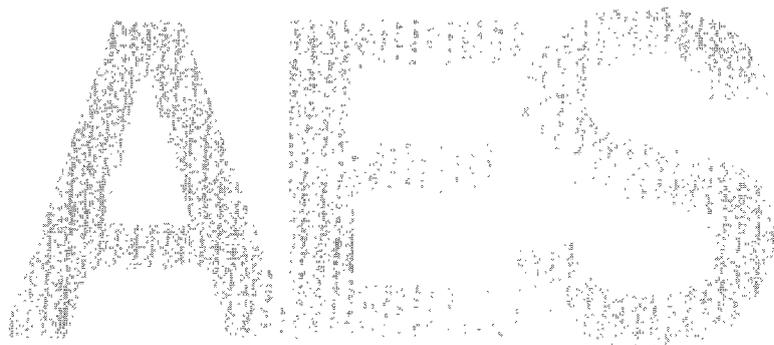


A 20 dB AUDIO NOISE REDUCTION SYSTEM  
FOR CONSUMER APPLICATIONS

1850 (J-6)

Ray Dolby  
Dolby Laboratories  
San Francisco and London

**Presented at  
the 70th Convention  
1981 October 30–November 2  
New York**



*This preprint has been reproduced from the author's advance manuscript, without editing, corrections or consideration by the Review Board. The AES takes no responsibility for the contents.*

*Additional preprints may be obtained by sending request and remittance to the Audio Engineering Society, 60 East 42nd Street, New York, New York 10165 USA.*

*All rights reserved. Reproduction of this preprint, or any portion thereof, is not permitted without direct permission from the Journal of the Audio Engineering Society.*

**AN AUDIO ENGINEERING SOCIETY PREPRINT**

A 20 dB AUDIO NOISE REDUCTION SYSTEM  
FOR CONSUMER APPLICATIONS

Ray Dolby

Dolby Laboratories, San Francisco and London

ABSTRACT

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 demands made on transmission channel uniformity are generally unchanged from those associated with the B-type system. However, 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 1967-68 and first applied to open-reel recording (K1H Models 40 and 41). However, by this time there was a general feeling that for widespread use a more convenient tape format was required. 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 (1 7/8 ips) 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 1/4 inch 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 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 which were comparable with the best discs. 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 discs and B-type encoded cassettes when produced and reproduced under the best conditions:

<u>Discs</u>	<u>Cassettes</u>
Mold-grain noise	Hiss
Hiss	High-frequency overload distortion
Ticks and pops	

The low-frequency mold grain noises (rumbling and rushing sounds) produced by discs are evidently unnoticed by most listeners; perhaps these noises are masked by the ambient noises of typical listening environments. Disc processing hiss is variable but usually not too obtrusive. The main audible defect of most discs 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 discs; the continuing presence of this hiss has evidently been the main factor in causing the cassette to fall just short of discs 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 discs. For economic reasons, most duplicators are reluctant to use tapes which 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 1979-80. 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 have become available, especially during the last two years. Many cassette deck manufacturers have been requesting 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 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 high-frequency response of the recorder.

## 1. STAGGERED ACTION DUAL-LEVEL FORMAT

In the development of the A-type noise reduction system in 1965-66 (4), it was 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 1967-69, 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 towards the retention of the dual-path, 10 dB, low-level format which had proved to be successful in the A 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 ratio 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 "bi-linear" characteristic because there are two portions of substantially constant gain. Such characteristics may be distinguished from other popular types of characteristics, namely:

- (a) A logarithmic or non-linear characteristic with either a fixed or changing slope and with no linear portion: the gain changes over the whole dynamic range.

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

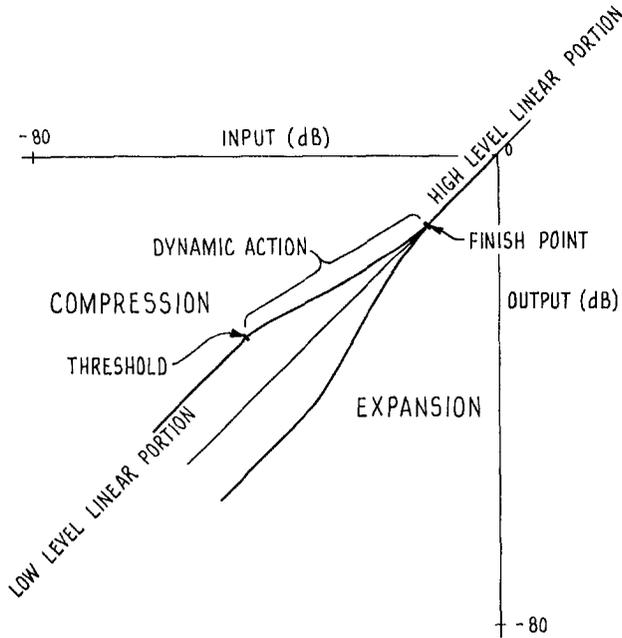


Fig. 1 Bi-linear compression and expansion characteristics.

