppression level, and control circuit smoothing time constants are set to the B-type values.

For C-type operation, the spectral skewing and de-skewing networks are switched in. The low-level stage is connected in series, with its preset low-level area of action, including overshoot suppression; the variable-filter quiescent cutoff frequency is preset to 375 Hz; and the antisaturation network is connected in the main signal path. In the high-level stage the fixed and variable filter frequencies are both lowered to 375 Hz. The latter changes, together with the retention of control circuitry with B-type characteristics, result in a modified spectral distribution and slightly higher level at the side chain output. Overshoot suppression for these conditions is

[1] Thanks are due to K. J. Gundry for reviewing the saturation properties of contemporary high-performance cassette tapes and for recommending this 50 μs/70 μs characteristic for use in the C-type noise reduction system.

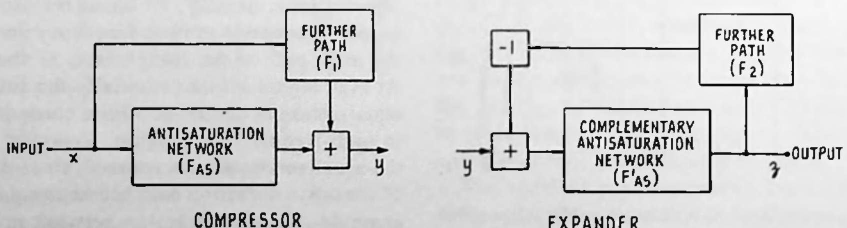


Fig. 9. Placement of antisaturation networks in main signal paths of compressor and expander.

optimized by setting the suppression threshold 3 dB higher than in B-type operation. The potential maximum overshoot is therefore somewhat greater than in the B-type system.

In the development of the C-type circuit, the opportunity was taken to incorporate full-wave rectification in the control circuitry of both stages (for economy, half-wave rectification is normally used in the B-type circuit). This significantly reduces distortion caused by control signal ripple modulation and makes it possible to decrease the smoothing time constants used. Halving the time constants eliminates the last vestiges of noise tails upon abrupt cessation of high-amplitude high-frequency signals (which generate the largest control signals). The attack time constants are also reduced, which tends to offset the higher overshoot suppression level of the high-level stage, as well as the (somewhat lower) overshoot contribution of the low-level stage. As shown in Fig. 11, the smoothing time constants of the high-level stage are made switchable in order to retain compatibility with the B-type characteristics.

A minimum amount of staggering is used in separating the areas of action of the two circuits, consistent with maintaining the overall compression ratio at a maximum of about 2. Fig. 12 shows the single-tone compression characteristics of the high-level stage; spectral skewing

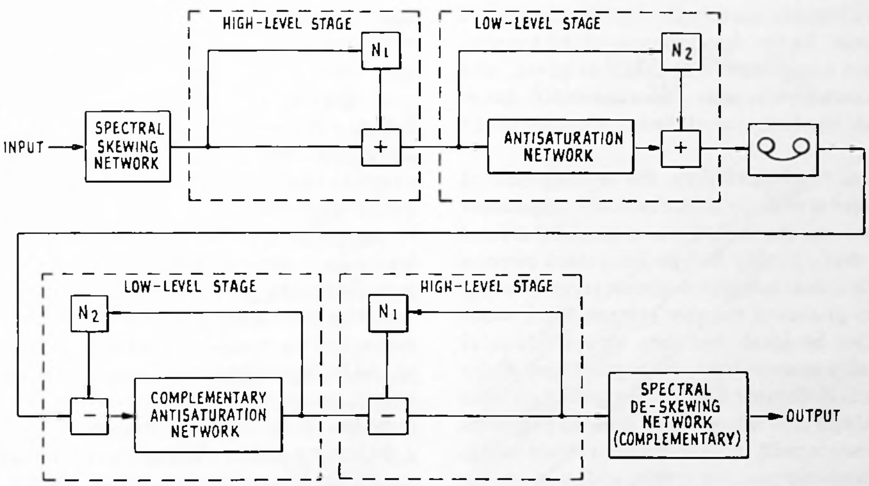


Fig. 10. Basic block diagram of C-type compressor and expander. N_1 and N_2 are the noise reduction networks.

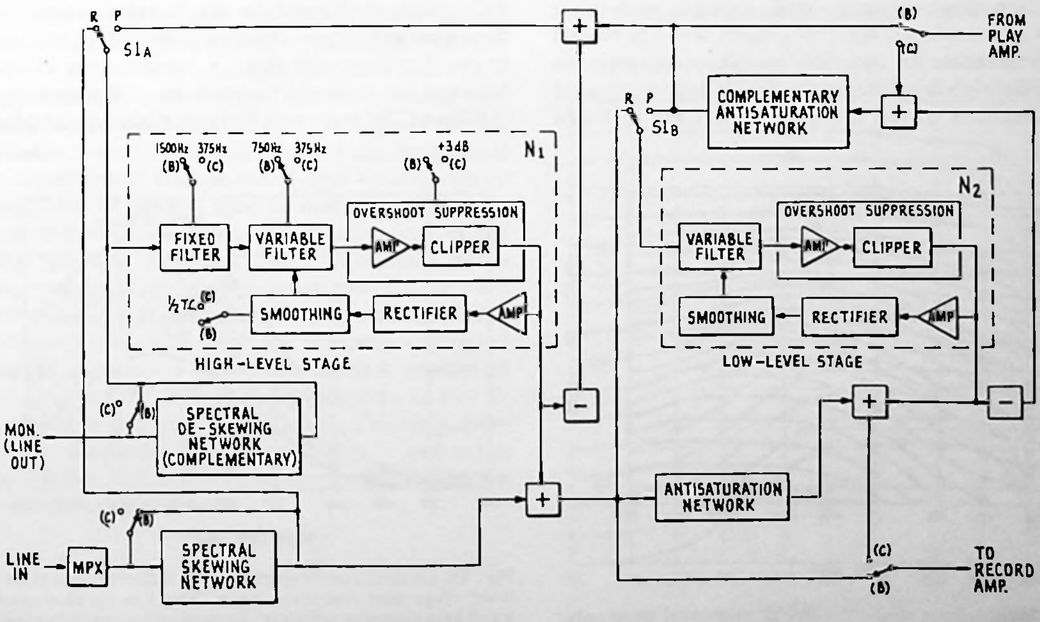


Fig. 11. Switchable record/play and B-type/C-type processor.

is omitted for clarity. Note that the frequencies from about 1 Hz to 8 kHz include areas which have compression ratios in the region of 2. Thus the low-level stage must have an action area arranged to be well clear of these. Above and below this frequency range the compression ratios are generally lower, so that some overlapping of the characteristics is possible in these ranges.

Staggering is achieved by increasing the effective amplification employed in generating the control signal of the low-level stage. As previously discussed, this increased gain is provided partly in a fixed way and partly in a variable way, by virtue of the low-level signal boosting of the high-level stage.

Fig. 13 shows the fixed and variable elements of the amplification difference used in the low-level stage of the C-type system. In the development of the system, the variable-gain component was taken as given, and the fixed-gain component was experimentally determined to provide the best overall fitting together of the characteristics.

For convenience and flexibility, the development of the C-type system was done with discrete components and FETs. However, the design was organized around the possibility that existing B-type integrated circuits could be used, in a dual integrated-circuit layout, during the introductory phases of the new system. Such adaptations would not be ideal, but they would be useful and also generally demonstrate the capabilities of the new system until dedicated C-type integrated circuits could be developed and introduced. Interim performance compromises would include higher circuit noise, incorrect overshoot-suppression levels, and an extension of the operating conditions of the variable resistance elements into ranges not originally designed or specified in B-type integrated circuits (resistances a factor of 2 higher than those normally used, together with those a factor of 2 lower).

To compensate for incorrect overshoot-suppression level in the high-level stage, the dual B-type integrated circuit versions of the C-type system have included cross-coupling of the transient components of the low-level stage control signal to the control circuit of the high-level stage. Audible distortion under transient signal conditions is thereby avoided, but some transient spectral alterations are an unavoidable by-product on some kinds of program material. Nothing is done to lower the overshoot suppression level of the low-level stage, but this does not cause any audibly significant effects. Thus there is a certain incompatibility between the recordings made with early (1981–1982) dual B-type integrated circuit embodiments and the subsequently phased-in dedicated C-type integrated circuit designs (1982–). The incompatibility is in the direction of subjectively favoring dedicated C-type circuit recordings which are decoded with early dual B-type circuit adaptations: on some program material the sound is boosted or brightened on a transient basis; the reverse combination produces a dulling effect.

In adapting the C-type design initially to the available B-type integrated circuits and generally to integrated circuit technology, it was also necessary to match the external impedances used in the variable filter to the variable-resistance characteristics of the integrated circuits. The many B-type integrated circuit manufacturers have always had difficulties in designing and manufacturing variable resistances to function accurately over a ratio of some 1000:1. Therefore, it was essential to specify C-type filter impedances which would be as compatible as possible with such integrated circuit resistance characteristics, especially bearing in mind the increased requirements of the C-type circuit.

Referring to Fig. 11 and references [1] and [2], the fixed filter is simply a series capacitor and shunt resistor; there is no problem with this stage. The variable filter is a series-connected parallel combination of a resistor R and a capacitor C (with a turnover frequency of $1/2\pi RC$) which is shunted by the variable-resistance R_v; this combination provides a variable shelf characteristic. In the B-type circuit there is a one-octave difference between the turnover frequencies of the two sections (1500 and 750 Hz), which yields a quasi two-pole filter

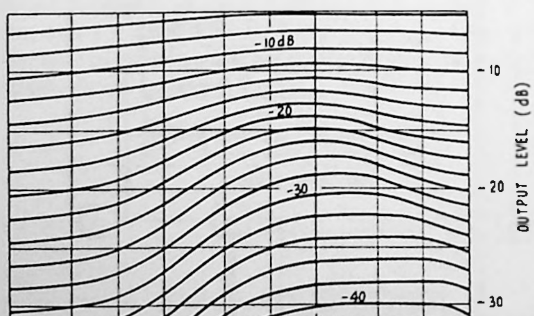


Fig. 12. Input–output characteristics of high-level stage only, without spectral skewing.

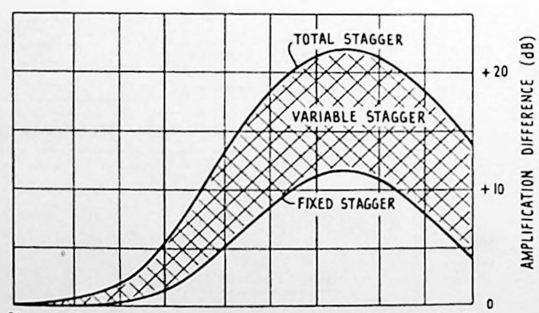


Fig. 13. Control-circuit amplification difference between high-level stage and low-level stage. The low-level stage has a fixed gain increase which is augmented by a variable increase caused by the signal-processing action of the high-level stage.

with a more steeply rising noise reduction characteristic in the presence of signals than a simple single-pole filter might provide.

In lowering the fixed filter cutoff frequency in the C-type circuit by two octaves, the available integrated circuit characteristics made it seem unlikely that the variable-filter turnover frequency could be lowered by a similar amount. For this reason, and a further reason to be discussed, the variable-filter turnover frequency was lowered by only one octave (to 375 Hz).

The component values used in the variable filter are a compromise which stretches the capability of R_v at both ends of the range, regarding limiting values and their repeatability, as well as repeatability within the range (especially at high resistances). Attention in dedicated C-type integrated circuit designs has been directed to these matters.

Using the same turnover frequency (375 Hz) for the fixed and variable sections causes this particular filter configuration to perform in the same manner as a single-pole variable filter.[2] Replacement of the combination with a single-pole filter saves a resistor and a capacitor; this saving can be realized in the low-level stage, which does not have to be switchable to B-type operation.

The performance limitations of available integrated circuit variable resistances thus was one consideration in selecting a C-type filter arrangement which, by itself, is not as efficient as the filter of the B-type circuit with respect to noise modulation. However, the steepness compounding effect of the two-stage arrangement used in the C-type system more than compensates. The resulting noise modulation margin of safety, while not quite that of the B-type system, is adequate, if not good, on nearly all program material, especially taking into account the lower real noise level achieved in the presence of signals.

Even if available integrated circuit characteristics had made it possible to lower the variable-filter turnover frequency by a full two octaves to retain quasi two-pole performance in each stage, it is unlikely that such a choice would have been made. Throughout the development, the midband modulation effect, transposed two octaves lower than in the B-type system, was a hazard borne in mind at least as much as noise modulation (thereby stimulating the development of the spectral skewing technique). It is inevitable that a steeply rising noise reduction characteristic (in each stage) results in a greater susceptibility to the midband modulation effect. Even with the advantages afforded by spectral skewing, it would be difficult to predict all the possibilities for error in the mass production of C-type machines and prerecorded tapes. Thus it is to be hoped that the modest filter characteristics used in the C-type system will in due course prove to have been a good design compromise.

[2] Thanks are due to the audio group at Sony Corporation for pointing this out.

6 PERFORMANCE

6.1 Characteristic Curves

Fig. 14 shows the overall input–output transfer characteristic of the C-type compressor at 1 kHz. The high-level stage by itself is also shown, in order to demonstrate how the actions of the two stages blend together without increasing the maximum compression ratio.

In Fig. 15 the overall single-tone compression characteristics can be seen. The reduced drive to the tape at very high frequencies and high levels significantly extends the frequency response which can be obtained routinely. High-frequency distortion under test and real signal conditions is also notably reduced.

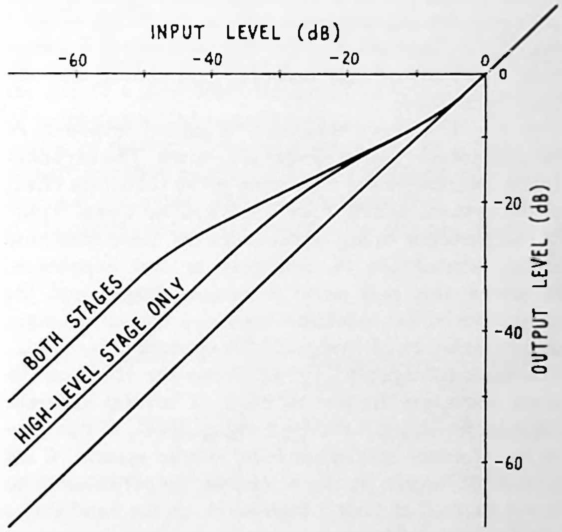


Fig. 14. Input–output transfer characteristic of C-type compressor at 1 kHz. High-level stage also shown.

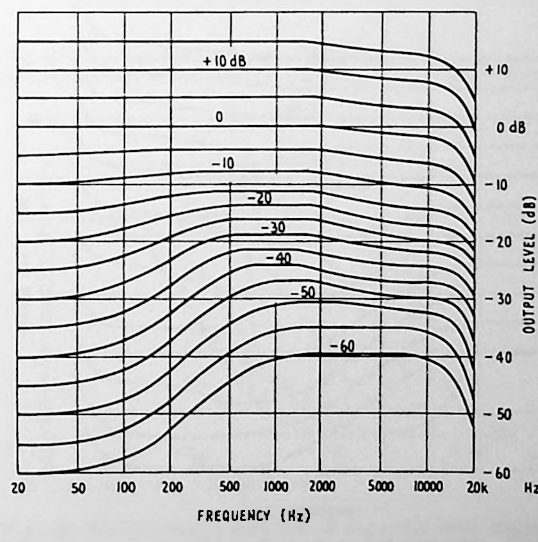


Fig. 15. C-type compression characteristics.

