PREPRINT NO. 543

AN AUDIO NOISE REDUCTION SYSTEM

by
Ray M. Dolby
Dolby Laboratories
London, England

PRESENTED AT THE

33rd CONVENTION

OCT. 16 – OCT. 19, 1967

AUDIO ES ®

AN AUDIO ENGINEERING SOCIETY PREPRINT

Convention Price . . . . . . $ .25
By Mail to Members . . . . $ .50
By Mail to Non-Members . $ .85

AN AUDIO NOISE REDUCTION SYSTEM

Ray M. Dolby
Dolby Laboratories, London, England

A noise reduction system which is suitable for use with high quality audio recording or transmission channels is described. A special signal component, derived from four band splitting filters and low level compressors, is combined with the incoming signal during recording or sending. During reproduction, the additional component is removed in a complementary way; any noises acquired in the channel are attenuated in the process. Practical features of the system include: 10 dB (unweighted) noise reduction; imperceptibility of signal modulated noise effects; level frequency response (overall); accuracy of reproduced signal dynamics; low distortion; low internal noise level; stability of characteristics.

Presented April 25, 1967, at the 32nd Convention of the Audio Engineering Society, Los Angeles.

To be published in Journal of the Audio Engineering Society, October 1967.

in accordance with the signal envelope (6-9); in some systems, pilot tones are used in the expansion process (10,11). An electronically switched two-channel (low level - high level) noise reduction system has also been developed recently by Mullin (12).

REQUIREMENTS

Referring to Fig. 1b, it is possible to draw up a set of requirements which any noise reduction system must meet if it is to be used without reservation in high quality recording or transmission channels. The system will necessarily be of the complementary type.

A. Overall signal quality requirements

1. The output signal not to be perceptibly different from the input in frequency response, transient response, and dynamics; stereo signals thus to be perceptibly free of image wandering or shifting.
2. The system not to introduce perceptible non-linear distortion of transient or steady state signals at any level or at any frequency or combination of frequencies; the overload point to be substantially above the normal peak signal level.
3. The system to have low internal noise level; the system not to generate any additional perceptible noises in the presence of signals.
4. All the above requirements to be met in tandem operation of the system (i.e., with multiple processing and de-processing of the signal).

B. Requirements in relation to recording or transmission channel

1. The output from the record/send processor to be suitable for transmission through one channel of normal audio bandwidth.
2. Correct operation not to be dependent upon linear phase-frequency response in the channel.
3. Normally encountered errors or fluctuations in gain and frequency response of the channel not to cause audibly significant changes in the system output.
4. The system not to modify significantly the overall steady state or transient overload characteristics of the channel.

C. Interchangeability requirements

1. The operating characteristics of the system to be fixed and reproducible.
2. The processing units to be sufficiently stable with time, temperature, and other factors to permit interchange of recordings or channels.

D. Noise Reduction Requirements

1. The amount of noise reduction to be perceptibly similar for all types of noises encountered.
2. The noise reduction action to be perceptibly free of signal modulated noise effects with any normally encountered combination of program material and noise.

COMPANDORS

Of the possible noise reduction methods which have been investigated, the syllabic compressor and expandor (compandor) technique (Fig. 2) has been the subject of the most development effort. Since the noise reduction system to be described in this paper may be roughly classified as a compandor, it is worth noting some of the limitations of previous approaches.

Well known compandor difficulties - which by now are regarded as classical - include poor tracking between record/send and reproduce/receive, both statically and dynamically; high sensitivity to gain-errors in recording or transmission; inadequate dynamic range (high noise level vs. high distortion); overshooting with transient inputs; audible modulation product generation under dynamic conditions; distortion of low frequencies by control signal ripple modulation; and production of noticeable signal modulated noise effects.

A comparison of compandor performance with the previously listed requirements for high quality applications shows that the normal compression and expansion approach is unsuitable. Compandors have thus been found to be usable without qualification only in relatively low grade, narrow band applications such as telephone circuits.

NOISE REDUCTION SYSTEM

The general principles of a noise reduction system which is capable of meeting the listed requirements will now be described.