An advantage of a bi-linear 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 non-linear 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, i.e., in the region of 40 dB below the nominal peak level. This leaves a high-level linear region of some 20 dB to deal with overshoots.

Note that bi-linear compressors and expanders determine the two end regions of constant gain by means of fixed, pre-set 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 multi-stage 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 (e.g., 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 bi-linear 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 bi-linear, with all of the attendant advantages. Thus the action staggering possibility of bi-linear compressors and expanders represents a further advantage of this type of device.

At any given frequency, the thresholds and dynamic regions of the compressors or expanders 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 bi-linear compressor feeds the low-level bi-linear compressor connected in series. During playback a pair of series connected bi-linear expanders receive the input from the signal channel and provide an overall noise reduction system output at the output of the high-level expander.

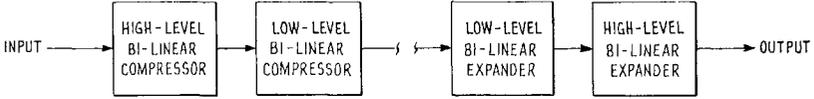


Fig. 2 Basic block diagram of staggered action bi-linear noise reduction system.

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 and 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,

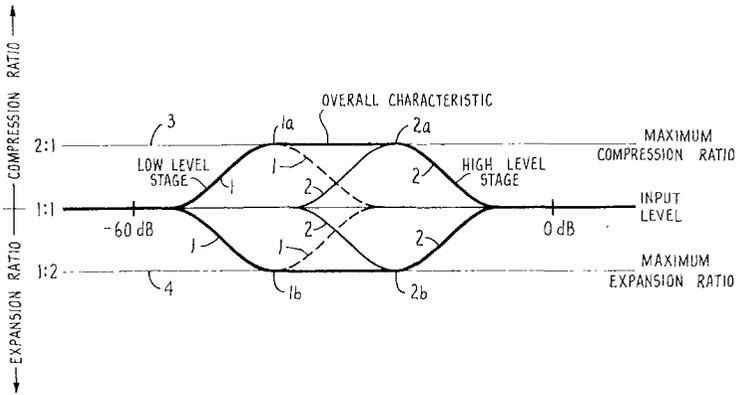


Fig. 3 Action staggering principle.

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 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 maximum expansion ratios are not increased beyond those of single devices, and the advantages of single bi-linear devices are retained. Consequently, any errors occurring within the range of dynamic action caused by the devices in series should not substantially exceed those 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, 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 characteristics, 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 rearranged 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.

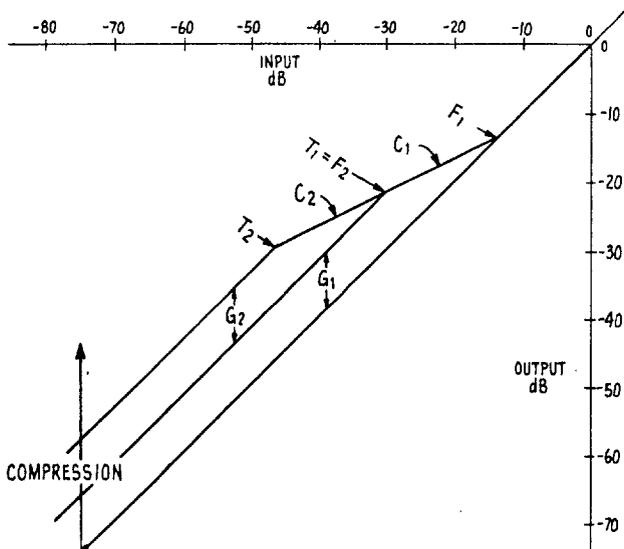


Fig. 4 Idealized construction to show parameter relationships (eq. 1). C is compression ratio, T is threshold, F is finishing point, G is stage gain.

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 (e.g., 1.5:1) would permit an expander to track the compressor more easily, but, on the other hand, the dynamic action would be spread over a wider range of 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 both by 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, thus leading to the simplest circuitry.
2. A signal-dependent variable threshold effect is achieved, which with sliding band stages reduces noise modulation effects.