The corresponding expansion curves are given in Fig. 16. While the curves of Fig. 15 and 16 do not appear to be symmetrical, or complementary, it should be noted that the expander is normally not fed by an unprocessed signal; rather it is always supplied with a signal from the output of a compressor. Consideration of these curves will show that the expander characteristic restores the compressed signal to its original state. Reference to Eq. (4) shows that the restoration is theoretically exact in all respects: frequency response, phase, and dynamic properties. This ideal can be achieved in practice to any extent desired in the tolerancing and matching of components and operating conditions. However, from Figs. 14 and 15 it should be noted that the maximum compression ratio of the C-type system prevails over a significantly greater range of amplitudes and frequencies than in the B-type system. For optimal reproduction it is thus essential to maintain high standards of tape recorder gain setting and frequency response.

In Fig. 17 the sub-threshold frequency responses of the compressor and expander are shown. The expander curve determines the maximum noise reduction effect of the system, which is obtained with no signal input.

The presence of signals reduces the noise reduction effect attributable to compression and expansion. However, this real noise reduction merges into the subjective noise reduction provided by the masking effect, on which all compressor/expander noise reduction systems depend. Fig. 18 shows the effect on the noise reduction frequency band of several different frequencies applied at a high signal level, in this case at the nominal maximum level of the system, 0 dB (reference level). It is, of course, impermissible to boost a signal at such a high level, so the band slides upwards to eliminate the boosting action. Low-level signal components at higher frequencies continue to be boosted. The complementary attenuation provided by the expander then produces whatever real noise reduction action is possible under those signal conditions.

The curves of Fig. 18 were obtained by mixing a sweeping probe tone at a level of −65 dB into the compressor input signal and detecting the tone at the output with a tracking wave analyzer.

Fig. 19 shows the operation of the compressor in response to a 500 Hz signal over a range of levels; the corresponding expander characteristic is given in Fig. 20. The compressor and expander progressively slide the frequency band upwards with increasing signal levels; the masking effect, working in cooperation with the expander, creates the overall illusion of a low, virtually unchanging noise level.

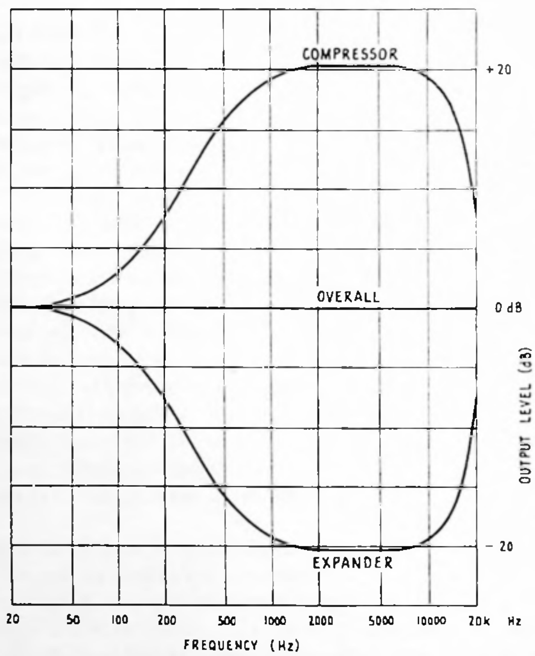


Fig. 17. Low-level (sub-threshold) frequency response of C-type compressor and expander.

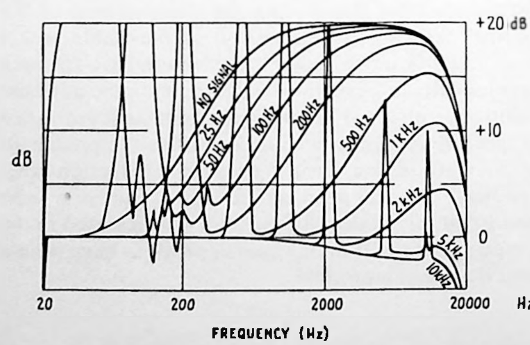


Fig. 18. Sliding band action of compressor in response to high-level (0 dB) signals at the frequencies shown (−65 dB probe tone).

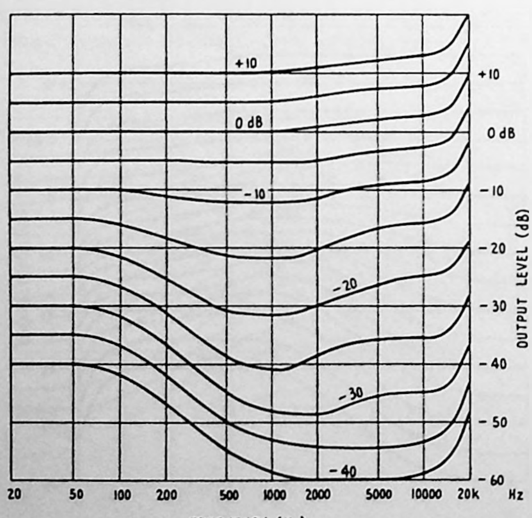


Fig. 16. C-type expansion characteristics.

Figs. 12–20 have shown the response to a single tone or to the simulation of a dominant signal together with other signals at much lower levels. Often, however, the signal will comprise a complex combination of frequencies and levels. The action of the spectral skewing network in tilting or altering the spectrum of very high-frequency signals has been discussed. A similar, but variable, type of skewing action affecting the operation of the system also takes place at lower frequencies, by virtue of the two-stage sliding band layout. With complex input signals the high-level stage alters the spectral content of the signal. Thus the low-level stage is actuated by a signal which is not only different in level but different in spectral balance. This has the tendency of spectrally spreading out the chance for error in the decoding function. If the high-level stage is controlled by signal components in a certain frequency range, then the low-level stage will tend to be controlled by signal components somewhat higher in frequency. Thus the spectral shifting effect reduces the overall dynamic and frequency response errors of the decoded result when the tape recorder has an uneven frequency response.

6.2 Compatibility

In the design of the B-type system, the subjective acceptability of encoded recordings when used with conventional players was a matter of some concern; the dual inventory production of cassettes was not judged to be commercially feasible. Thus the B-type system represents a three-way balance of: 1) operating characteristics which are practical, economical, and technically safe in implementation, 2) a noise reduction effect which is sufficient to be acknowledged as useful, and 3) an amount and type of dynamic action which could be judged as "compatible" on simple players.

The conditions relating to the development of the C-type system were somewhat different. For one thing, the majority of listeners were already adequately catered to with the B-type system, which gave greater latitude in the design of the new system. At least in the beginning, the C-type system would be an audiophile or critical listener system. Therefore the main consideration had to be the provision of a usefully increased amount of noise reduction without provoking undesirable side effects under full encode/decode conditions, taking into account the strengths and deficiencies of practical production model tape recorders and tapes. Thus no specific design concessions were made to the issue of compatibility (except to the further consideration of providing a circuit which would be B–C switchable). However, a certain compatibility happens to be a useful by-product of the design philosophy used in producing these noise reduction systems, namely, that the best treatment of the signal is the least treatment. If the action of the system is constrained to the bare minimum, with respect to the signal levels handled and the frequency ranges covered, then an inevitable consequence is that the bulk of the encoded signal is simply the original input signal. The compatibility of this or that kind of C-type encoded program material (with B-type reproduction, or with a tone control, or with nothing) is a matter of opinion. It may well be possible that the single-inventory production of certain types of C-type recordings or signal sources is workable. However, in relation to the various potential applications and markets, this matter must be judged on a case-by-case basis.

7 CONCLUSION

A new high-performance noise reduction system, designated C-type, has been developed for consumer applications. Primarily designed for use with cassette tapes, the system is switchable between the B-type and C-type modes.

The new system avoids many of the problems associated with large amounts of dynamic action through the use of a dual-level staggered action arrangement of series-connected compressors and expanders. Specific problems relating to cassette recording are addressed: an antisaturation method extends useful high-level high-frequency response and reduces distortion, and a technique called spectral skewing further reduces saturation

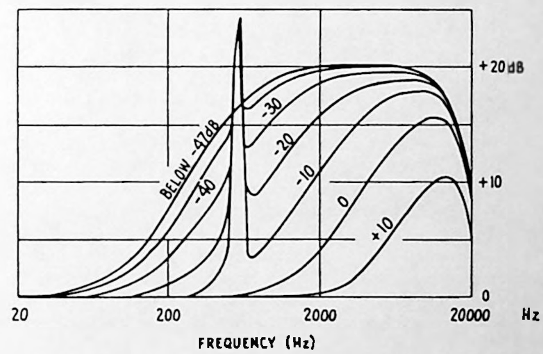


Fig. 19. Sliding band operation of C-type compressor with 500 Hz signals at the levels indicated (−65 dB probe tone).

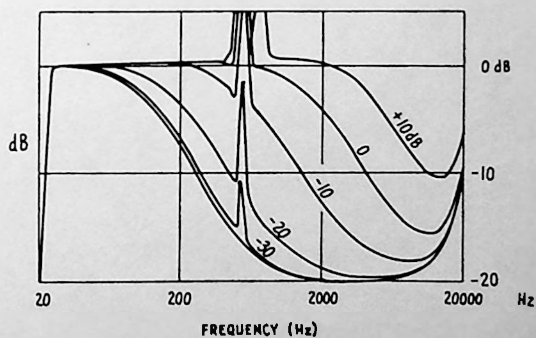


Fig. 20. Sliding band operation of expander with 500 Hz signal at the levels indicated (−45 dB probe tone).

and desensitizes the system to recorder errors at very high frequencies.

8 ACKNOWLEDGMENT

The author is grateful for the dedication and valuable contributions of many Dolby Laboratories staff members in the creation, implementation, and practical introduction of this new system, especially K. J. Gundry, I. Hardcastle, J. B. Hull, D. P. Robinson, E. A. Schummer, and S. J. Solari. Grateful acknowledgment is also made to a number of licensees whose persistence and enthusiasm provided the stimulation and encourgement to proceed with this venture.

9 REFERENCES

[1] R. M. Dolby, "A Noise Reduction System for Consumer Tape Applications," presented at the 39th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 18, p. 704 (1970 Dec.).

[2] R. Berkovitz and K. J. Gundry, "Dolby B-type Noise Reduction System," *Audio* (1973 Sept. and Oct.).

[3] K. J. Gundry, "Headroom Extension for Slow-Speed Magnetic Recording of Audio," presented at the 64th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 27, p. 1026 (1979 Dec.), preprint no. 1534.

[4] R. M. Dolby, "An Audio Noise Reduction System," *J. Audio Eng. Soc.*, vol. 15, pp. 383–388 (1967 Oct.).

THE AUTHOR

Ray Dolby was born in Portland, Oregon, in 1933, and received a B.S. degree in electrical engineering from Stanford University in 1957. From 1949–52, he worked on various audio and instrumentation projects at Ampex Corporation, and from 1952–57 he was mainly responsible for the development of the electronic aspects of the Ampex video tape recording system. After he was awarded a Marshall Scholarship, followed by a National Science Foundation graduate fellowship, he left Ampex in 1957 for further study at Cambridge University in England where he received a Ph.D. degree in physics in 1961, and was elected a fellow of Pembroke College. During his last year at Cambridge, he was also a consultant to the United Kingdom Atomic Energy Authority.

In 1963, he took up a two-year appointment as a United Nations adviser in India, and returned to England in 1965 to establish Dolby Laboratories in London. Since 1976 he has lived in San Francisco, where his company has established further offices and laboratories.

Dr. Dolby holds a number of patents and has written papers on video tape recording, long wavelength X-ray analysis, and noise reduction. He is a fellow and past-president of the AES, and a recipient of its Silver Medal Award. He is also a fellow of the British Kinematograph, Sound and Television Society, the SMPTE, and a recipient of its Samuel L. Warner Memorial Award and Alexander M. Poniatoff Gold Medal. In 1979 he and his colleagues received the Scientific and Engineering Award of the Academy of Motion Picture Arts and Sciences.

A NEW ANALOG RECORDING PROCESS FOR USE WITH CONSUMER RECORDING FORMATS

Stan Cossette
Dolby Laboratories, Inc.
San Francisco, California 94103

Abstract

A new analog recording process intended for use with consumer recording formats is described. The process, designated S-type, uses compression during recording and complementary expansion during playback to increase the dynamic range of the recording system. The resulting system provides up to 24 dB of noise reduction above 400 Hz and up to 10 dB of noise reduction below 200 Hz. Design goals for the new process are discussed and performance results are presented. Performance improvements of the new recording system when applied to the compact cassette format are presented as well.

1. Introduction

The Spectral Recording process was introduced by Dolby Laboratories in 1986 and has found great acceptance in professional format recording and transmission. Shortly after its introduction, requests were received for a consumer version of the process. Since the Spectral Recording (SR) process circuitry involves some 1,600 components as well as an extensive alignment procedure, it was not clear whether this request could be met. It was also questionable whether the resulting circuit would be inexpensive enough to allow widespread consumer use.

The initial experiments were conducted using modified SR circuits and gave us encouragement to continue. From our past experience with consumer noise reduction systems we compiled a list of required performance attributes for the new system.

1. Dynamic range improvement of consumer recording formats allowing subjectively noise and distortion-free reproduction at normal listening levels.

2. Low noise modulation.

3. Tolerance to record/play-path level and frequency response errors.

4. Subjectively good results when improperly decoding encoded program material.

5. The resulting system must be inexpensive (must be IC-based).

Design of the new system involved the use of psychoacoustic principles, computer modeling, and advanced analog integrated circuit techniques.

The resulting signal processor, labeled S-type, is an attempt at realizing these goals. The processor involves the use of well known companding techniques whereby the useful dynamic range of a recording device is increased by compressing the signal before recording and expanding the signal in a complementary manner during playback. The expansion process restores the input signal dynamic range and suppresses low-level artifacts added by the recording process such as noise and distortion. The S-type encoder uses 5 compressor stages to separately treat signals of different level and frequency. For high level signals, the encoder acts as a passive filter stage which attenuates signals at the audio frequency extremes thereby increasing headroom. Low-level signals are boosted by up to 24 dB by the encoder which leads to up to 24 dB of noise reduction during decoding.

2. Principles of Design

2.1 Least treatment

S-type employs an important design principle, expressed by R. M. Dolby in his report on SR as the principle of least treatment [1]. This principle, applied to an encoder design, can be illustrated as follows: Take the case of a simple sine wave signal at frequency F1. While the signal is at low levels, the encoder simply provides constant signal boost. As the signal increases in level, the encoder boost must be reduced at that frequency to prevent overloading of the recording medium. Figure 1 shows a simple compressor response in the presence of a high-level signal. In the simple compressor the boost at all frequencies has been modified or "treated" due to the presence of this tone. The premise of the least treatment principle is that there is great benefit in applying signal treatment only when and where necessary. Figure 2 shows a compressor response which more closely adheres to this principle.

The benefits of applying this principle are many and take on different forms depending on whether one is discussing encoder or decoder performance. The benefits to an encoder are that it results in a very dense, compressed signal which is relatively free from negative side-effects normally associated with such compression. Normally, the most objectionable side-effect of compression is the level-modulation of one part of the spectrum by another, more dominant, part of the spectrum. Application of the least treatment principle minimizes this effect. Another benefit is that signals encoded in this way are "easy" to decode. This is due to the spectrally-independent compression which makes errors in one part of the spectrum less likely to affect signal levels in other parts of the spectrum.