A. Differential method

In normal compression or limiting, a primary object is to modify high level signal dynamics. It is thus unfortunately necessary to subject the signal as a whole to the hazards of passage through a variable gain system. In applying compression techniques to the noise reduction problem, in which the objective does not include modification of signal dynamics, it is unnecessary and undesirable to operate upon high level signal components; noise amplitude in a high quality channel is only of the order of 0.1% of maximum signal amplitude. It would clearly be preferable to generate a small correction or differential component which could be appropriately subtracted from the signal, thereby cancelling or reducing noise while leaving the larger aspects of the signal untouched.

The differential treatment of the signal in the present noise reduction system is illustrated in Fig. 3. The networks (operators) G_1 and G_2 are signal multipliers controlled by the amplitudes, frequencies, and dynamic properties of the signals fed into them. During reproduce/receive, the network G_2 passes low level components (assumed to be noise) back to the subtractor, which partially cancels these components in the signal from the channel. In the process of reducing noise, G_2 and the subtractor also partially cancel low level signal components. To compensate for this cancellation, the network G_1, which has the same characteristics as G_2, adds an identical component prior to recording/sending.

These operations may be expressed in the following way. If the input to the record processor is x (some function of time), the signal in the channel is y, and the output signal from the reproduce processor is z, we have

$$y = \left[1 + G_1(x)\right] x \qquad (1)$$

and $z = y - zG_2(z)$ or $z = \left[\dfrac{1}{1 + G_2(z)}\right] y$ (2)

Combining (1) and (2),

$$z = \left[\frac{1 + G_1(x)}{1 + G_2(z)}\right] x$$

The solution of interest is: $G_1 = G_2$; $z = x$.
The output signal will evidently be equal to the input signal if the record and reproduce differential networks, the operators G_1 and G_2, are identical (subject to the conditions that $G(Z) \neq -1$ and that the resultant functions (1) and (2) are continuous and single valued).

The prime requirement of any high quality noise reduction system -- that the signal should be unchanged overall -- is thus satisfied, and it is necessary only to choose an operator which yields a record/send signal which is compatible with the channel and which produces satisfactory noise reduction properties.

1. Steady State Properties

Referring to the steady state transfer characteristics shown in Fig. 4., the noise reduction requirement, together with the desirability of interfering as little as possible with high level signal components, dictates a reproduce (expansion) curve of the type shown in (b); that is, the gain at low levels must be reduced, while a unity gain condition should prevail at high levels. The required differential component transfer characteristic, shown in (c), is then determined, being linear up to the compression threshold, rising slightly with increasing input, and finally decreasing with larger inputs. Such a characteristic is formed in practice by deriving the compressor control voltage from a combination of feed-forward and feed-back signals.

The record (compression) transfer characteristic, shown in (a), is complementary to the reproduce characteristic, amplifying low level signal components in order to compensate for the corresponding deficiencies produced by the noise reduction action during reproduction.

Comparison of the differential method of forming the compression and expansion laws with the conventional approach depicted in Fig. 2 shows that the scheme has several advantages. Non-linear and modulation distortion are both reduced since the compressor (limiter) contribution is negligible at high levels. System noise problems are alleviated; the variable gain device can be worked at higher levels than would be possible if it were called upon to pass the whole dynamic range.

Tracking accuracy problems between units are also reduced, since the transfer characteristic is largely determined by two readily controlled factors: the compression threshold and the addition or subtraction coefficient of the differential component. At low and high levels the possibility of mistracking is minimum, and in the transition region it is not a difficult design matter to hold the error to a small fraction of a decibel.

A further tracking characteristic concerns compatibility of the system with the audio channel; to a first order, gain variation in the channel manifests itself only as a level change at the output, not as an alteration of signal dynamics. For the parameters used in the present system the maximum tracking error, having a decibel value approximately equal to that of the decibel error in gain, occurs at about 30 dB below peak operating level, where its effect is unobtrusive. The method is thus in practice tolerant of moderate errors in gain. This tolerance is of special significance in stereo, as it enables the noise reduction system to operate without control signal interconnections.