## 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 practical advantage of relatively straightforward adaptation of existing B-type integrated circuits. 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 3 3/4" ips 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 - 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  $\mu$ S equalization, a quiet residential listening environment, and volume settings corresponding to rather loud listening conditions--i.e., 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 Hz 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 non-standard. 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 pre-amplifiers, 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 wide-band, 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 Mid-Band Modulation Effect

Compressor/expander complementarity requires not only that the expander have essentially the inverse characteristics of the compressor, but also that the transmission channel between the compressor and expander preserve relative signal amplitudes, and preferably phases, at all frequencies within the bandwidth of the signals compressed. As received

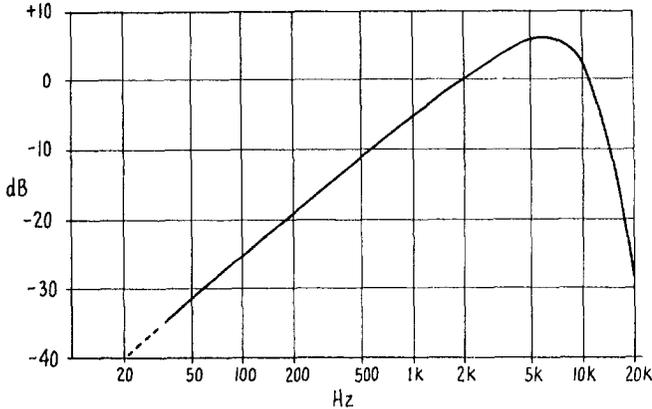


Fig. 5 CCIR noise weighting characteristic (CCIR/ARM).

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, e.g., in the several hundred Hz region. For discussion purposes, this effect will be referred to as the mid-band modulation effect.

In wideband companders an amplitude error at the controlling frequency will manifest itself to the same degree in all other portions of the spectrum; this may or may not be acceptable. In sliding band companders (i.e., 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 mid-band 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 mid-band 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:

- a) 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.
- b) Attenuation of tape recorder bias which may leak into the signal circuits, in order to avoid encoder/decoder mistracking.
- c) Attenuation of rf or supersonic components in the encoder input signal which may otherwise result in audible intermodulation products and/or bias birdies.
- d) Attenuation of supersonic tape noise or other transmission channel noise at the decoder input, in order to avoid encoder/decoder mistracking.
- e) A signal bandwidth definition to promote complementarity of the encoder/decoder--i.e., to reduce the mid-band modulation effect.

Strictly speaking, if an ideal channel exists between the encoder and 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 listed considerations above.

### 3.2 Spectral Skewing

The solution to the mid-band 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--i.e., 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 with about a 10-12 kHz corner frequency--that is, within the useful bandpass of the system, but somewhat below the frequency at which the record/playback response becomes highly unreliable. 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 Figs. 6A-F. Fig. 6A shows the spectrum of a signal that might provoke the mid-band modulation effect (such a signal might be generated by a wideband percussive sound). The compressor control circuit pre-emphasis results in an energy spectrum as shown in Fig. 6B. After rectification, the peak in the pre-emphasized AC control signal spectrum provides the DC signal which controls the sliding band action of the compressor.

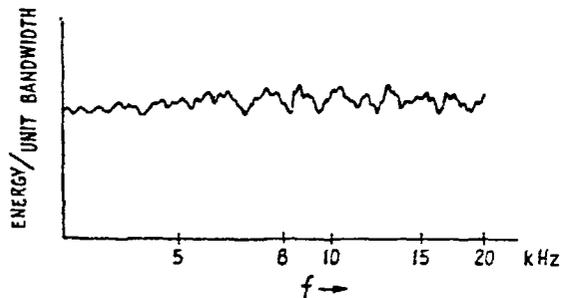


Fig. 6A Representation of spectral distribution of signal having a significant wideband component.

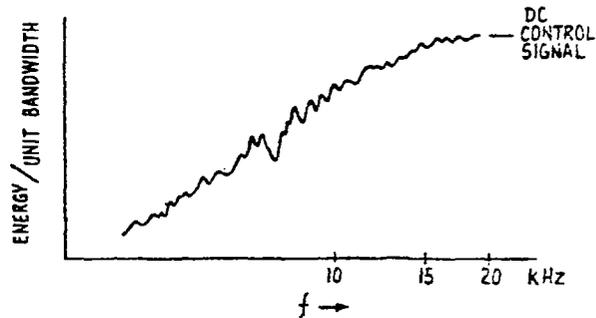


Fig. 6B The signal of Fig. 6A after control amplifier pre-emphasis in the compressor.

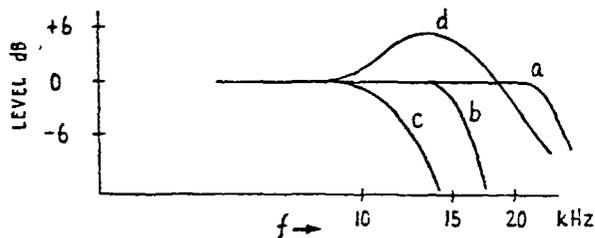


Fig. 6C Four different tape recorder frequency responses (a,b,c,d).