To appreciate the benefits as applied to a decoder one must think now of the complement of the action described for an encoder. That is, a decoder will attenuate or suppress non-dominant signals as much as possible. In a typical record path, these signals include noise, print through, distortion harmonics, modulation noise, and hum. Therefore a decoder that follows this principle should provide a steady noise floor even in the presence of signals.

The results of designing a compandor according to this principle are low distortion, low noise modulation, and excellent tolerance to level or frequency response errors in the recording or transmission channel. Furthermore, when there are errors during decoding, due to errors in the recording path or perhaps due to improper decoding (using another, similar process), the program material remains quite pleasant and relatively free from audible artifacts. In these cases, careful comparison to the original source material would reveal a difference but without this comparison it is unlikely a problem would be noticed. This claim, of course, depends on many factors such as the nature of the error, the type of program material, training of the listener and type of decoding used. However, this claim is based upon extensive listening tests using many different types of source material decoded in the presence of common errors which were known to cause problems with previous companding systems.

2.2 Bilinear Compression

While the goal of least treatment is easily visualized, its attainment, especially under economic constraints, is difficult. The S-type recording process achieves its performance by borrowing and modifying some of the circuit techniques used in previous noise reduction processors developed by Dolby Laboratories. The circuit uses dynamic signal processing in signal side-chains which are summed with a passive main path (see figure 3). This puts less performance demands on the side-chain itself since its effect will be increasingly diluted as signal levels increase. This allows a more cost effective design. It also allows a bilinear compression effect to be achieved which has many advantages [2-3]. A bilinear compressor provides compression over a limited range of input levels. Outside this range it has fixed gain or attenuation. Furthermore, the processor uses two staggered-action bilinear stages to achieve the required compression. This allows increased compression effect (boost) to occur without increased negative effects from the side-chain.

Figure 4 shows the response of the compressor of figure 3 to a fixed frequency signal which is swept in level from -60 dB to +10 dB. By plotting the output level of the compressor versus input level, the bilinear characteristics can be illustrated. At low input signal levels the compressor provides a fixed 10 dB of boost. At a certain level, the side chain output begins to fall. This is known as the compressor threshold. As the side chain output continues to fall, a point is reached such that the side chain contribution is an insignificant fraction of the composite output. This is marked on the graph as the finishing point. As the input signal is increased above the finishing point, the compressor curve becomes linear again with a gain of 0 dB. The region between the threshold and the finishing point is called the compression or "active" region. The side chain output is shown at the point V_2 to illustrate how the side chain output becomes a very small fraction of the composite output signal at high levels. 0 dB on the chart is known as the processor reference level. Since the compression effect is dependent upon signal level, a complementary expandor must be calibrated to this reference level.

2.3 Action Substitution

The S-type process uses techniques known as action substitution and modulation control to further improve its performance. As described in the paper by R. M. Dolby [1], action substitution is a way of combining a fixed and sliding band such that the advantages of each are retained while the disadvantages are minimized.

Figure 5 shows a family of curves that represent the response of a fixed band compressor. The compressor is intended to treat frequencies above 400 Hz. Each response curve shows the compressor gain in the presence of a signal of a different level. As the signal level increases, the compressor gain decreases. It can be seen that the characteristic of a fixed band is that it provides the same amount of boost throughout its passband. Unfortunately, it follows that any signals within the band that require reduction of this boost will cause loss of boost for the entire band (hence the problem with a wideband compandor).

A sliding-band compressor will slide in response to a dominant signal until the boost is correct at the dominant signal frequency. Figure 6 shows the response of such a compressor to a signal which is increasing in level or frequency. The quiescent or "no signal" response is identical to the fixed band compressor of figure 5. Its response in the presence of signals is quite different, however. The fourth response curve from the left represents a reasonable response to a high level, 500 Hz signal because there is no gain at 500 Hz. On this same response curve, however, there is 11 dB of gain at 10 kHz. It can be seen from figure 6 that a high-frequency sliding band compressor provides boost primarily at frequencies above the dominant signal.

Figure 7 shows the block diagram of an action substitution compressor. The output signal can be expressed as:

$$V_3(s) = V_1(s)\,[H_1(s) + H_2(s) - H_1(s)H_2(s)]$$

In this type of compressor, the greater output of either stage dominates and the resultant response is effectively improved over either stage alone. Figure 8 shows the response of a fixed and sliding band combined in this manner. Note that the high frequency boost has been improved over that of the fixed band alone, while the midrange boost has been increased over that provided by the sliding band alone.

In the S-type process, a low-pass filter is added at the fixed band output. This allows a significant amount of midrange boost to occur in the presence of dominant high frequencies. This filter also minimizes the effect of very high frequencies, which are the most difficult to reproduce, on the attenuation of the fixed band. This in turn provides the circuit with more resistance to high frequency response errors during playback (see figure 9). Use of action substitution provides a very frequency-selective compression effect in a very cost effective manner.

2.4 Modulation Control

Modulation control, introduced in the Spectral Recording process, is used in the S-type process to minimize the effect that high level signals have on the compressors. Modulation control takes advantage of the side-chain/main-path configuration in that once the amount of attenuation required in the side-chain is attained, no further attenuation is necessary. This requires that the control signal depart from its normal, directly proportional, relationship with the compressor output signal. This is accomplished by creating modulation control signals which oppose the control signals under certain conditions.

Figure 10 shows a family of response curves for a high frequency fixed band compressor in the presence of a 100 Hz tone. The number associated with each curve indicates the level of 100 Hz tone which produces the response shown. In figure 10a note that the compressor gain has been reduced to only 1 dB at 1 kHz due to the presence of a +20 dB, 100 Hz tone. When modulation control is added, the same compressor provides 5 dB of gain at 1 kHz under the same condition (see figure 10b).

Figure 11 shows a family of response curves for a high-frequency sliding band compressor in the presence of a 100 Hz tone. The level of the 100 Hz tone is indicated on each resultant response. Without modulation control, the band continues to slide even after the gain at 100 Hz has been reduced to 0 dB (see figure 11a). This results in a gain of only 4.5 dB at 1 kHz in the presence of a 20 dB tone. Figure 11b shows an improvement in the gain at 1 kHz under the same condition to 8.5 dB. By reducing the amount of gain reduction, or modulation, that a compressor produces in response to a signal, the goal of least treatment is more closely attained.

2.5 Anti-Saturation and Spectral Skewing

These techniques were introduced in the C-type process [3]. Anti-Saturation provides attenuation for high-level signals at the extremes of the audio band which reduces the chance of overloading the medium. Spectral skewing is also provided at the audio frequency extremes and desensitizes the compressors to signals at these frequencies. This means that the compressor action will be controlled by the area of the spectrum which is most reliably reproduced.

The S-type processor uses two stages of high-frequency anti-saturation and spectral skewing. This has the disadvantage of requiring more components but has the advantage of reducing high frequency compression ratios and desensitizing the circuit to signals above the audio range.

3. Circuit Block Description

The signal processing aspects of S-type can be understood by examining the block diagrams. In these diagrams, only the major elements are shown. Practical elements such as signal buffering, gain, attenuation and D.C. biasing circuitry are omitted to enhance clarity.

3.1 The S-type Processor

The S-type processor block diagram is shown in figure 12. The processor is designed so that a decoder can be created by simply putting an encoder in the feedback path of a high gain amplifier. This assures complimentarity as can be seen by the compandor transfer function.

$$H_c(s) = H_e(s)\ H_d(s)$$

$$H_d(s) = \frac{1}{1/A + H_e(s)} \Rightarrow \frac{1}{H_e(s)} \quad (\text{for } A \Rightarrow \infty)$$

so

$$H_c(s) = 1 \quad (\text{for large A})$$

where: $H_e(s)$ = Encoder transfer function
$H_d(s)$ = Decoder transfer function
A = Amplifier gain

Consider the case where S1 is in the encode position. The signal first passes through the high-level stage which provides up to 10 dB of boost for signals below 200 Hz and up to 12 dB of boost for signals above 400 Hz. There are three compressors in the side chain: the low-frequency fixed band (LF/FB), the high-frequency fixed band (HF/FB) and the high-frequency sliding band. (HF/SB). The HF/FB and HF/SB are combined to form an action-substitution compressor. The main path of the high level stage contains the anti-saturation filter which has the characteristic shown by its block. The side chain signal is amplified and summed with the main path signal to form the high-level stage output.

The low-level stage provides an additional 12 dB of boost for signals above 400 Hz. The side chain contains a fixed-band and sliding-band compressor in an action-substitution configuration. The main path includes the low-level antisaturation network which is a simple low-pass shelf with the characteristics shown. The side chain signal is amplified and summed with the main path signal to form the low-level stage and encoder output.

The encoder output is fed to the modulation control (MC) block where four control signals are created. MC1 is fed in opposition to the sliding-band control signal. MC2 is fed in opposition to the sliding-band overshoot suppression signal. MC3 and MC4 are fed in opposition to the steady-state and overshoot suppression control signals in the high-frequency and low-frequency fixed band stages respectively.

3.2 High Frequency Stage

There are two high-frequency stages in the S-type side chain: the high-level, high-frequency stage and the low-level, high-frequency stage. The low-level stage provides compression (or expansion) for signals roughly in the range of -60 dB to -30 dB. Signals above this range receive a fixed, frequency-dependent attenuation from this stage. The high-level stage acts as a compressor (or expander) for signals roughly between -30 dB and 0 dB. Signals outside this range receive a fixed gain or attenuation.

Except for the levels at which they operate, the stages are so similar that their operating principles can be described as if they were identical. In the following block diagram description, the stages will be treated as one except where noted. The high frequency stage block diagram is shown in figure 13 and has two main sub-stages, the fixed-band compressor and the sliding band compressor. Each sub-stage can be subdivided into an audio signal path and a control signal path.

3.2.1 The Audio Signal Path

The input to the stage first passes through a band-defining filter with corner frequencies at 400 Hz and 12.8 kHz. The filter slopes are 6 dB per octave except in the LLS where the low-pass filter has a 12 dB/octave slope. The high-pass filter defines the operating frequency band for the stage, while the low-pass filters keep the operation within the audio band and provide a spectral skewing effect for high audio frequencies as well as desensitizing the control circuits to super-audio frequencies.

The filter output is fed to the input of the fixed and sliding band compressors which are connected in an action-substitution configuration. The output of the sliding band therefore represents the stage output. The high-frequency fixed band stage (HF/FB) acts as a variable attenuator whose attenuation increases with increasing signal level. The high-frequency sliding band (HF/SB) acts as a variable single-pole, high-pass shelf whose corner frequency rises with increasing control signal. The stopband attenuation of the sliding band falls at 6 dB per octave down to 200 Hz. The fixed band output is fed to a low-pass shelf with corner frequencies of 3.2 kHz and 12.8 kHz. This filter desensitizes the fixed band compressor to very high audio frequencies which allows the compressor to provide mid-frequency signal boost even in the presence of relatively high-level, high frequencies.

3.2.2 The Fixed-Band Control Signal Path

As previously mentioned, the amount of attenuation in the fixed band and the amount of "slide" of the sliding band increases as the magnitude of the control signal increases in each band.

Given this relationship, it is interesting to note that if the principle of least treatment is applied to each compressor circuit, it results in a desire to keep the control signal at the minimum possible value at all times. It is useful to analyze the control paths from this point of view and may help to further explain some of the circuit functions.

The control path signal for the HF/FB is taken from the HF/FB output. It is first high-pass filtered to desensitize the control path to out-of-band signals. The signal is then split into two paths, the main rectifier path and the passband rectifier path. The main rectifier path full-wave rectifies the signal and subtracts the results from MC3 (also full-wave-rectified). After subtraction, the signal is fed to one input of the maximum selector circuit, whose output is the greater of its two inputs. The passband rectifier path is similar to the main rectifier path except it is high-pass filtered and is not subtracted from MC3. The passband rectifier is required for complex signal conditions which would result in a large MC3 signal. Under these conditions, it is necessary to create another control path signal which can still provide the appropriate attenuation for signals within the band. The 800 Hz high-pass filter at the passband rectifier input provides the appropriate frequency weighting to insure that only signals substantially above the high frequency stage corner frequency of 400 Hz are used to determine the correct attenuation.

The output of the maximum selector is passed through two integrator stages with time constants of 8 ms and 160 ms respectively. The dual integrator provides a well-smoothed control signal which in turn keeps modulation distortion at very low levels.

After integration, the signal is passed to the non-linear stage which determines the control voltage vs. attenuation characteristics. In the fixed bands an exponential law was chosen of the form:

$$V_{out} = \exp\left(\frac{V_{in} - V_{os}}{V_t}\right)$$

where: V_{os} = scaling constant

$$V_t = \frac{kT}{q} = 26\text{mV (at } 300°\text{K)}$$

This law provides three advantages in the design of a fixed-band compressor.

a. It has a slow, smooth onset of compression. This is valuable in controlling compandor errors due to level mismatching.

b. The compression ratio rises with input level. This allows for a well-defined finishing point and allows the threshold to be as high as possible.

c. The term Vos allows the shape of the curve to be changed to provide excellent staggered stage combining.

3.2.3 Transient Signal Handling

The previous paragraphs describe the operation of the circuit under steady-state or slowly changing conditions. When the input signal changes level rapidly by 10 dB or more, there is a lag in the response of the integrators and the overshoot suppression (O/S) circuit begins to act. This circuit operates very much like a diode which can directly charge the final integrating capacitor. This diode will conduct whenever the peak voltage at its anode rises sufficiently above the integrated average on the second integrating capacitor. This threshold behavior has two advantages:

1. It prevents the O/S from operating on slowly changing signals or signals which exhibit small, rapid changes that would not lead to much overshoot.

2. Due to the fact that its conduction is based on the voltage across it, the circuit will respond more aggressively to high-level transients than low-level transients.

MC3 is fed in opposition to the O/S signal so that the resultant O/S will track the behavior of the main rectifier path.

For large signals which have a rapid decay, a condition could exist where the attenuation of the band remains excessively high for tens of milliseconds after the cessation of the signal. To prevent this, a fast recovery diode is connected to discharge the second integrating capacitor to the faster decaying voltage of the first integrating capacitor. Again, the threshold behavior of the diode causes it to conduct only during very large transients.

3.2.4 The Sliding Band Control Path

The HF/SB control is taken from the high frequency stage output. Since the stage output is the combined fixed and sliding band output, a portion of the HF/FB output signal is subtracted from the control path input. This subtraction raises the sliding band threshold, as long as the HF/FB output is not greatly attenuated, which is helpful in controlling the circuit's sensitivity to high frequency noise. This causes the HF/SB control signal to be proportional to the signal across the HF/SB variable resistor rather than the stage output.

After subtraction, the signal passes through a single-pole high-pass filter. This filter basically determines how the HF/SB corner frequency varies with input frequency. The corner frequency of this filter is a trade off between high frequency attenuation vs. excessive sensitivity to super-audio signals.

The signal then passes through a full wave rectifier, is opposed by MC1, and then on to two stages of integration. A single control path suffices in the sliding band because, unlike the HF/FB, there is no fixed, definable passband. Also, there is a fundamental difference in the way in which SB modulation control operates vs. FB modulation control.

The non-linear control element in the HF/SB is a fixed power law described by:

$$V_{out} = KV_{in}^{n}$$

where: n = power law
k = scaling constant

This is preferable for a sliding band stage because the compression ratio is constant with frequency.

There are two O/S signals in the HF/SB, the primary O/S and the secondary O/S. The primary O/S signal is derived from the main rectifier output which is opposed by a smoothed version of MC1 called MC2. It is necessary to use a smoothed modulation control signal in order to provide a reliable bucking effect. The secondary O/S signal is derived from the HF/FB O/S output. This signal is used to improve the O/S effect for signals in the 200 - 800 Hz region.