A related matter is the tracking behavior of the system with channels having non-linear phase-frequency response. For a given rms value, the peak and average values of a complex wave depend on the phase relationships of the various frequency components. With a channel of uncertain phase response it is in principle necessary to control the compression and expansion operations using the rms value of the signal, a procedure which at best is inconvenient. However, in practice a combination of peak and average values is a sufficiently accurate indicator of the rms value to permit the use of relatively simple rectification and smoothing circuits; in the present system such circuits are used. Good tracking is thus obtained even when the signal has suffered considerable phase distortion.

A further channel compatibility aspect concerns the possibility of overloading channels with frequency dependent overload characteristics. The overload properties may be further complicated if pre-emphasis is used. Since the pre-emphasis is usually based on the energy probability distribution with frequency for normally encountered sounds, it is evident that any practical noise reduction system should not interfere unduly with this distribution. The compression of comparatively high level signal components thus must be avoided; the transfer characteristic of the present system satisfies this condition (Fig. 4a).

2. Dynamic Properties

Overshoots, arising because of control circuit time lag, normally have amplitudes equal in value to the degree of compression. Such overshoots waste the dynamic range of the audio channel if they are passed linearly, and if they are clipped by the channel various undesirable side-effects can be created: for example, blocking of amplifiers, break-through from groove to groove with discs, interference with other channels if modulated carriers are used. Controlled clipping of the output signal in the compressor itself is a method of avoiding these difficulties, but this approach has the disadvantage of reducing the overload margin.

The usual solution is to make the attack time as short as possible and either to clip within the device or to depend upon the shortness of the overshoot to minimize side effects with clipping in the channel. Unfortunately, the use of short attack times results in side effects in the signal. Rapid changes in gain cause significant modulation products to be generated, which may or may not be cancelled by reciprocal treatment following transmission.

In the present system it is not only possible to confine overshoots to low values but to use relatively long attack times, thereby minimizing modulation distortion. Referring to the differential network portion of Fig. 5, the method used is to follow the compressor circuit (linear limiter) with a conventional symmetrical clipper (non-linear limiter).

A suddenly applied signal is thus momentarily passed without attenuation to the clipper; the differential component is confined to an amplitude which results in negligible overshoot when added to the main signal. In the present system the clipping level has been chosen to limit the overshoot to 2 dB with peak amplitude step inputs.

The addition of the low amplitude clipped signal to the large amplitude pure main signal results in momentary distortion of a few percent, but the degradation is so small and of such short duration (about 1 ms) that it is masked by transient components present in the input signal as well as attenuated subjectively by the relatively slow loudness growth characteristics of the ear (13). In practice, the clipper circuit is rarely called upon to perform its function, the compressor operating linearly except with the most percussive types of program material.

Regarding modulation distortion, it is evident that at high levels such effects are negligible because of the diminished influence of the differential component. By the use of non-linear control signal smoothing circuits, distortion is minimized at low levels as well. A relatively long attack time is used for small variations in signal amplitude, the gain changes produced being slow enough that they do not generate audible modulation products. The time constant is decreased in accordance with the size of the amplitude transition, and for steps large enough to cause the compressor output to exceed the clipper threshold the attack time is reduced to such an extent that the modulation/clipping distortion produced is masked by the transient components present in the input signal.

The use of long attack times as described not only reduces modulation distortion but tends to improve the noise reduction action as well. Since the amount of noise reduction depends upon the amplitude of the differential component in relation to that of the signal in the main path, it is an advantage if short transients of moderate amplitude are prevented from causing unnecessary compression of the differential component. Overshoots of several dB may be produced under these conditions, but they are of such low amplitude as compared with peak level that they are handled linearly in all respects.

While the attack behavior is undoubtedly the most important dynamic aspect of the system, particularly in relation to ensuring compatibility of the record/send processor output with the audio channel, the decay or recovery time is of equal significance when the noise reduction properties are considered. The problem in this regard is to reduce the recovery time to such a value that noise reduction following cessation of high amplitude signals is provided adequately by the residual masking phenomenon (14), in which the sensitivity of the ear is momentarily reduced. The noise reduction action of the system must thus be restored in a time of the order of 0.1 seconds, during which residual masking prevails.

The use of short recovery times in normal compressors results in high distortion at low frequencies. Furthermore, modulation distortion, which was discussed previously from the point of view of attack time, is a product of short recovery time as well as short attack time.