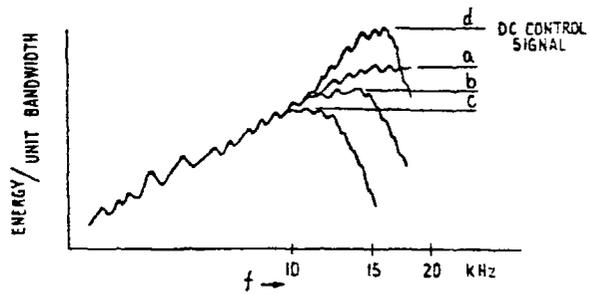


Fig. 6D The signal of 6A, compressed and then sent through recorder channels a,b,c,d, resulting in the above signal distributions at the point of rectification in the expander; different DC control signals a,b,c,d are thereby produced.

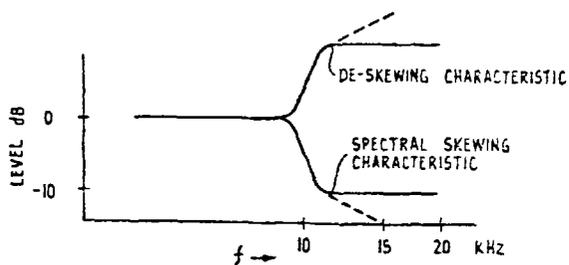


Fig. 6E Idealized spectral skewing characteristic.

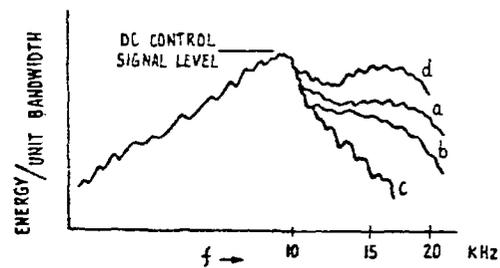


Fig. 6F As Fig. 6D, but with spectral skewing treatment at the input of the compressor; the same DC control signal is produced in the expander with the four different recorder responses a,b,c,d.

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

An idealized spectral skewing characteristic, Fig. 6E, causes the compressor and expander to generate the same DC control signal in each case, as shown in Fig. 6F, which results in accurate decoding not only of the high-frequency signals but 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. 6E. 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 Figs. 8A-D, which show 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 10-12 kHz. Hence, at most levels the spectral skewing network will ensure that there will be a significantly reduced discrepancy in the decoder control signal caused by uncertainties in response at extremely high frequencies.

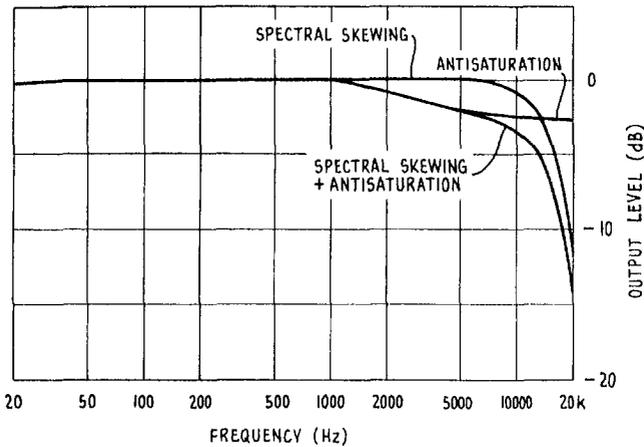


Fig. 7 Spectral skewing and antisaturation characteristics used in C-type system. Overall antisaturation effect is produced by combination of the two characteristics.

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.

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 systems 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 mid-band 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 (e.g., 20 kHz), without provoking the mid-band modulation phenomenon (but a switchable 19 kHz notch should be provided for recording FM broadcasts).

#### 4. SATURATION REDUCTION

Inspection of Fig. 8D shows that high-frequency saturation is a serious problem in cassette recording. Usable peak levels at lower frequencies are some 8 to 10 decibels 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 de-sensitizes the compressor to the frequency components likely to cause trouble during decoding but it 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 mid-band modulation effect would be reduced. The CCIR weighing 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. 6E. 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. 8D (0 dB curve).