3.3 Low Frequency Stage

The low frequency stage block diagram (see figure 14) is very similar to the high frequency fixed band (HF/FB). The audio path is simply comprised of a single pole, 200 Hz, low-pass filter followed by a variable attenuator. As in the HF/FB, the attenuator is such that its attenuation increases as the control signal increases.

The control path input is again taken from the compressor stage output. The signal is first passed through a low-pass filter to further desensitize it to high frequencies. The signal is then split into main and passband rectifier stages. MC4 is used to oppose the main rectifier signal. The control signal is again smoothed by a double integrator with somewhat longer time constants than the HF/FB stage. An exponential control law is used and provides the advantages described for the HF/FB. The O/S circuits work in a manner very similar to those in the HF/FB and require no further explanation.

3.4 Modulation Control

The modulation control circuits, shown in figure 15, are fed from the encoder output. The filter blocks shown were empirically found to produce the required effect. The rectifiers are full wave rectifiers. The circuit outputs are fed to the destinations noted on the right hand side of the diagram.

4. Processor Performance

4.1 Encoder Response

There are a number of ways to characterize the performance of a compressor. The simplest way is to measure its output response when given a constant level, swept frequency sine wave input. It is useful to produce a series of these response sweeps over a range of input levels. This type of measurement reveals how the processor responds to signals of a given frequency and level. The results are useful for estimating the amount of boost or attenuation steady-state signals will receive during encoding. Compression ratios can be observed by looking at the curves as well. Vertical spacing of adjacent sweeps is directly related to compression ratio. Since the input signal was varied in 10 dB steps, a spacing of 10 dB in the output curves indicates no compression while curves spaced 5 dB apart indicate a compression ratio of 2:1. See figure 16.

4.2 Compression Ratio Versus Level

A better way of examining the compression ratio of the encoder is to apply a fixed frequency, swept level sine wave as the input. This type of response was illustrated in figure 3. Compression ratio is defined as:

$$\text{Compression Ratio} = \frac{\Delta V_{in}(dB)}{\Delta V_{out}\ (dB)}$$

where: $\Delta V_{in}(dB)$ = A specified change in input level in decibels
$\Delta V_{out}(dB)$ = The change in output level in dB due to ΔVin (dB)

This can be easily calculated from the results of a fixed-frequency level sweep. The results can be plotted versus the input or output level.

It is important to recognize that proper operation of a compandor system is dependent upon the linearity of the recording medium. An advantage of the bilinear compression characteristic is that it only requires that the medium be linear over the signal range that is being actively processed. Outside this active range, no compression is taking place so non-linearity of the medium does not lead to further errors. One goal of the design of the S-type system was to control this active range to be below the limits of linearity of the recording medium and above the noise floor. In order to examine the processor's performance in this respect, it is most useful to plot the compression ratio versus the output level since this is the level which will be recorded.

Figure 17 shows this type of plot for an S-type encoder at several frequencies. Compression ratios below 1.2:1 can be considered essentially linear (no compression). The results reveal that compression starts in the -40 dB to -30 dB region and ends in the -10 dB to 0 dB region. The highest compression ratios are confined to a small output level region where all consumer formats are quite linear. Note also that the 1 kHz compression ratio curve becomes linear at a higher level than the lower or higher frequency compression ratio curves. This corresponds well to the headroom characteristics of consumer formats (see figure 20). The 0 dB processor output level is defined to correspond to a reference recorded fluxivity of a given format.

4.3 Noise Reduction Effect Versus Frequency

At signal levels below threshold the compressor provides a fixed, frequency-dependent boost and the decoder provides a complementary attenuation. By examining the low-level boost of the encoder, a noise-reduction effect versus frequency plot can be created. This is shown in figure 18.

4.4 Response to Non-Dominant Signals

The previous performance graphs show the response of the processor to dominant signals at the frequency of the dominant signal. This does not indicate the response of the system to low-level, non-dominant signals. This can be examined by summing two signals and applying them to the processor, one which is above threshold and is fixed in level and frequency and one which is below threshold and fixed in level but swept in frequency. By examining the output versus frequency and rejecting the dominant signal, the non-dominant signal boost characteristics are revealed. In figure 19 the results are displayed in a manner which shows the relative gain received by the sub-threshold signal. A family of curves illustrates how the characteristics change as the fixed-frequency signal level changes. Each curve is labelled with the level of the fixed frequency signal. Note that while encoder boost is lost near the dominant signal, it is retained at other frequencies. This illustrates the application of the least treatment principle.

4.5 Computer Model

Because the S-type process is based on simple transfer functions which are easy to implement with conventional analog technology, it is possible to define the steady-state transfer function and predict its response. A mathematical model for an S-type encoder has been constructed and is being used as a reference to predict circuit performance. The model was written in C language and can simulate a theoretically perfect processor. Both dominant and non-dominant signal responses can be examined. Discrete and integrated circuit performance has been found to match the model perfectly to the extent that the circuitry performs according to its design equations. We have chosen to define the steady-state response of the S-type process as the result obtained from this model. All graphs of processor performance in this section were generated using this model.

5. Compandor Performance with the Cassette Recording Format

While the S-type process was not designed solely for cassette recording, it is useful to examine its effect on this common format. In this way, some of the improvements that have been previously discussed can be quantified and their benefits illustrated. The cassette deck used for the following tests was a good quality, 3 head deck equipped with Dolby HX. The tape used was IEC type IV (metal) and the deck was aligned to have a flat frequency response (±1.0 dB) with this tape. Playback sensitivity was adjusted so that 200nWb/m fluxivity played back at processor reference level (0 dB). Record sensitivity was adjusted so that encoder reference level was recorded at 200nWb/m.

5.1 Headroom Improvement from Anti-Saturation

The headroom of a recording format is often defined as a level which produces a certain amount of 3rd harmonic or total harmonic distortion. This measurement is really only useful for frequencies up to about 6 kHz since above this frequency the 3rd harmonic will not be reproduced. Another method is to measure the input level at which the output level falls below the input level by a certain amount. This measurement can be easily made at all frequencies. Since the nature of magnetic recording is such that saturation at any frequency can affect the linearity at all frequencies, one must be aware how this headroom changes with frequency. In figure 20 the 1 dB loss or 1 dB "squash point" is plotted versus frequency with and without S-type. The resultant headroom versus frequency curve conforms well to the expected peak level versus frequency for most popular music [4].

5.2 Distortion Reduction

The S-type process reduces distortion in two ways - it decreases the level during recording and it suppresses harmonics during playback. Figure 21 shows this improvement for a 0 dB (200nWb/m), 200 Hz signal. While it is arguable that the graph illustrates an ideal case since there are no other signals present; it can also be argued that a more complex signal would mask the harmonics so that distortion reduction is not necessary in that case.

5.3 Dynamic Range Improvement

Figure 22 is a composite of two different measurements. The headroom is measured as in figure 20. The noise floor is measured with a swept frequency, 1/3rd octave filter with and without S-type for type IV (metal) tape and referenced to 200nWb/m. The graph is intended to illustrate the relative dynamic range improvement versus frequency and cannot be used to determine the actual signal dynamic range. For instance, the headroom has been improved by 6 dB at 10 kHz and the noise floor has been lowered by 21 dB for a 27 dB dynamic range improvement at that frequency.

5.4 Dynamic Range Versus Hearing Threshold

In order to quantify how well the resulting recording system will perform, the dynamic range of the recording system can be compared to the dynamic range of the human auditory system. In order to do this, one must be able to quantify the ability of a listener to detect noise in the absence of a music program. This measurement is known as the threshold of hearing and can be expressed as a sound pressure level (SPL) versus frequency. By choosing a certain maximum peak SPL for playback of a recording, the dynamic range of the recording system can be superimposed over the hearing threshold curve. The noise floor of the recorder must be further processed to account for the difference between the bandwidth of the measurement filter and the apparent detection bandwidth of the ear versus frequency (critical bandwidth) [5].

The results of this process are shown in figure 23. A peak playback level of 100 dB SPL was chosen as a typical playback level. Higher or lower playback levels would move the recorder dynamic range limits up or down respectively while the hearing threshold curve would remain fixed. Noise which is below the threshold of hearing on this graph should be inaudible.

These curves, of course, say nothing about the audibility of noise in the presence of signals. Under these conditions, the processor takes advantage of the psycho-acoustic phenomenon of masking whereby the presence of noise in a certain frequency band is masked (made inaudible) by audio signals in the same frequency range. This phenomenon is fairly well understood for the case of a single sinewave masking bands of noise. In this case, a rough generalization can be made that masking extends from 1/2 to 1 octave below the tone to approximately 1 to 2 octaves above. The masking effect of complex tones is not well understood making the audibility of noise in the presence of program material difficult to predict. It is our experience, however, that the noise floor of the resultant recording, if audible, remains at an apparently constant level regardless of program material. The lack of noise-modulation was one of the design goals of the S-type processor and has been mentioned as an expected benefit of the principle of least treatment.

6. Conclusion

A new signal processor that extends the useful dynamic range of consumer recording formats has been described. The design goal of least treatment was described and results were presented to show how well the new process achieves the goal. The processor reduces many of the undesirable side-effects of recording to inaudible levels. The processor is also tolerant of level and frequency response errors and is designed to minimize mistracking under those conditions. Furthermore, the processor is designed to minimize the audibility of mistracking or improper decoding should it occur, resulting in a rugged system capable of good results in a wide variety of recording and playback environments.

7. Acknowledgement

The author would like to thank the following contributors to the project: Mark Davis and Doug Mandell for their work formulating and implementing the computer model, Ken Gundry for many hours of useful instruction on the finer points of noise reduction, and Ray Dolby for his development of the principles upon which this process is based and for his encouragement and patience as I learned to apply them.

8. References

[1] R. M. Dolby, "The Spectral Recording Process," J. Audio Eng. Soc., Vol. 35, pp. 99-118 (1987, Mar.).

[2] R. M. Dolby, "An Audio Noise Reduction System," J. Audio Eng. Soc., Vol. 15, pp. 383-388 (1967 Oct.).

[3] R. M. Dolby, "A 20 dB Audio Noise Reduction System for Consumer Applications," J. Audio Eng. Soc., Vol. 31, pp. 98-113 (1983 March).

[4] R. A. Greiner and J. Eggers, "The Spectral Amplitude Distribution of Selected Compact Discs," J. Audio Eng. Soc., Vol. 37, pp. 246-275 (1989 April).

[5] L. D. Fielder, "Evaluation of the Audible Distortion and Noise Produced by Digital Audio Converters," J. Audio Eng. Soc., Vol. 35, pp. 517-535 (1987 July/August).

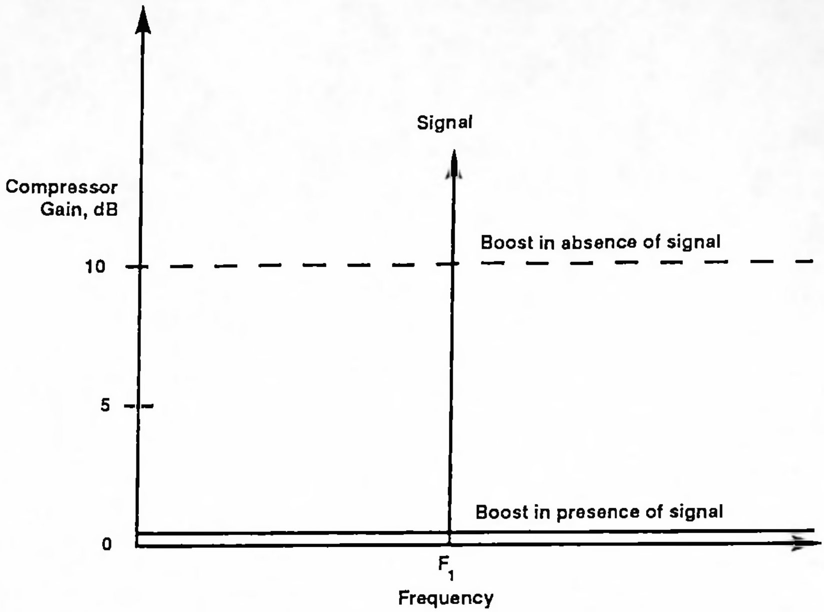


Fig. 1 Simple Compressor Action

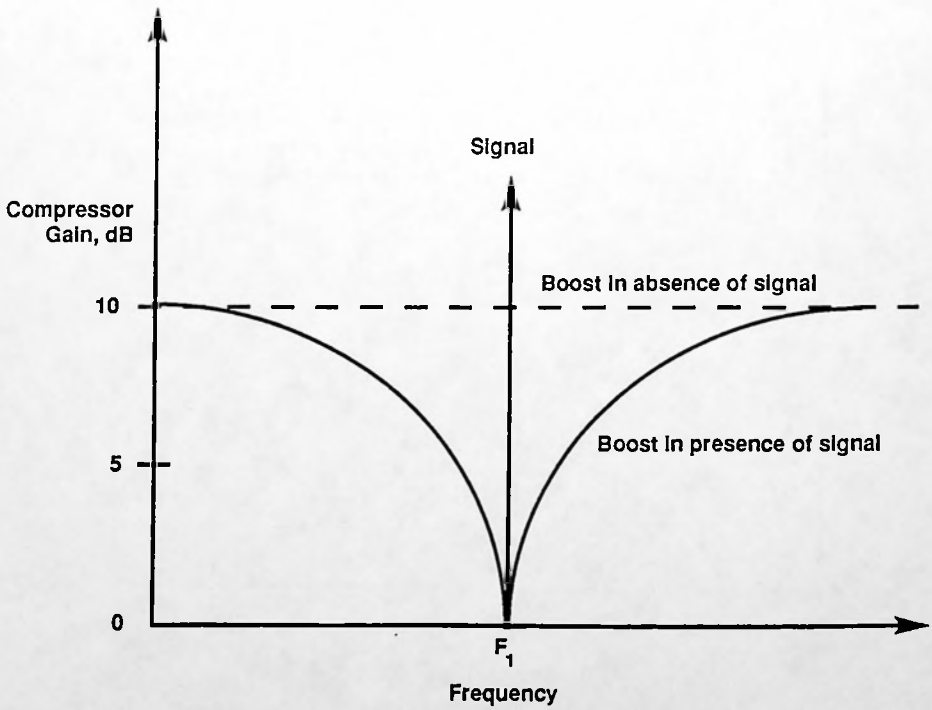
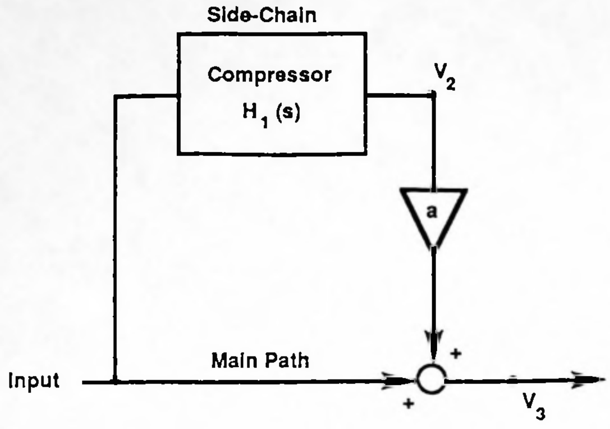