As with the attack aspects of the matter, the recovery problem in the present system is solved jointly by the differential method itself and by suitable choice of characteristics of the non-linear control signal integration circuitry; the smoothing time-constant is made long under equilibrium conditions but is decreased appropriately for large abrupt reductions in signal amplitude. In this way, low frequency distortion in the record/send processor output is readily held to neglibible values at high and low levels and to moderate values (a fraction of a percent) at intermediate levels, while the recovery time is made sufficiently fast that perceptible noise modulation effects are avoided following cessation of high amplitude signals.

Because of the undistorted character of the record/send processor output signal, the system does not depend upon subsequent distortion cancellation during reproduction for correct operation. Phase errors in the audio channel are thus untroublesome; the signal may be re-recorded a number of times or be sent through transmission lines, both being important applications in which non-linear phase-frequency characteristics prevail. It follows also that the signal may be processed and de-processed repeatedly with negligible cumulative distortion effects.

It may be remarked that some of the operating characteristics discussed, which have generally been attributed to the differential method, are in fact properties of the overall compression law produced

(Fig. 4a). Good tracking, high tolerance of gain-errors in the audio channel, avoidance of steady state overloading of highly equalized channels, and negligible formation of modulation products at high amplitudes are features of the overall transfer characteristic, not of the method of forming it. However, the weaknesses of a direct approach to such a transfer characteristic would appear in the usual forms: overshoots, high noise level, high non-linear distortion, difficult reproducibility, and poor stability.

B. Band Splitting

The advantages of the differential method of deriving the compression law are dependent upon there being a large ratio between the maximum amplitude of the signal in the main path and the maximum amplitude of the differential component; it follows that the compression threshold must be set at a low value. Unfortunately, a low compression threshold is detrimental to good noise reduction properties. With moderate and high level signals the noise reduction action technically disappears, so that if only one full frequency compression band were used, an unacceptably high degree of program modulated noise would be evident. This difficulty has been overcome in the present system by splitting the differential component into four frequency bands (Fig. 5). Dependence is placed on the masking effect for subjective noise reduction in portions of the spectrum occupied by signals having amplitudes appreciably higher than the compression thresholds.

Beginning with the early studies of Wegel and Lane (15), investigations of the masking effect have been concerned almost exclusively with the masking of pure tones by tones or noise (e.g. ref. 16). The considerable body of results available is unfortunately not very relevant in the noise reduction system application, in which the masking of a band of noise by one or more tones is of interest. A closer approach to the condiditions required is the masking of one band of noise by another band of noise (17). But systematic research into the use of wideband noise as the maskee has only recently been undertaken (18), and it would appear that it will be some time before sufficient work has been done to permit the choice of noise reduction system design parameters simply by reference to published psychoacoustic data.

In applying the masking phenomenon to the noise reduction system design problem, the number of bands used (circuit complexity) must be balanced against other parameters and the overall system performance requirements. In the present development it was found that for normal high quality audio channels the use of four bands yields satisfactory noise reduction properties while permitting the compression thresholds to be set at a value low enough to obtain the advantages of the differential method as previously discussed.

In the present system, using four bands with compression thresholds of 40 dB below peak operating level, the frequency divisions are as follows: band 1, 80 Hz low pass; band 2, 80 Hz - 3 kHz band pass; band 3, 3 kHz high pass, band 4, 9 kHz high pass. Bands 1, 3, and 4 are conventional 12 dB per octave filters, while band 2 has a frequency response which is complementary to that of bands 1 and 3. The outputs of all the bands are combined with the main signal in such a way as to produce a low level output from the record/send processor which is uniformly 10 dB higher than the input signal up to about 4 kHz, above which the increase in level rises smoothly to 15 dB at 15 kHz.

The figure of 10 dB represents a compromise between a number of design factors, not the least of which is the desirability of minimizing sensitivity to gain-errors in the audio channel. At the high end of the spectrum an extra 5 dB of noise reduction is obtained without appreciable disadvantage in this regard; bands 3 and 4 contribute approximately equally in this region, so that with normally encountered sounds the output of band 3 is usually compressed substantially before the threshold is reached in band 4. The maximum compression ratio is thereby reduced, and the possibility of occurance of program modulated frequency response with gain-errors is decreased.