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

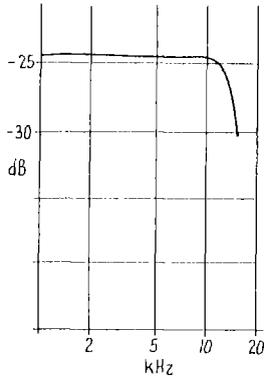


Fig. 8A Recorder A

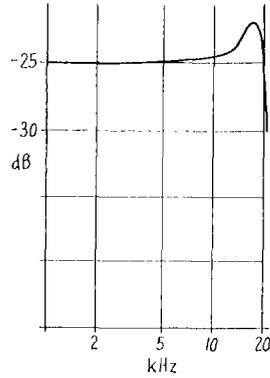


Fig. 8B Recorder B

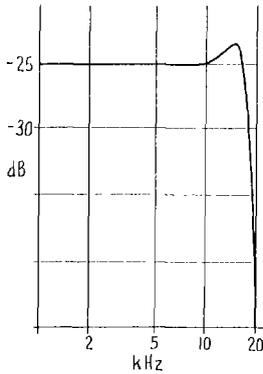


Fig. 8C Recorder C

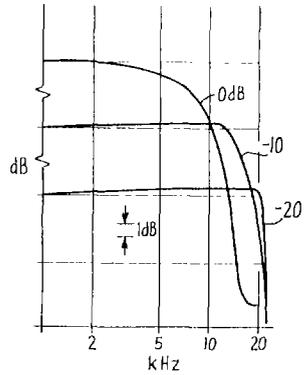


Fig. 8D Recorder D

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

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 anti-saturation 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 for the main and 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.

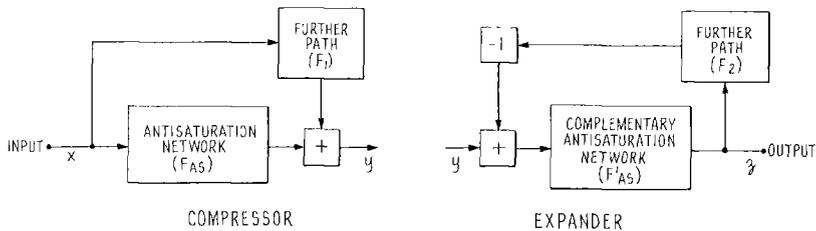


Fig. 9 Placement of antisaturation networks in main signal paths of compressor and expander.

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  $F_{AS}$  be the transfer characteristic of the antisaturation network. Let  $F'_{AS}$  be the required compensating characteristic in the decoder.

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

$$\text{and } z = yF'_{AS} - zF_2F'_{AS} \quad (3)$$

$$\text{thus, } z = \frac{F'_{AS}F_{AS} + F_1F'_{AS}}{1 + F_2F'_{AS}} x \quad (4)$$

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

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 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 (2 resistances and one reactance) with time constants of 70  $\mu$ S and 50  $\mu$ S, corresponding to turnover frequencies of about 2.3 kHz 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. Thanks are due to K.J. Gundry for reviewing the saturation properties of contemporary high performance cassette tapes and for recommending this 50  $\mu$ S/70  $\mu$ S characteristic for use in the C-type noise reduction system. 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 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 anti-saturation 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 play modes with the required complementary characteristics.

For B-type operation the low-level stage and the 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.

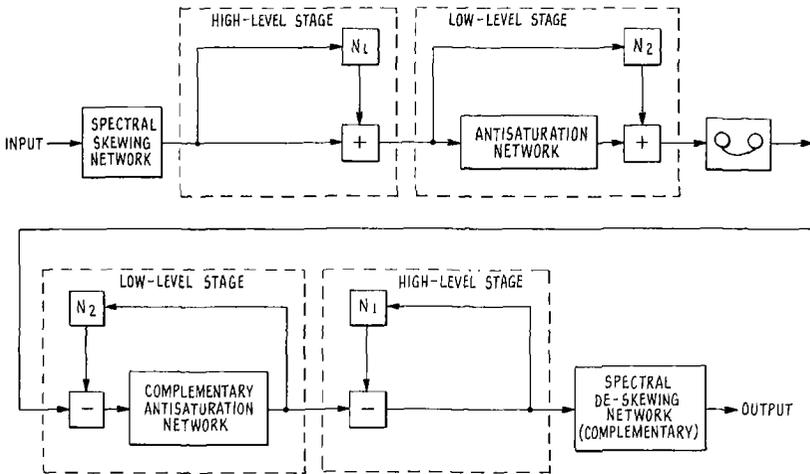


Fig. 10 Basic block diagram of C-type compressor and expander.  $N_1$  and  $N_2$  are the noise reduction networks.

For C-type operation, the spectral skewing and de-skewing networks are switched in. The low-level stage is connected in series, with its pre-set low-level area of action, including overshoot suppression; the variable filter quiescent cutoff frequency is pre-set 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 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 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.

