An Introduction to Noise Reduction

by Ken Gundry
Senior Staff Engineer

The complete nine part series on Noise Reduction.

Part I: Noise Modulation

Six years with the BBC and nearly 18 years with Dolby Laboratories have given Ken Gundry a lot to make noise about which he will be doing regularly in "Ken's Corner". His first article begins a series on noise reduction; as he will be making specific references back to it in future articles, you may wish to save this issue.

Understanding noise modulation is fundamental to understanding audio signal processing in general and noise reduction systems in particular. Therefore, noise modulation is the topic of this first article in a series on noise reduction and why NR systems sound as they do.

An ordinary amplifier has noise, with luck very low, which is independent of the signal. If we measure the noise in the presence of a signal and then in its absence, we find the two noise levels and their spectra identical. Most high-quality audio systems are of this nature; even the 78 rpm shellac pressing, with its very poor signal-to-noise ratio, had a surface noise which was substantially independent of the program.

Virtually all equipment intended to reproduce music had a program-independent noise characteristic until perhaps the 1950s, when the first serious attempts were made to design audio noise reduction systems. The phenomenon of noise modulation, whereby the noise level changes in accordance with the program (music, speech, etc.), now became an issue. It had first been noticed in the 1930s when the very first NR systems were attempted by the telephone industry.

The wide-band compander

Noise modulation is most easily explained and understood by examining the operation of the simplest form of NR system, the so-called wide-band compander (compressor/ expander). A wide-band compander's operation is in turn best understood by visualizing an audio tape recorder with a volume control to set recording level, and another one to set playback level.

The usual procedure is to adjust the first control so that during the loudest musical passages, the recording level meters read just "into the red". On playback, we adjust the second control so that those loudest passages are reproduced at some desirable playback level (for example, that which just doesn't cause the neighbors to complain!). When we do this, we find that during the loud passages we aren't particularly worried by tape hiss, although at times we may be able to hear it. On the quietest passages, we find that the tape hiss has the same level as it had during the loud music. But now it is obtrusive, and there may be moments when the music almost disappears into the hiss.

Compression . . .

To avoid losing the quiet music in the hiss, we could turn up the recording level. However, we can't just leave it turned up because the next loud passage that comes along will go hard "into the red", causing gross distortion in our recording. In principle, we could "ride the gain", that is, continually turn the level up when the music is quiet and down when it is loud. However, the composer and performers presumably intended the soft passages to be soft, and would most probably object to our making them louder than they should be, that is, compressing the music's dynamic range.

. . . and expansion

But let us accept for the moment that our tape recording is compressed. Could we not then restore the original dynamic range by expanding during playback? Could we not turn the playback control down on the quiet passages (making them as quiet as they should be), then back up to the original setting for the loud ones? Lo and behold, if we do that, we have a crude manual NR system. Turning down the volume on the soft passages has the effect of lowering the tape hiss just at those times it would otherwise be most obtrusive.

Obviously, such manual control is impractical. But replace it by some ingenious electronics and we would appear to have a useful NR system, that is, a wide-band compander. But is there a catch?

Yes, there is, and it lies in the statement a few paragraphs back that while not worrisome, noise is nevertheless audible at times during loud passages. When we use this simple NR system, that doesn't change: the hiss during the loud passages is the same as it would have been without the system (Figure 1). But during the quiet passages with this NR system, hiss is less than it was before, even inaudible (Figure 2). Thus by definition we have noise modulation. The noise level is dependent upon the program material; it goes up on loud passages and down on soft ones.

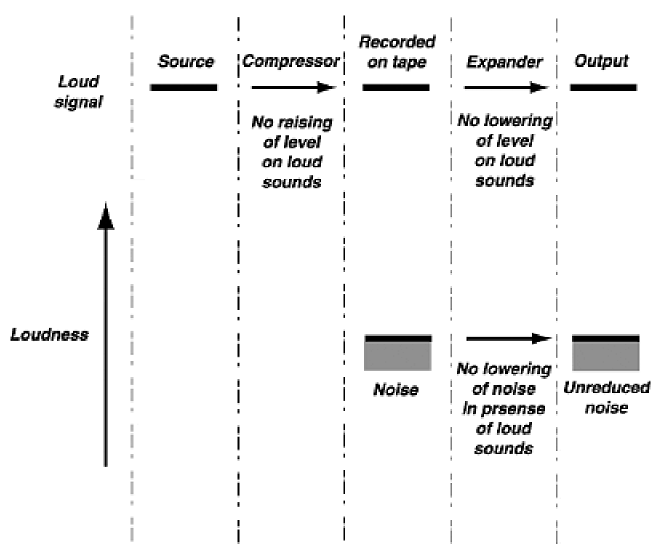


Figure 1: Wide-band compander on loud signals.