Because of the use of four bands, with consequent interactions between these bands, the noise reduction properties of the system under signal conditions are not altogether simple. These properties are, however, amenable to investigation and measurement by the use of low level probe tones.

The overall noise reduction action of the system may be summarized as follows: band 1 provides noise reduction in the hum and rumble frequency range; band 2, in the mid-audio range (broadband noise, crosstalk, print-through); bands 3 and 4, in the hiss range. With average orchestral music, band 1 is compressed fairly often; band 2, almost all the time, band 3, fairly often; band 4, rarely. The noise reduction action thus arises most of the time from low and high frequency pre-emphasis and complementary de-emphasis. The high frequency de-emphasis not only attenuates hiss but in magnetic tape recording it reduces high frequency modulation noise. High frequency sidebands of lower frequency signals suffering frequency modulation due to scrape flutter are treated similarly.

CONCLUSION

The general principles of a noise reduction system suitable for high quality use have been described. Low level signal components are amplified in four independent frequency bands prior to recording/sending, which is accomplished by adding the outputs of four filter and low level compressor channels to the main signal. During reproduce/receive the filter and compressor network is connected in a complementary way. Low level components are subtracted from the incoming signal, and the noise acquired in the audio channel is thereby subtracted or reduced as well.

The noise reduction system described is capable of performance of a high order not only with regard to noise reduction, but to signal quality, compatibility of the processed signal with the audio channel, and suitability of characteristics under non-ideal channel conditions.

REFERENCES

1. H.H. Scott, "Dynamic Noise Suppressor," Electronics 20, 96 (1947).

2. H.F. Olson, "Audio Noise Reduction Circuits," Electronics 20, 118 (1947).

3. N.J. Hudak, "Transistorized Audio AGC Amplifier," IEEE Trans. on Broadcasting BC-9, No. 1, 26 (1963).

4. A. Kaiser and B.B. Bauer, "A New Automatic Level Control for Monophonic and Stereophonic Broadcasting," IRE Trans. on Audio AU-10, No. 6, 171 (1962).

5. J.G. McKnight, "Signal-To-Noise Problems and A New Equalization for Magnetic Recording of Music," J. Audio Eng. Soc. 7, 5 (1959).

6. A.C. Norwine, "Devices for Controlling Amplitude Characteristics of Telephonic Signals," Bell System Tech. J. 17, No. 4, 539 (1938).

7. R.O. Carter, "Theory of Syllabic Compandors," IEE Procs. 111, No. 3, 503 (1964).

8. Several papers on compandor techniques appear in Transmission Aspects of Communication Networks, IEE Conference Procs., London 1964 (London: Institution of Electrical Engineers).

9. D. Aldous, "NoisEx Recording System," International Broadcast Engineer No. 10, 510 (July 1965).

10. L.H. Bedford, "Improving the Dynamic Range of Tape Recording," Wireless World 66, 104 (1960).

11. H. Fletcher, "The Stereophonic Sound-Film System," J. of the Society of Motion Picture Engineers 37, 331 (1941).

12. J.T. Mullin, "Advanced Tape Mastering System: Electronic Features," IEEE Trans. on Audio AU-13, No. 2, 31 (1965).

13. W.A. Munson, "The Growth of Auditory Sensation," J. Acoust. Soc. Am. 19, 584 (1947).

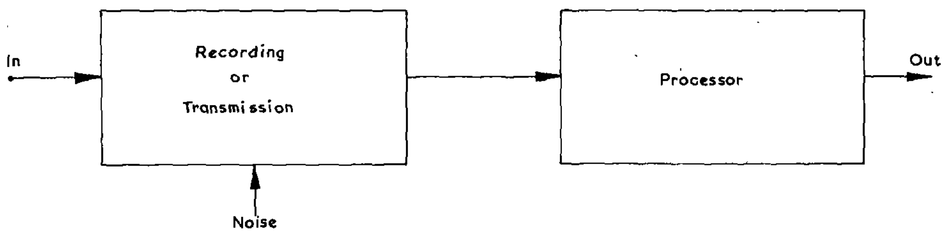
14. E. Lüscher and J. Zwislocki, "Adaptation of the Ear to Sound Stimuli," J. Acoust. Soc. Am. 21, 135 (1949).