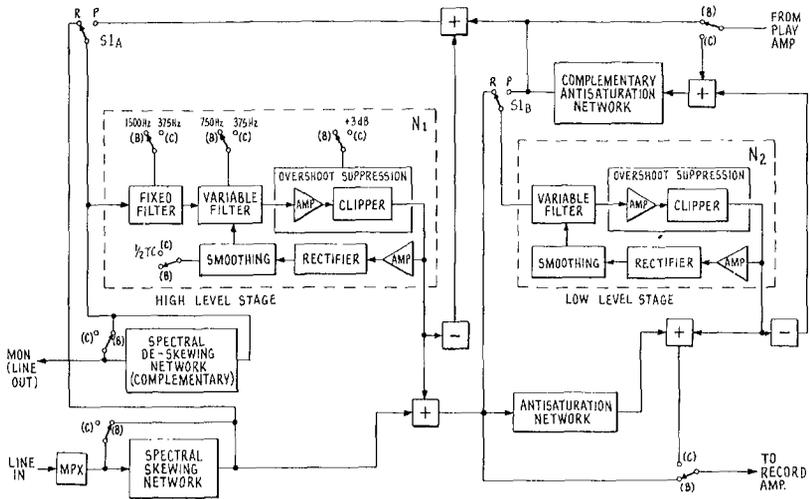


Fig. 11 Switchable record/play and B-type/C-type processor.

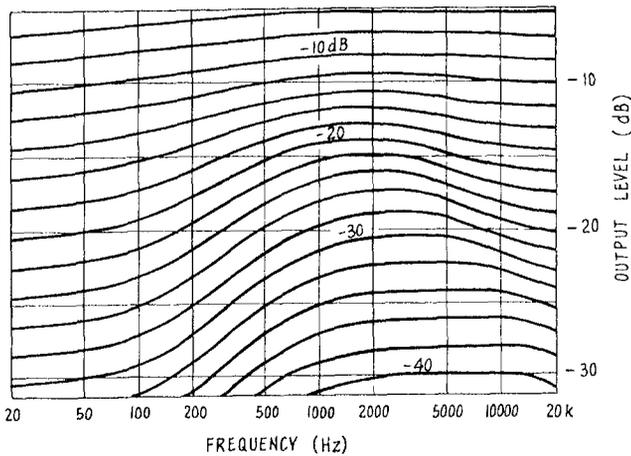


Fig. 12 Input-output characteristics of high-level stage only, without spectral skewing.

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 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 partly provided 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.

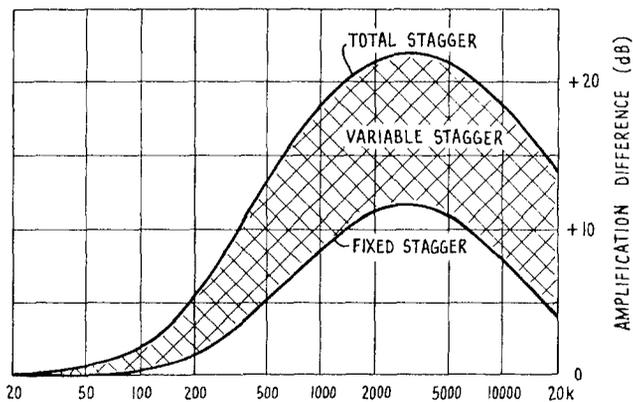


Fig. 13 Control circuit amplification difference between high-level stage and low-level stage. The low-level stage has a fixed gain increase which is augmented by a variable increase caused by the signal processing action of the high-level stage.

For convenience and flexibility, the development of the C-type system was done with discrete components and FET's. When the time came to see how available B-type integrated circuits could be adapted to the design, it was necessary to re-examine the component values used in the variable filter in relation to the variable resistance characteristics of the IC's. Knowing the difficulties which the many B-type IC manufacturers have had since 1971 in designing and manufacturing variable resistances to function accurately over a ratio of some 1000:1, it seemed essential to specify filter impedances which would be compatible with such IC resistance characteristics.

Referring to Fig. 11 and references 1 & 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 RV; 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 Hz & 750 Hz), which yields a quasi two-pole filter 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 IC 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 stretch the capability of RV 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 IC 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. Thanks are due to the audio group at Sony for pointing this out. 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 IC 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 IC 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 mid-band 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 mid-band 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 pre-recorded 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.

## 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.

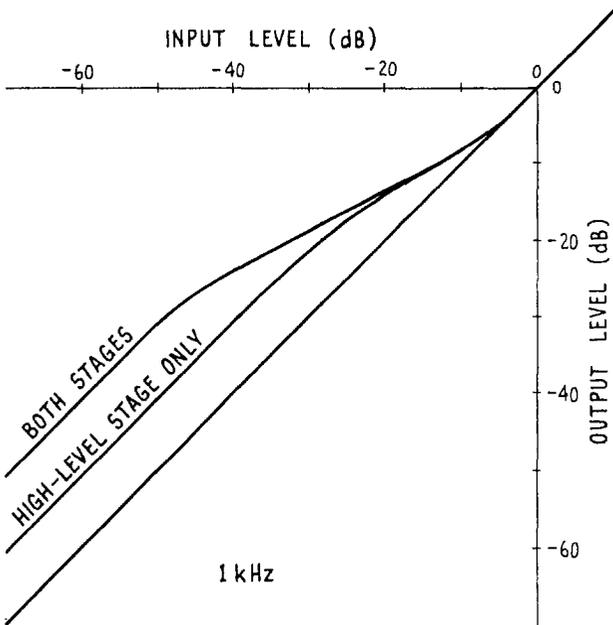


Fig. 14 Input-output transfer characteristic of C-type compressor at 1 kHz. High-level stage also shown.