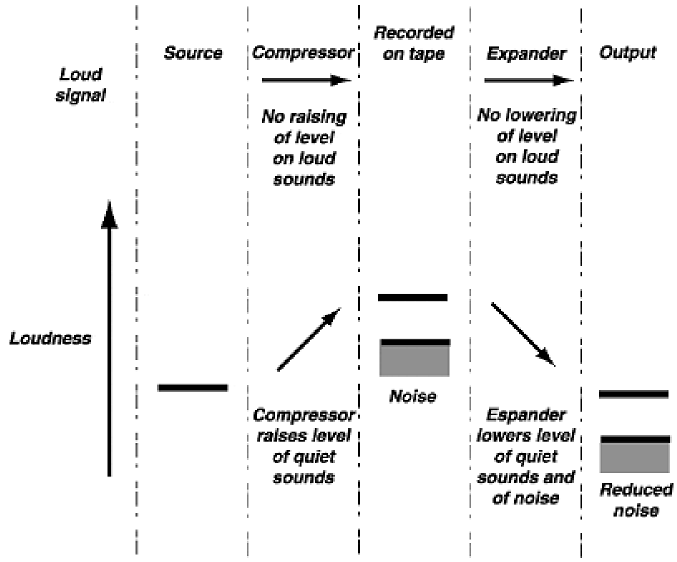


Figure 2: Wide-band compander on soft signals.

Of waterfalls and cable cars

Now that is a catch indeed. Noise modulation doesn't occur in nature. The sound of a waterfall, for example, isn't affected by the presence of other sounds. Indeed, we soon get used to a constant noise, and only notice it when it changes significantly or stops. I used to live on a cable car line in San Francisco, and did not notice the continuous low-pitched whine of the pulleys until the cable stopped at 2 a.m.–thereby waking me up. In other words, we are sensitive to changes in what we otherwise expect to be unchanging. (My theory is that evolution has left us with a sense that such changes warn of the approach of danger.)

Whatever the reason, a varying hiss level behind music is much more disturbing than a fixed one, even if the noise only rises to a level which was not very obtrusive in the first place. Any audible changes in the apparent hiss, however small, are immediately perceived as unnatural, something which wasn't a part of the original performance.

Therein lies the paradox of the simple wide-band compander system. Noise modulation is inevitable. It could only be innocuous if the highest noise level ever reached in the presence of a loud signal is in itself inaudible. However, because in the presence of a loud signal the compander provides no NR, that means the original noise level would have to be inaudible without the NR system! To put it another way, if without the NR system noise is at any time audible in the presence of loud signals, with it you will be able to hear noise modulation.

If we confine ourselves to companders based on simple, electronically-controlled volume controls, there is no solution to this paradox; all such companders give audible noise modulation under some operating conditions. However, there are alternative design approaches which are not susceptible to noise modulation. Watch this space.

Part II: The Ideal Noise Reduction System

In our last issue we saw why wide-band companders, basically electronic volume controls, inevitably give rise to noise modulation. We also discovered that if the noise of a medium without the noise reduction system is audible (as it obviously would be if it were a candidate for NR), then the noise modulation produced by a wide-band compander will also be audible.

Now I haven't been quite fair to such devices. On "busy" program material such as rock music or full symphony orchestra, a wide-band compander actually works quite well, and the noise modulation may not be apparent. However, on simpler sounds, for example solo clarinet or piano, the noise modulation can be obvious and disturbing. The significant difference is that the busy material spreads across a good part of the audible range of frequencies, whereas at any particular instant the solo instruments only occupy a small part of the range.

Masking

To understand the significance of this difference, we must know something about human hearing, and in particular the phenomenon known as masking.

The solid line in Figure 1 shows the sound pressure level at which a sinewave or a narrow band of noise is just audible, that is, the threshold of hearing. Sounds at levels above the curve are audible, those below it are not. This threshold is clearly very dependent on frequency. We are able to hear a much softer sound at say 4 kHz than at 50 Hz or 15 kHz. At 25 kHz the threshold is off the scale: no matter how loud it is, we can't hear it.