15. R.L. Wegel and C.E. Lane, "The Auditory Masking of One Pure Tone by Another and Its Probable Relation to the Dynamics of the Inner Ear," Phys. Rev. 23, 266 (1924).

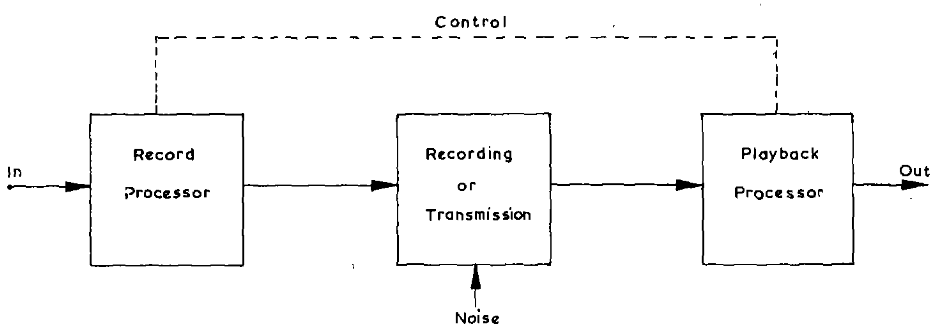
16. J.P. Egan and H.W. Hake, "On the Masking Pattern of a Simple Auditory Stimulus," J. Acoust. Soc. Am. 22, 622 (1950).

17. C.E. Bos and E. de Boer, "Masking and Discrimination," J. Acoust. Soc. Am. 39, 708 (1966).

18. I.M. Young and C.H. Wenner, "Masking of White Noise by Pure Tone, Frequency-Modulated Tone, and Narrow-Band Noise," J. Acoust. Soc. Am. 41, 700 (1967).

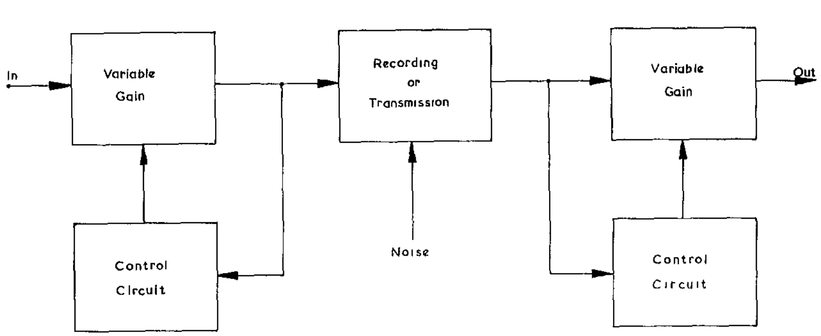


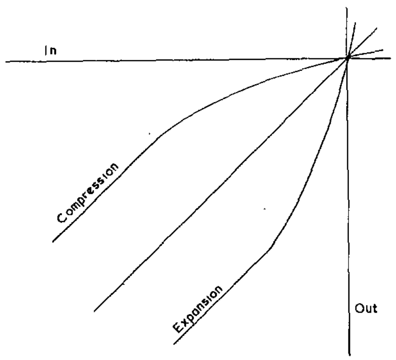
a. Non — Complementary.