Fig. 2 Least Treatment Compressor Action



$$V_3(s) = Vin(s)\ [aH_1(s) + 1]$$

Fig. 3 Bilinear Compressor Block Diagram

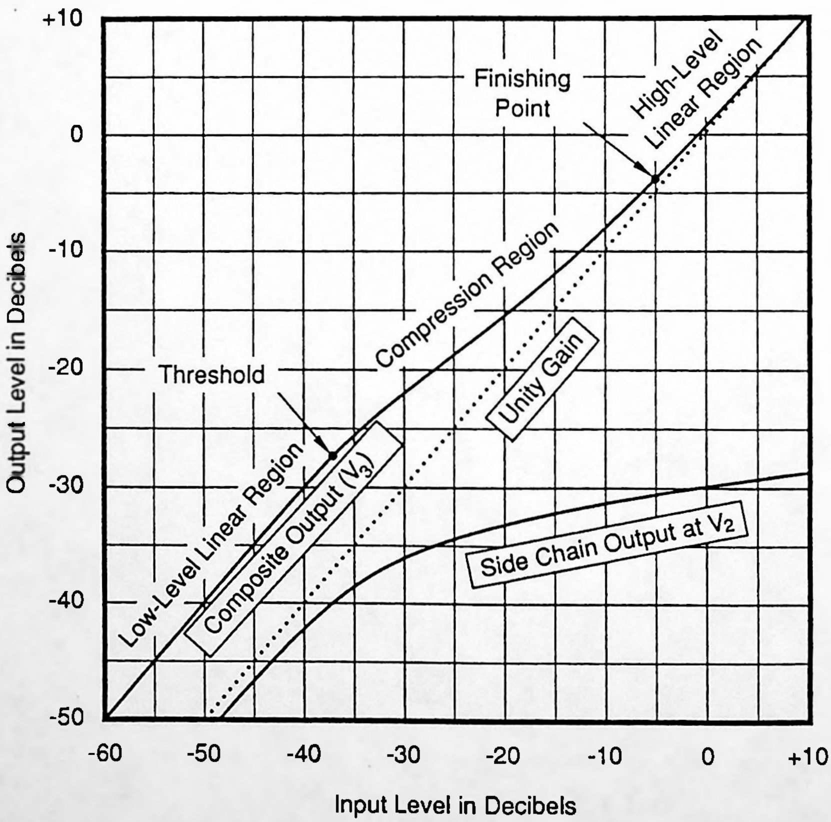


Fig. 4 Bilinear Compressor Characteristics

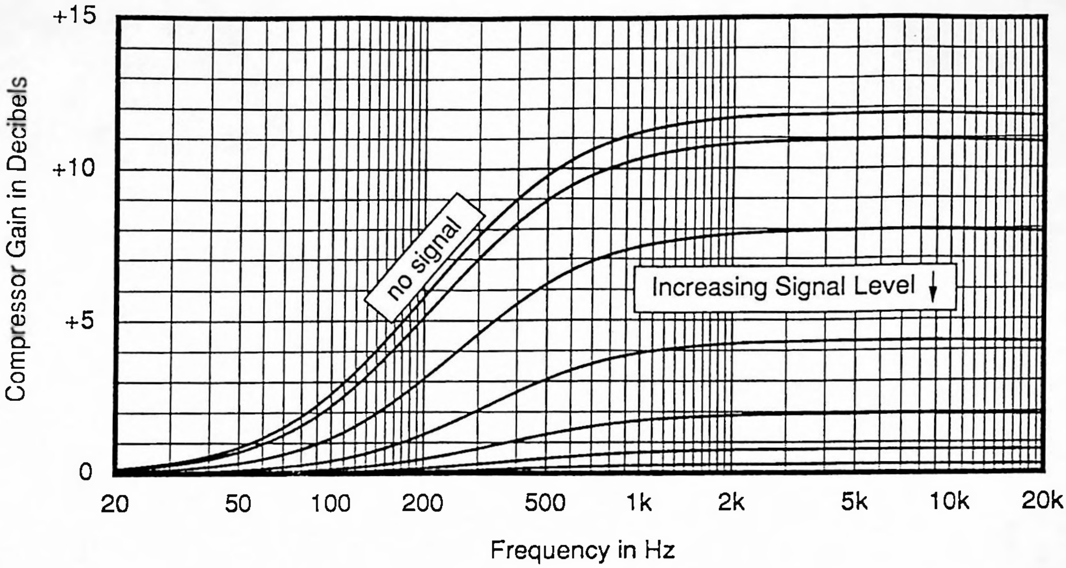


Fig. 5 Fixed Band Compressor Characteristics

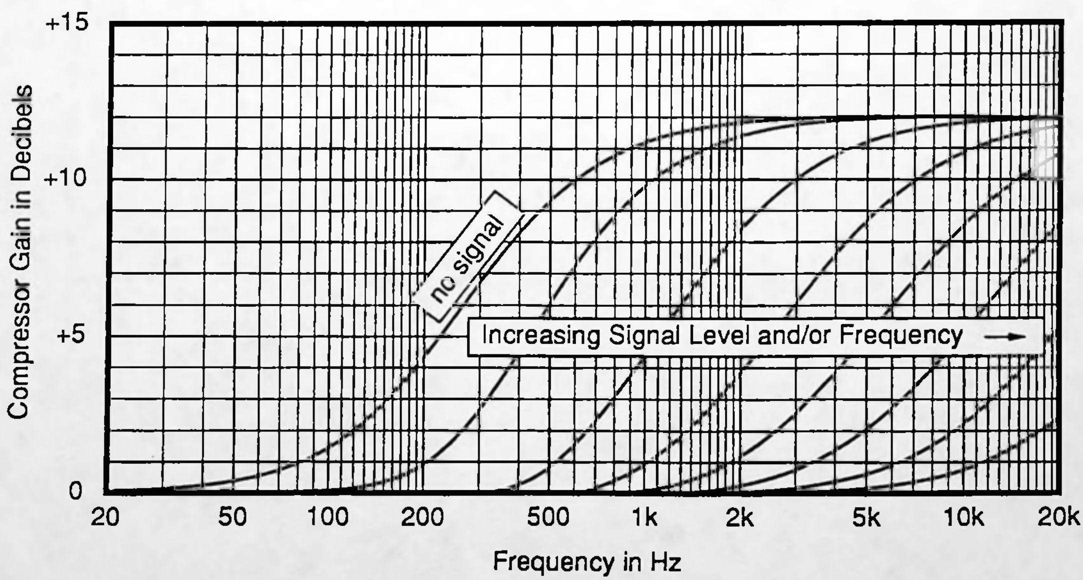


Fig. 6 High Freq. Sliding Band Compressor Characteristics

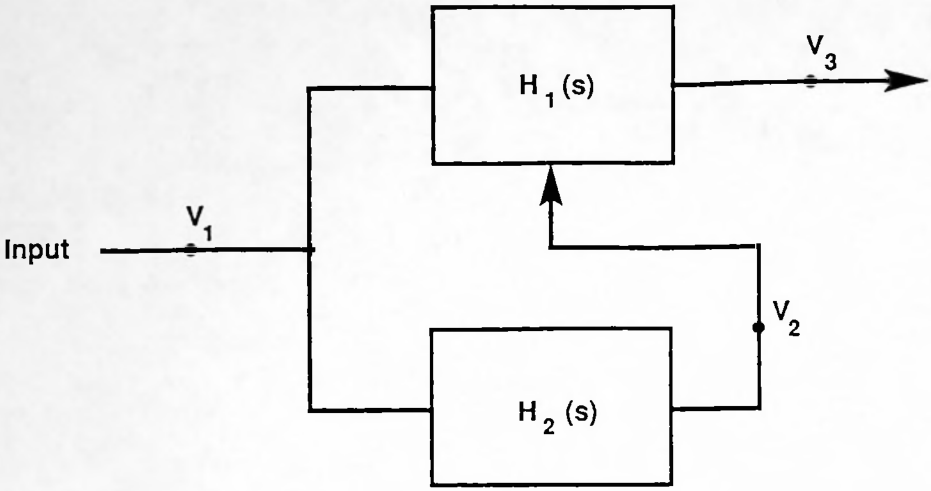


Fig. 7 The Action Substitution Configuration

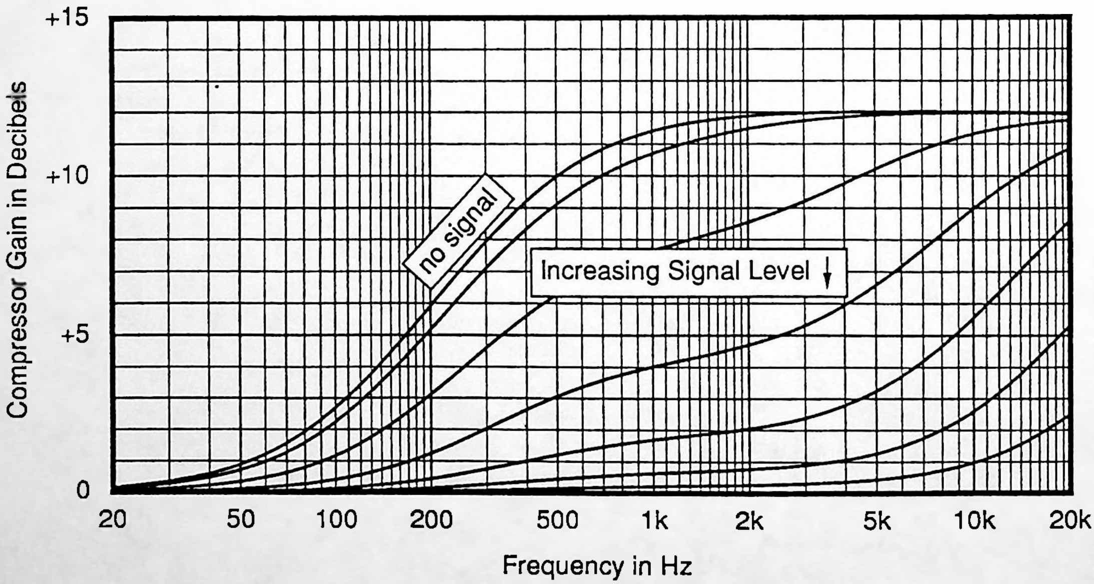


Fig. 8 Action Substitution Compressor Characteristics

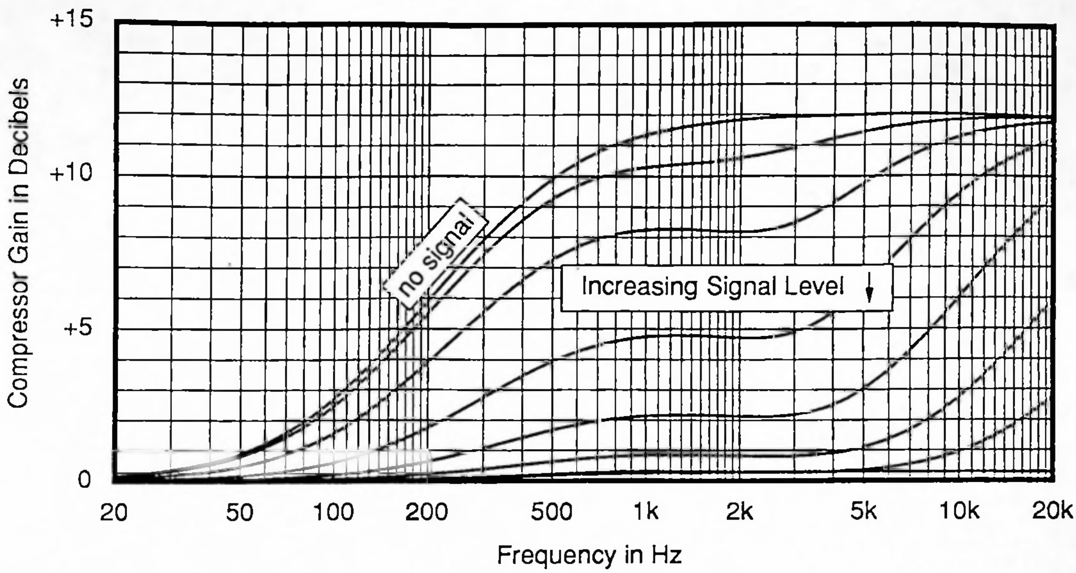


Fig. 9 S-type Action Substitution Compressor Characteristics