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.

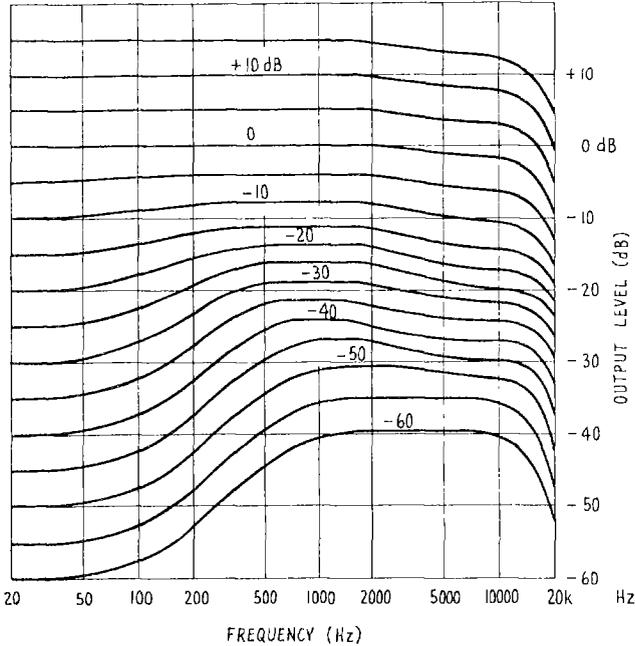


Fig. 15 C-type compression characteristics.

The corresponding expansion curves are given in Fig. 16. While the curves of Figs. 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.

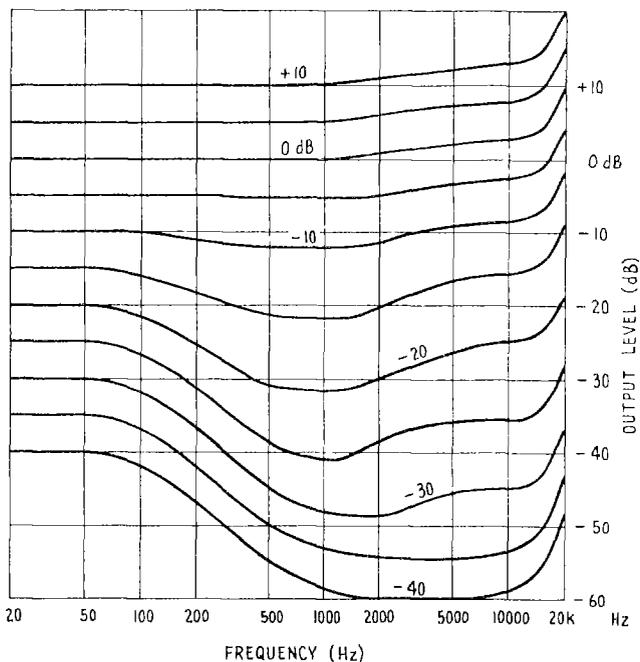


Fig. 16 C-type expansion characteristics.

In Fig. 17 the sub-threshold frequency response 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.