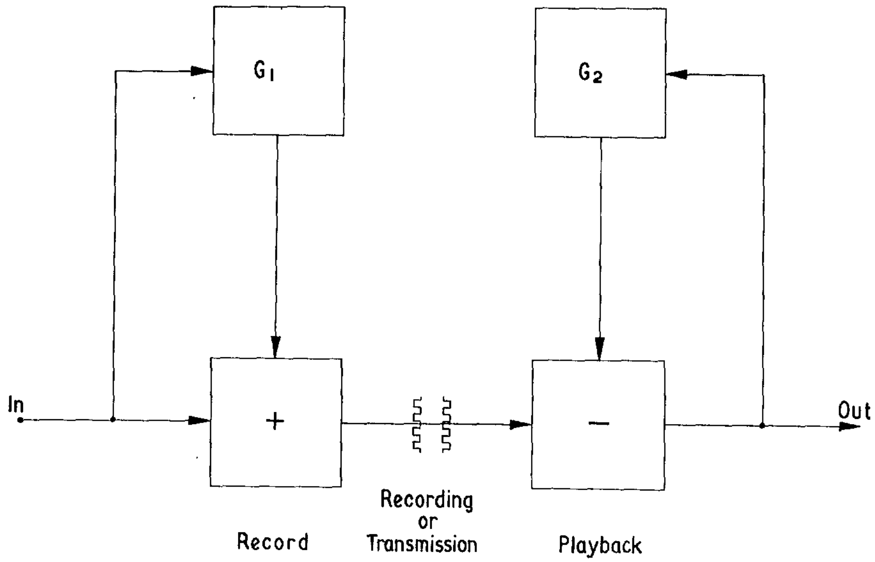
b. Complementary.

1. Noise reduction systems, illustrating two basic types, with reference to effects on signal. The control path shown in (b) is optional.

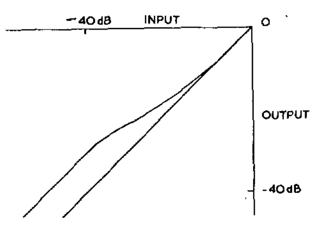




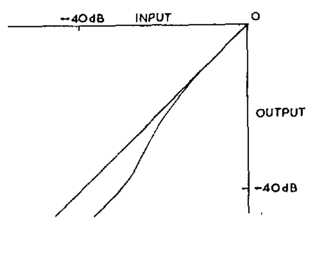
2. Layout and input-output transfer characteristics of compandor noise reduction system. Many variations on the above basic principle have been described (6).



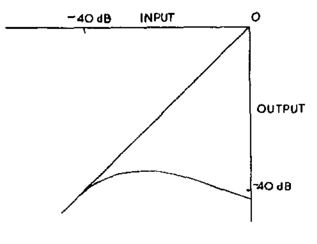
3. Basic layout of noise reduction system. In practice, the operators G_1 and G_2 comprise identical sets of four filters and low level compressors.



a Compression

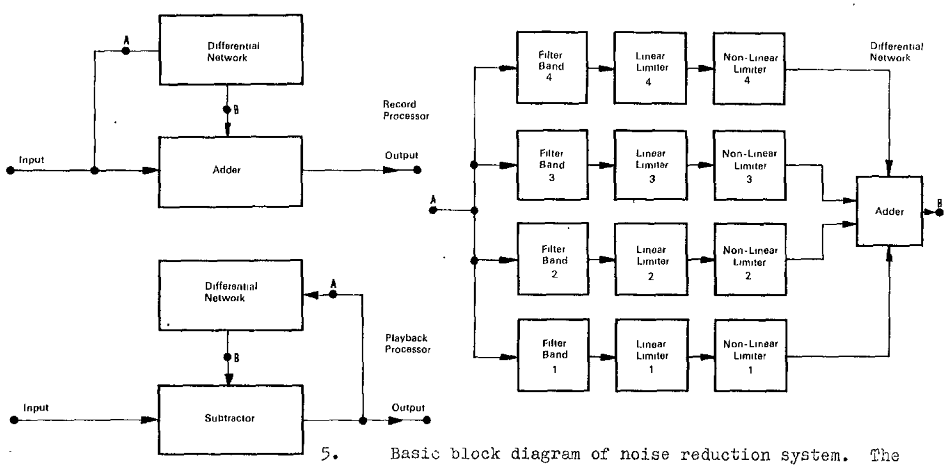


b Expansion



c Differential Component

4. Input-Output transfer characteristics of noise reduction system. The compression characteristic (a) is formed by adding the differential component (c) to the input signal; the expansion characteristic (b) is formed by subtracting the differential component from the recorded/transmitted signal in accordance with the negative feedback configuration shown in Figure 3.



5. Basic block diagram of noise reduction system. The differential network, shown on the right, is the same for both record and playback; the filters and compressors work under identical conditions, both statically and dynamically, in the two modes.