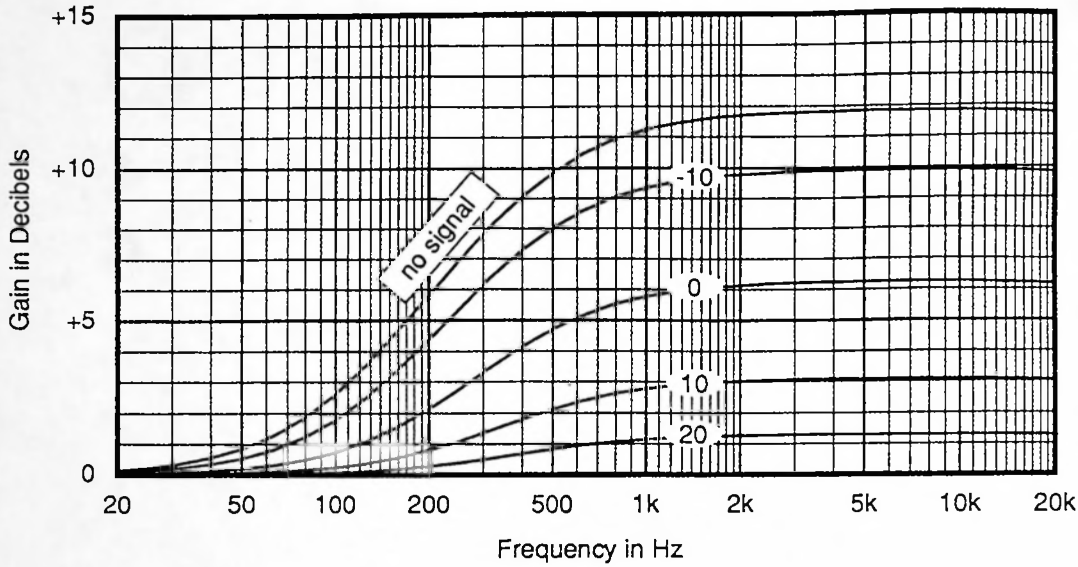


Fig. 10a Fixed Band Compressor Response to 100 Hz Tone

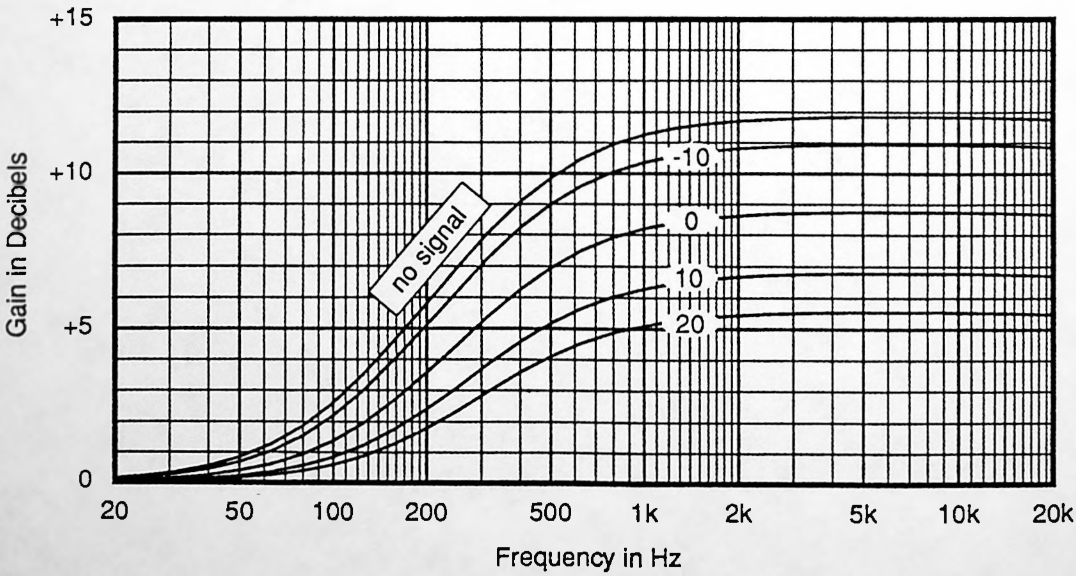


Fig. 10b Compressor Response to 100 Hz with Modulation Control

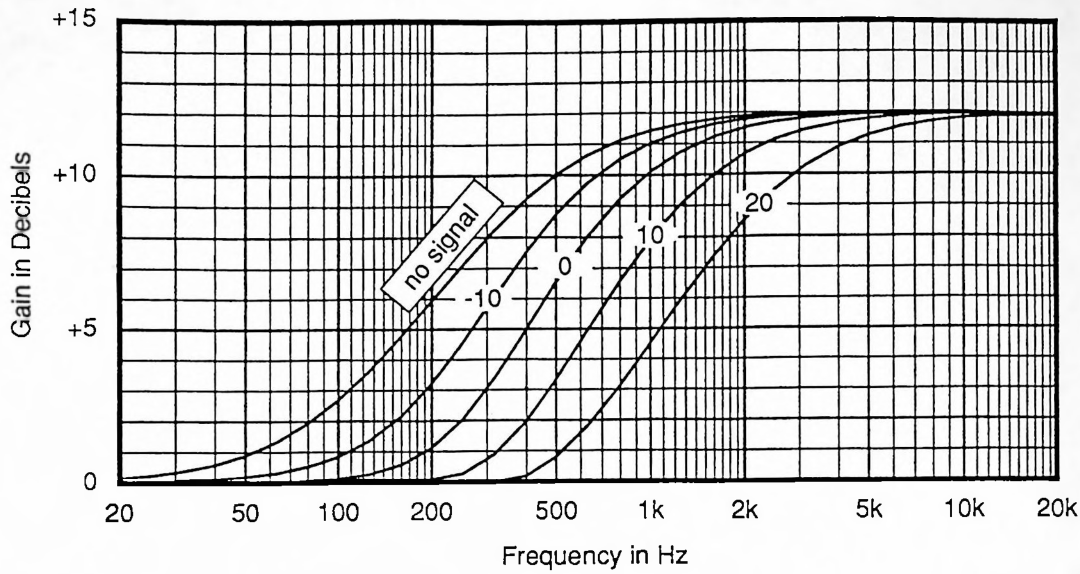


Fig. 11a Sliding Band Compressor Response to 100 Hz Tone

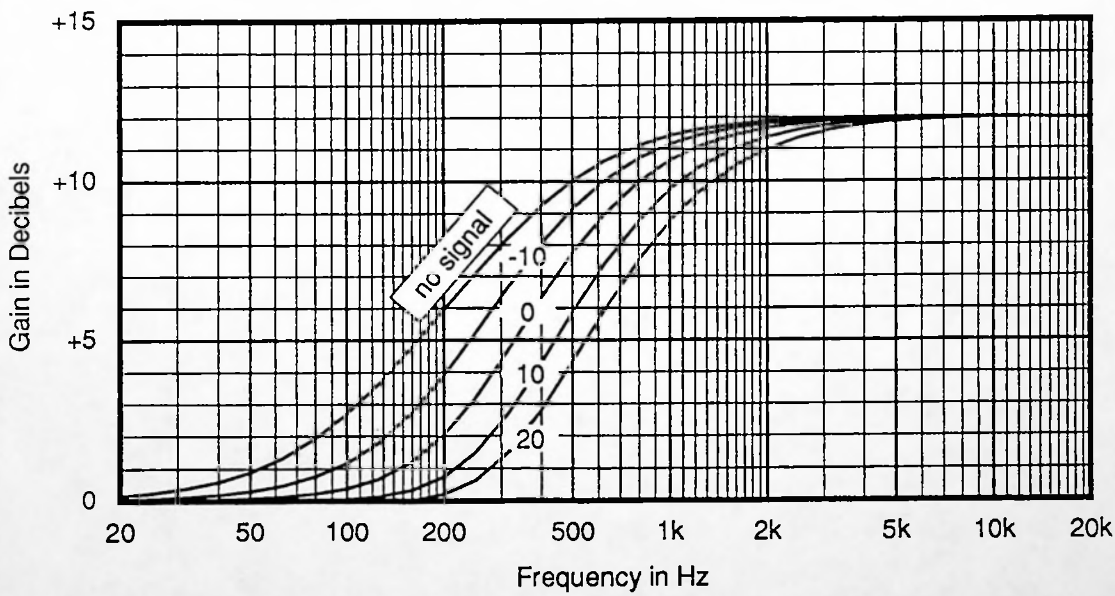


Fig. 11b Compressor Response to 100 Hz with Modulation Control

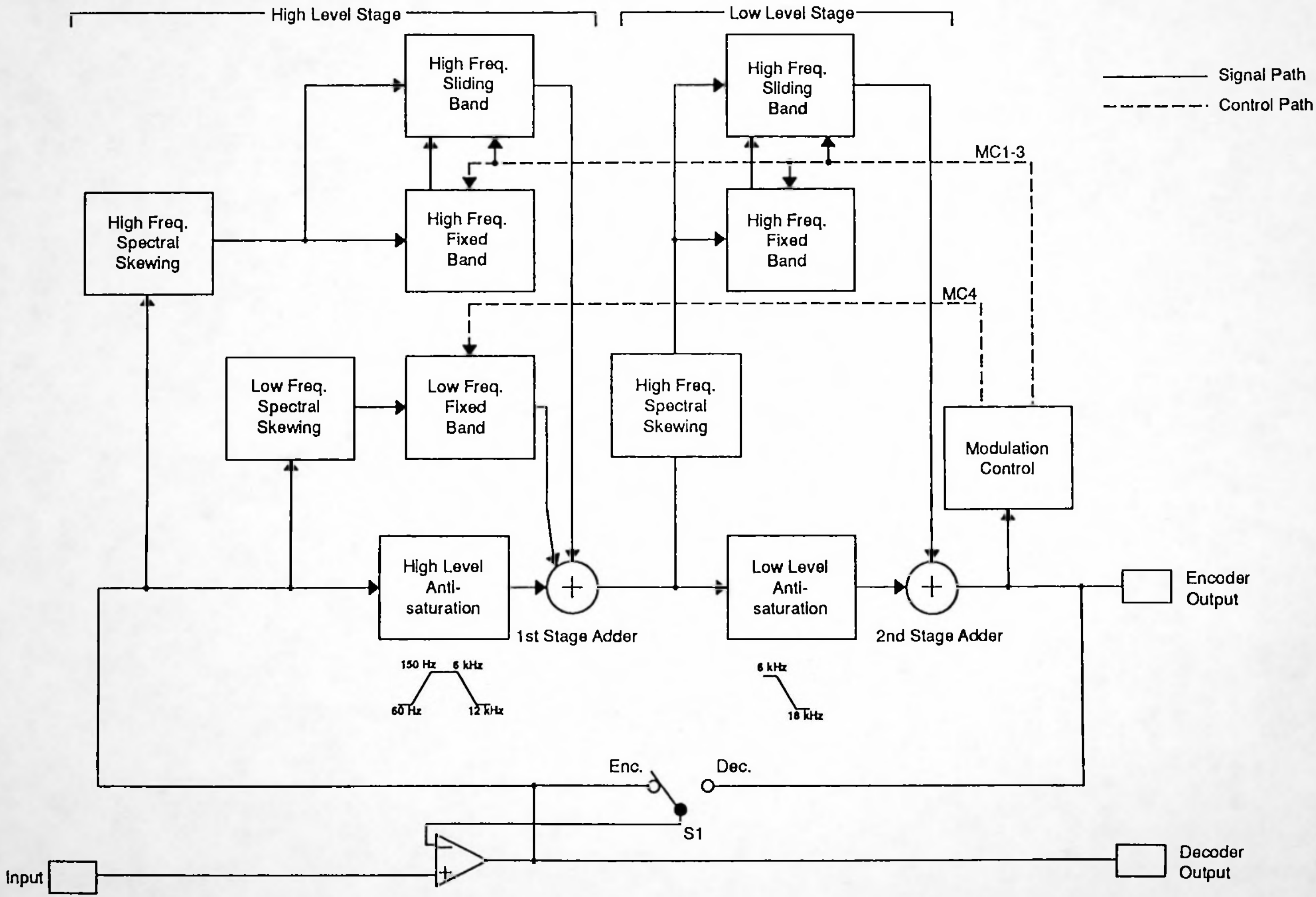


Fig. 12 S-type Processor Block Diagram

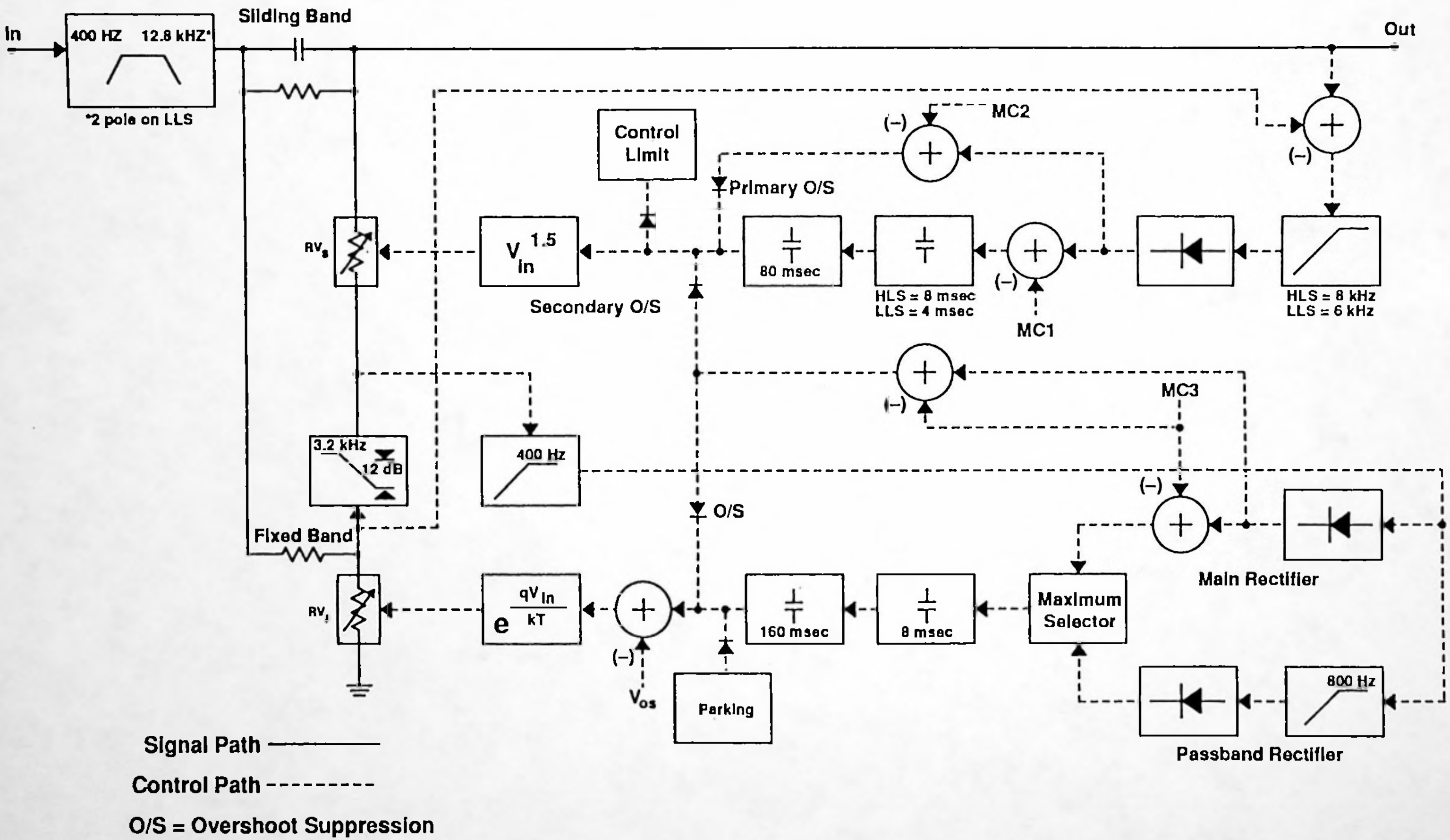


Fig. 13 High Freq. Stage Block Diagram

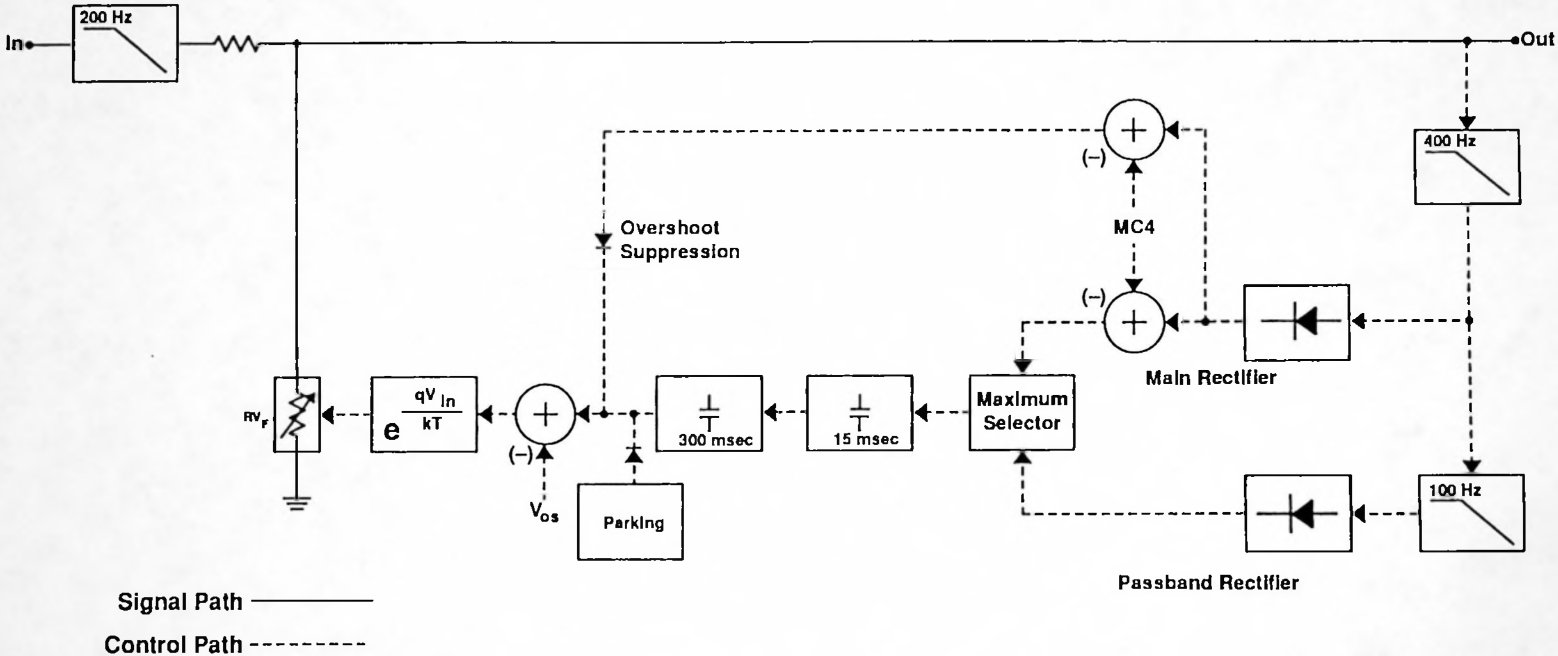


Fig. 14 Low Freq. Stage Block Diagram

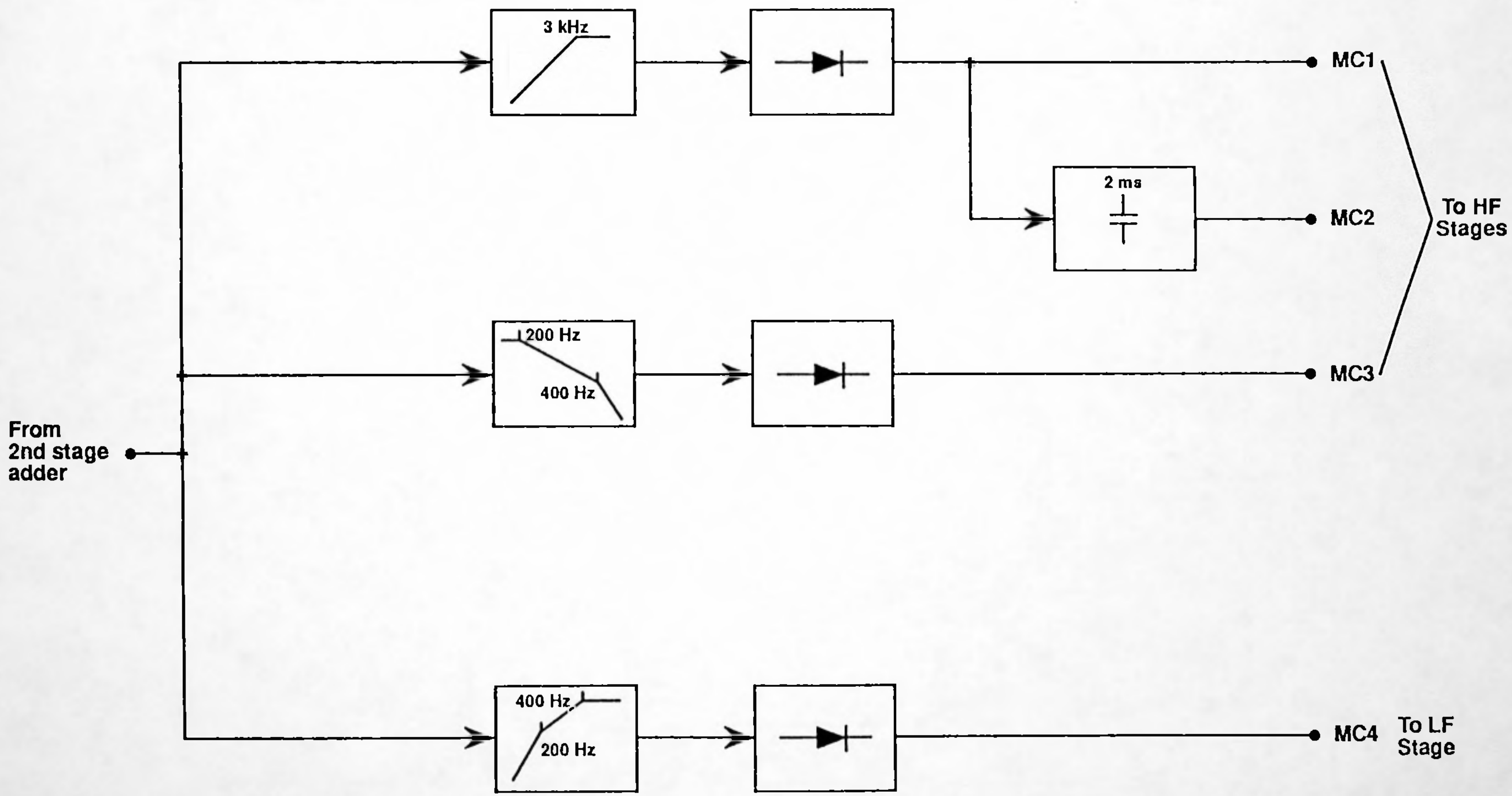