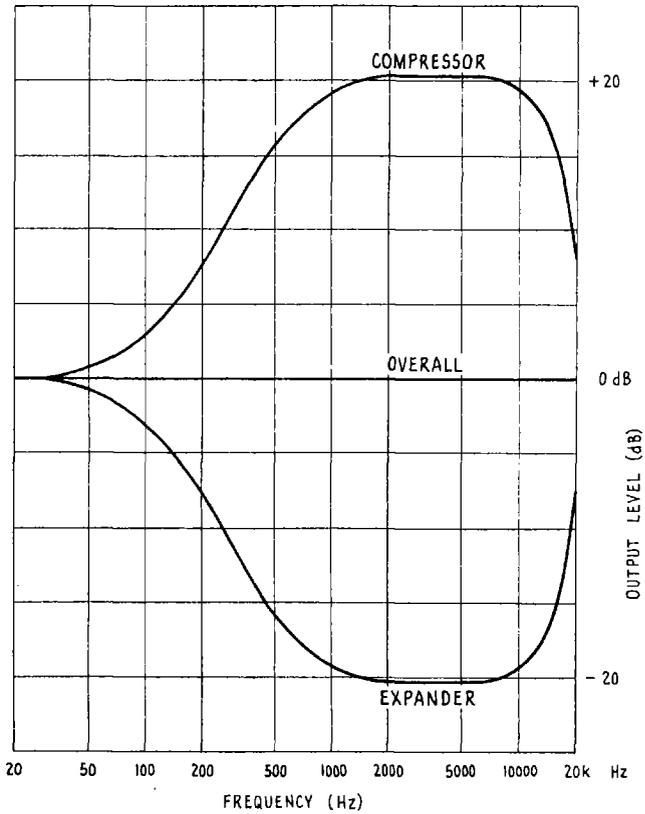


Fig. 17 Low-level (sub-threshold) frequency response of C-type compressor and expander.

These curves 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.

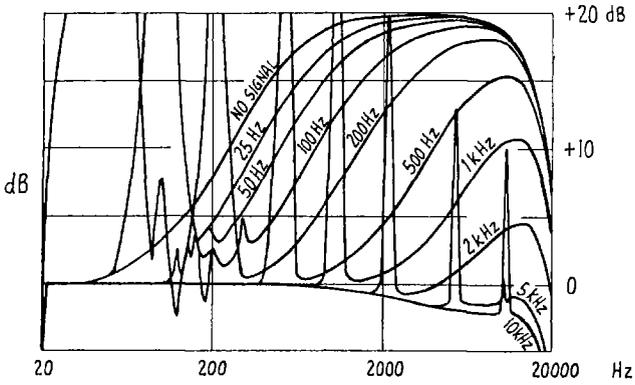


Fig. 18 Sliding band action of compressor in response to high level (0 dB) signals at the frequencies shown (-65 dB probe tone).

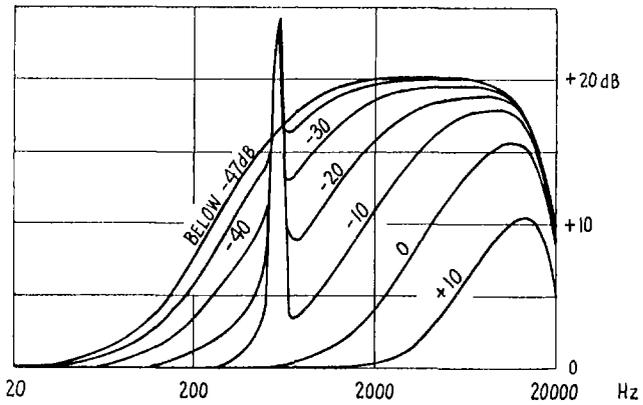


Fig. 19 Sliding band operation of C-type compressor with 500 Hz signals at the levels indicated (-65 dB probe tone).

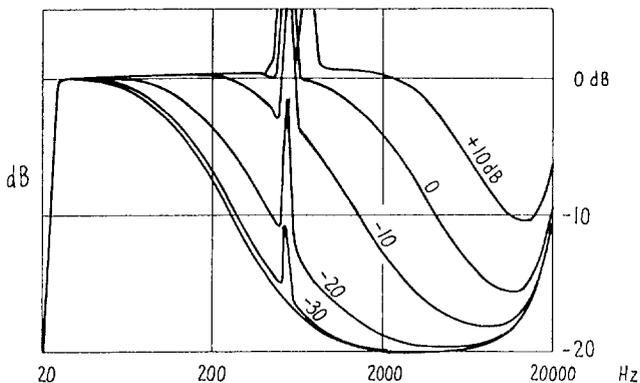


Fig. 20 Sliding band operation of expander with 500 Hz signal at the levels indicated (-45 dB probe tone).

The preceding figures 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 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 and desensitizes the system to high frequency recorder errors.

## 8. ACKNOWLEDGEMENTS

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. Harcastle, J.B. Hull, D.P. Robinson, E.A. Schummer and S.J. Solari. Grateful acknowledgement is also made to a number of licensees whose persistence and enthusiasm provided the stimulation and encouragement to proceed with this venture.

9. REFERENCES

- 1) R.M. Dolby, "A Noise Reduction System for Consumer Applications," Paper M-6, presented at the 39th Convention of the Audio Engineering Society, New York, Oct. 1970.
- 2) R. Berkovitz & K.J. Gundry, "Dolby B-type Noise Reduction System," Audio, Sept. & Oct. 1973.
- 3) K.J. Gundry, "Headroom Extension for Slow-Speed Magnetic Recording of Audio," Paper G-2, presented at the 64th Convention of the Audio Engineering Society, New York, Nov. 1979.
- 4) R.M. Dolby, "An Audio Noise Reduction System," J. Audio Eng. Soc., vol. 15, pp 383-388 (Oct. 1967).