Fig. 15 Modulation Control Circuits Block Diagram

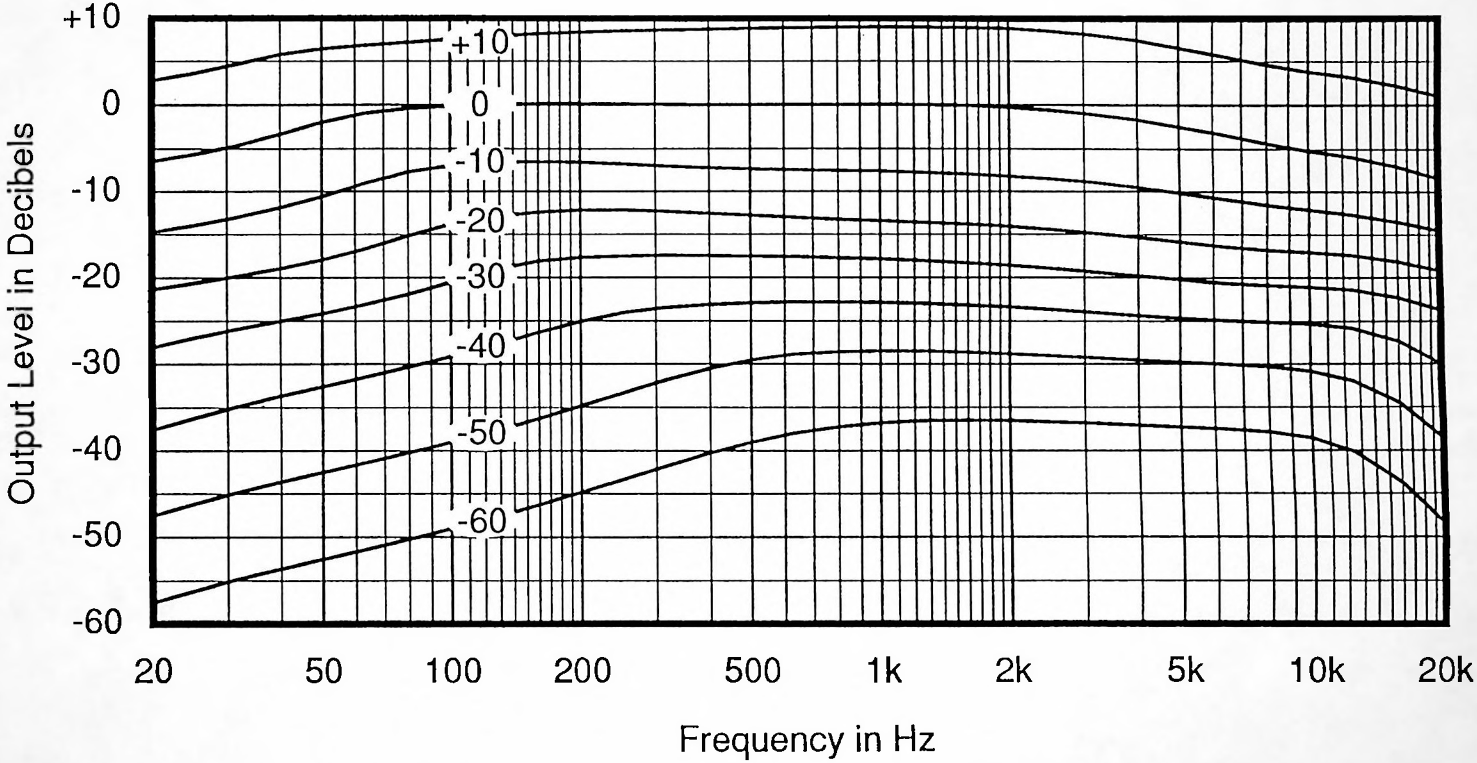


Fig. 16 S-Type Encoder Response versus Level and Frequency

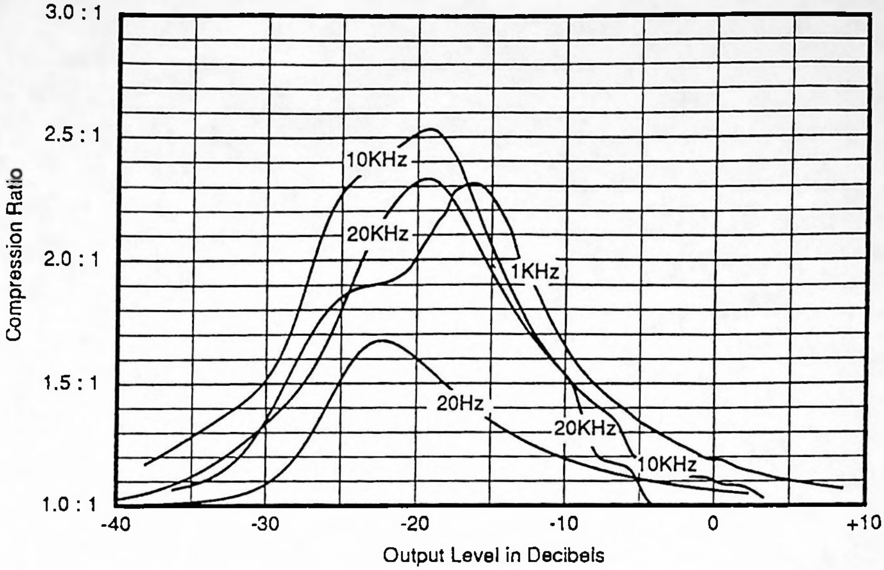


Fig. 17 Compression Ratio versus Output Level

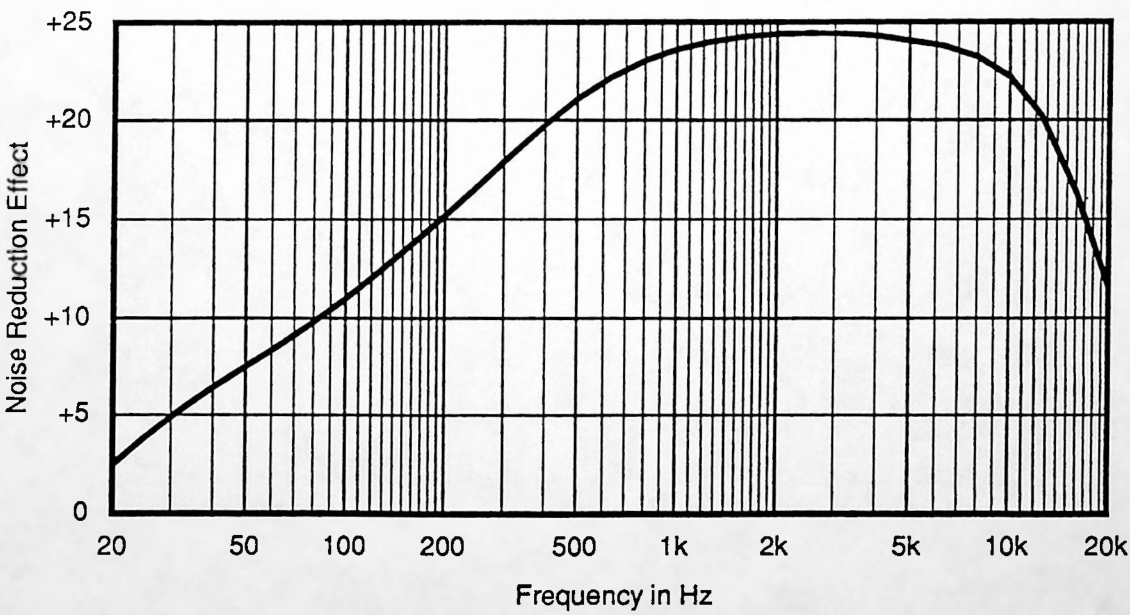


Fig.18 Maximum Noise Reduction Effect versus Frequency

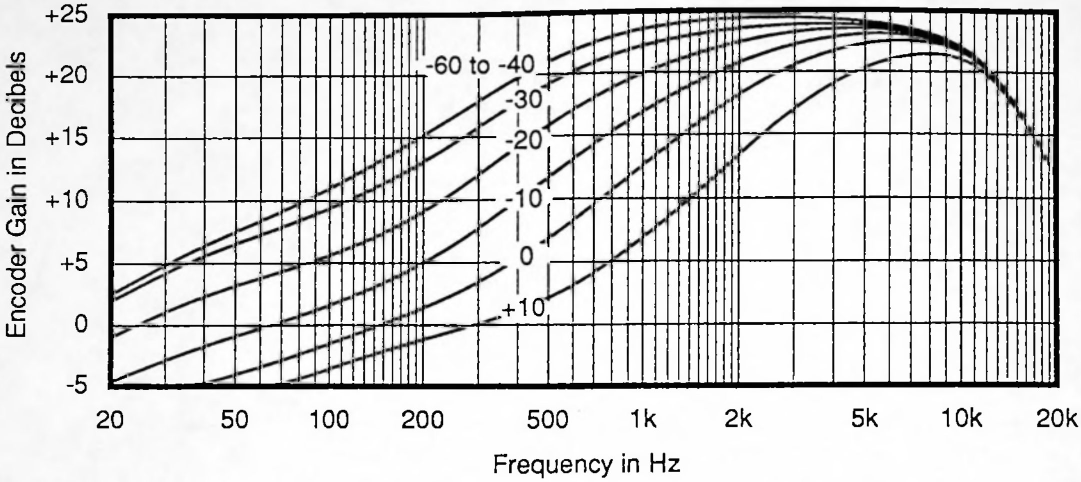


Fig. 19a Encoder Gain in the Presence of a 200 Hz Signal

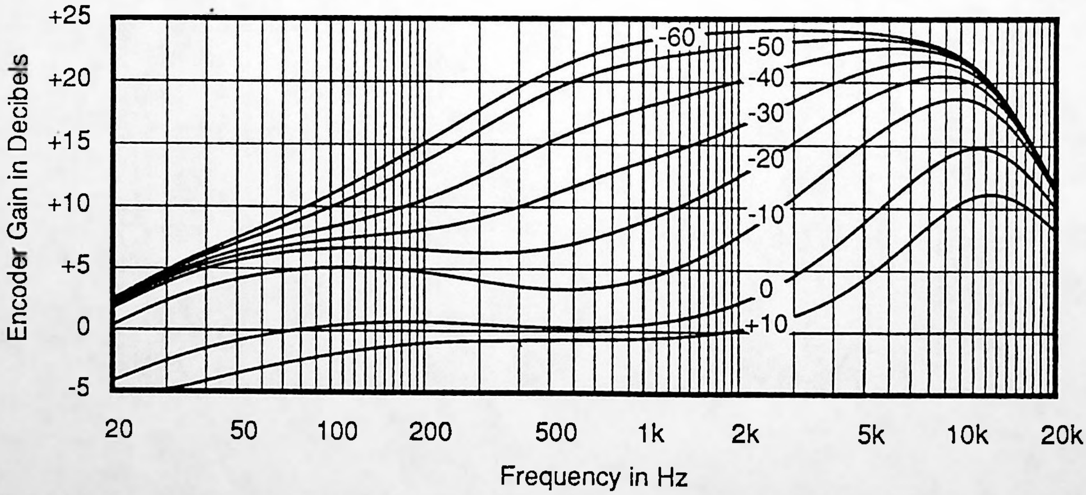


Fig.19b Encoder Gain in the Presence of an 800 Hz Signal

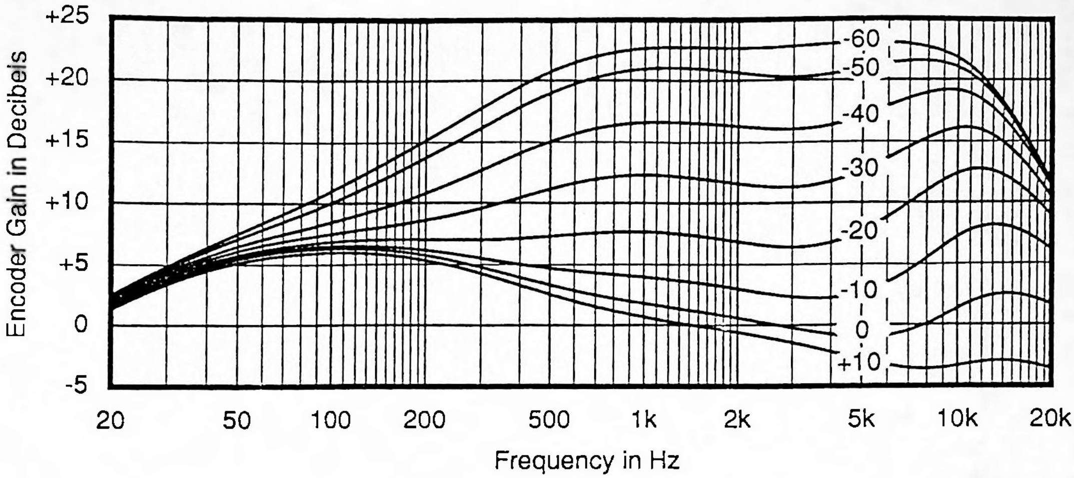


Fig. 19c Encoder Gain in the Presence of a 3 kHz Signal

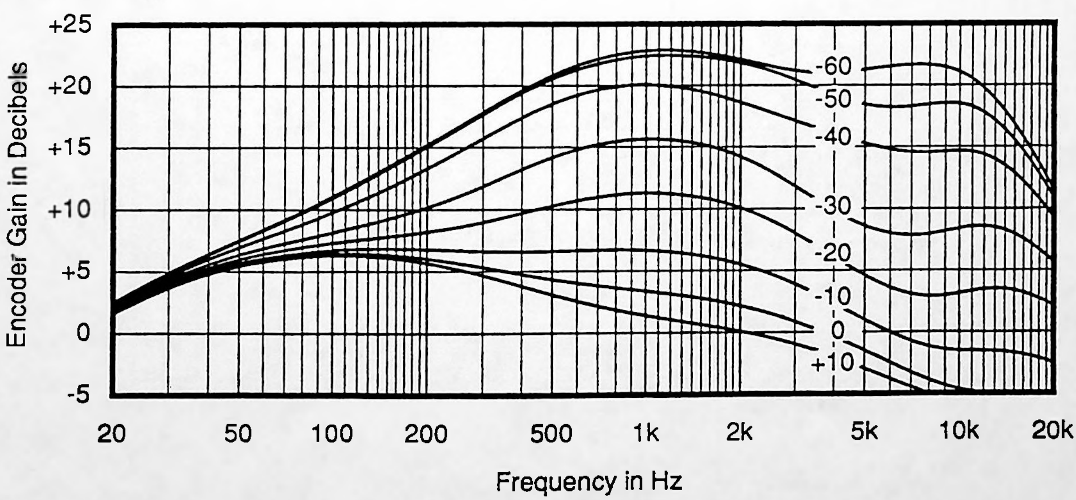


Fig. 19d Encoder Gain in the Presence of a 12 kHz Signal

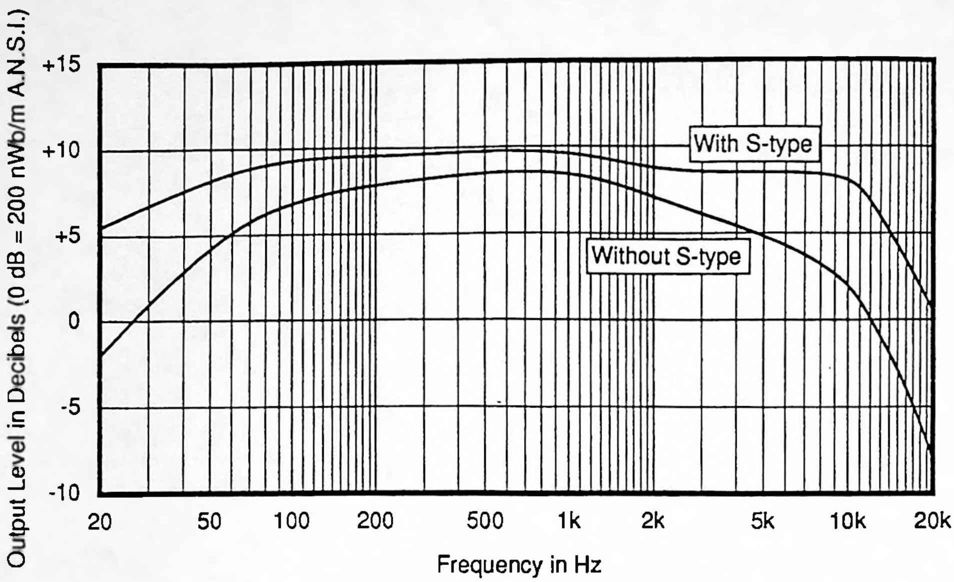


Fig. 20 Cassette Headroom versus Frequency for Type IV Tape

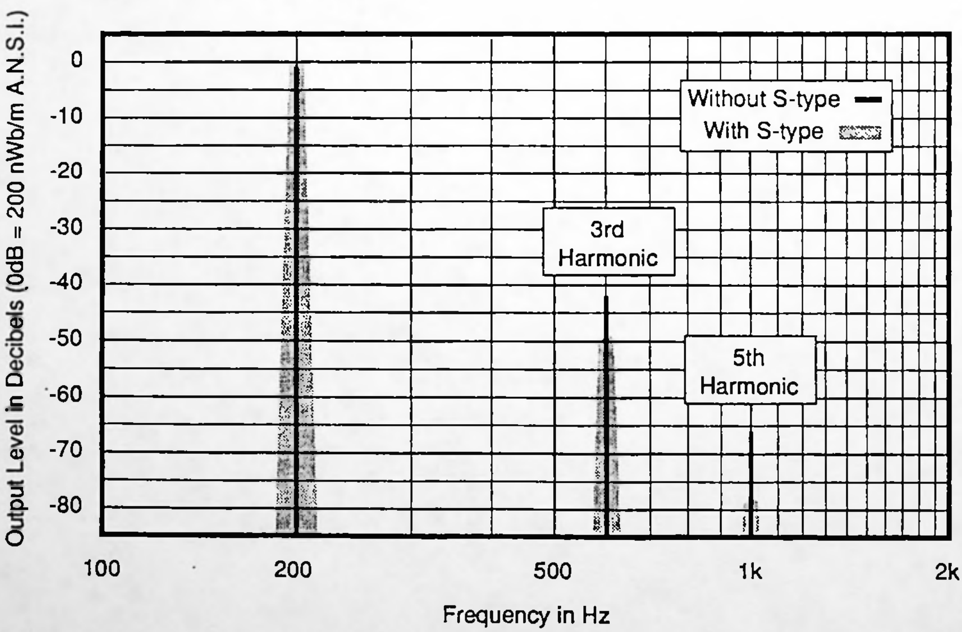


Fig. 21 Harmonic Distortion of 200 Hz Tone

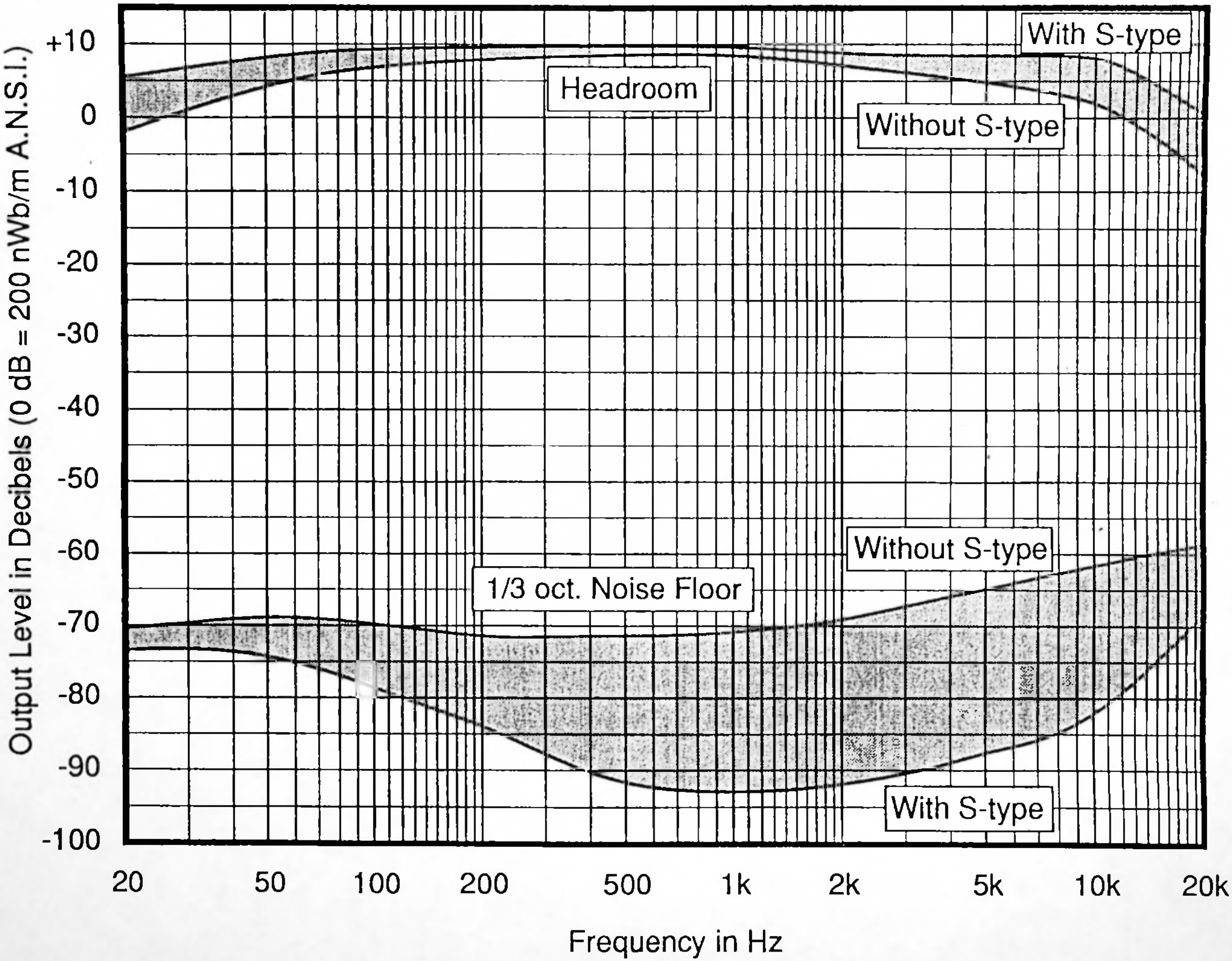


Fig. 22 Dynamic Range Improvement for Type IV Tape

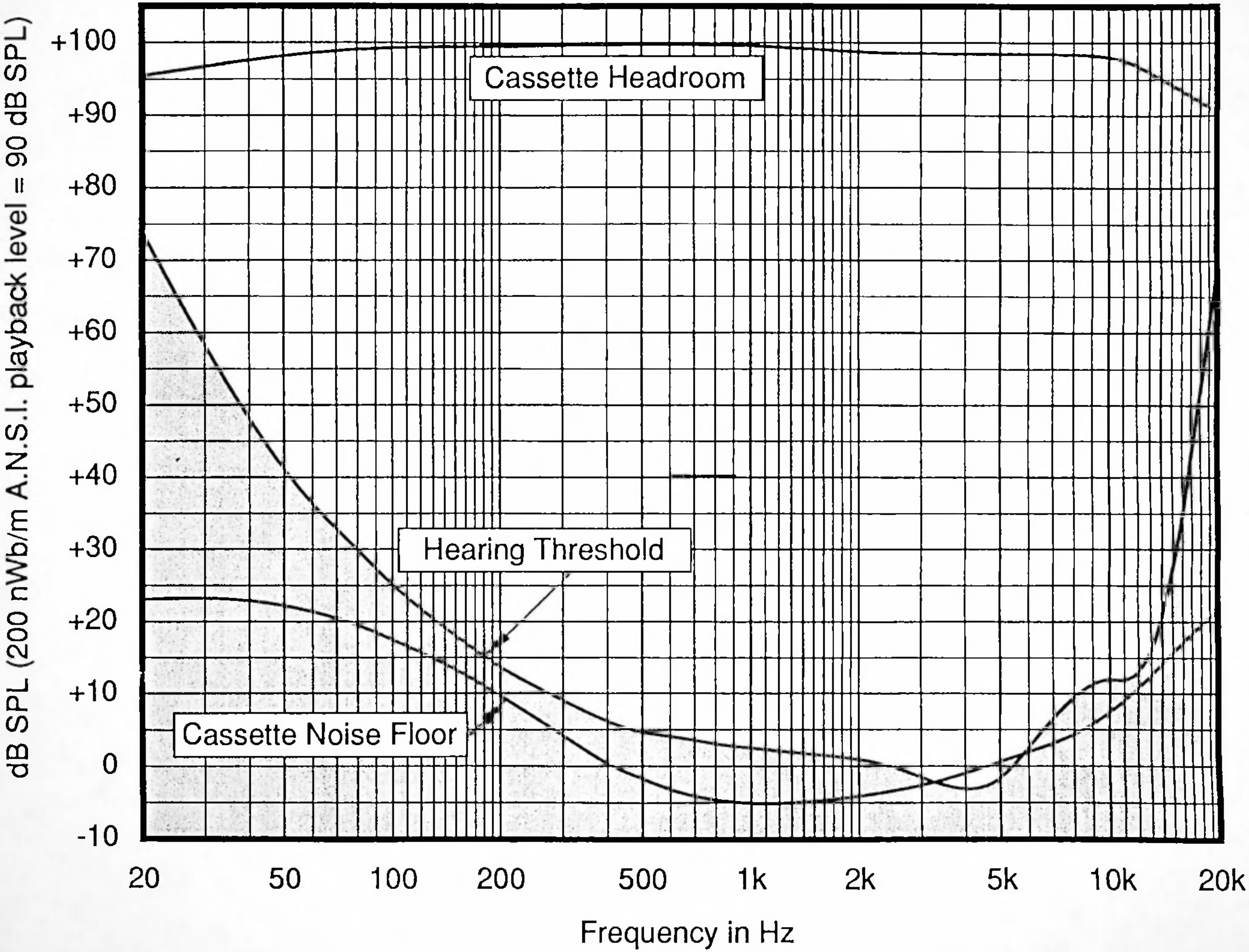


Fig. 23 Cassette Recorder Dynamic Range versus Threshold of Hearing

Dolby

DOLBY LABORATORIES

100 Potrero Avenue
San Francisco CA 94103-4813
Telephone 415-558-0200
Facsimile 415-863-1373
Telex 34409

346 Clapham Road
London SW9 9AP
Telephone 071-720-1111
Facsimile 071-720-4118
Telex 919109

A3MS402:
REV A