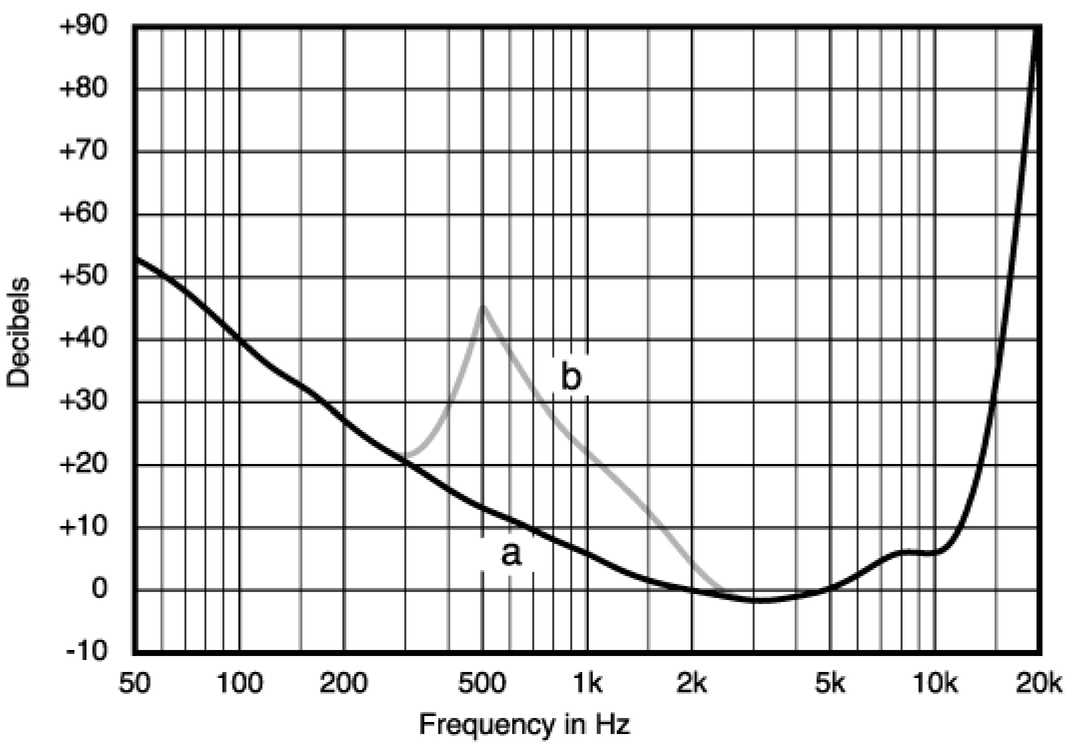


FIGURE 1 THRESHOLD OF HEARING WITH a) NO SIGNAL, b) LOUD 500Hz SIGNAL (AT +70 dB SPL). 0 dB = 20 μPa.

Now consider the threshold in the presence of a relatively loud signal at one frequency, say 500 Hz as shown by the dashed line. The threshold has risen dramatically in the immediate neighborhood of 500 Hz, modestly somewhat further away, and not at all at remote parts of the audible range.

This rise in the threshold is called masking. In the presence of the loud 500 Hz, we don't care about noise in its immediate region because it is hidden, or masked, by the loud signal. Further away, the noise level could rise somewhat above the no-signal threshold, yet still be below the new masked threshold and thus inaudible. However, in remote parts of the spectrum, any noise which was audible without the 500 Hz will remain just as audible with it.

How Masking Relates to Noise Reduction

As we have seen, masking is not dependent upon the mere presence of loud signals; it depends upon where they are spectrally. Thus if we design a noise reduction system to be level-sensitive only, noise modulation is inevitable.

Such a system, that is to say a wide-band compander, will let the noise rise on the busy music mentioned above. That's not likely to be a problem, because such music contains many frequencies distributed across most of the range, and therefore gives a masked threshold curve which is raised everywhere relative to the unmasked curve. All will be well on silence, too, because the compander will provide full boost in record and full cut (and thus full noise reduction) on playback.

But on the solo instruments, it's a different story. The wide-band compander will react to them as it does to the busy music, permitting the noise level to rise. However, in reality the solo instruments consist of relatively loud sounds confined to a small part of the spectrum, thus giving masked curves more like Figure 1. This means that the rise in noise level will be audible (i.e., not masked).

What the ideal noise reduction system would do

We can now see what a noise reduction system must do to eliminate audible noise modulation, and thus be truly effective. It must be sensitive not only to the level of the program material, but also its spectral distribution. In order to provide a constant noise reduction effect, we must provide constant noise reduction at all frequencies where there are no loud sounds and thus no masking. We can allow the noise level to rise in the presence of loud sounds only where it remains below the masking curve corresponding to those loud sounds.

Bear in mind that we reduce noise by turning up the gain during recording, and turning it back down by the same amount during playback. When there are loud sounds at particular frequencies, we cannot turn up the gain as much, if at all, or we would overload the recording medium. Therefore we cannot obtain as much noise reduction at those frequencies. Fortunately, however, that is exactly where we don't need as much noise reduction, because masking occurs at those frequencies.

Therefore, in the absence of an input signal, the ideal noise reduction system would apply a fixed boost during recording, and of course an equal cut during playback, this cut being what actually reduces the output noise. When a loud sound at a particular frequency comes along, loud enough that the fixed boost at that frequency would cause overload, the recording boost would be reduced only at that frequency and its immediate neighborhood (Figure 2). This is what at Dolby Laboratories we call "the principle of least treatment."

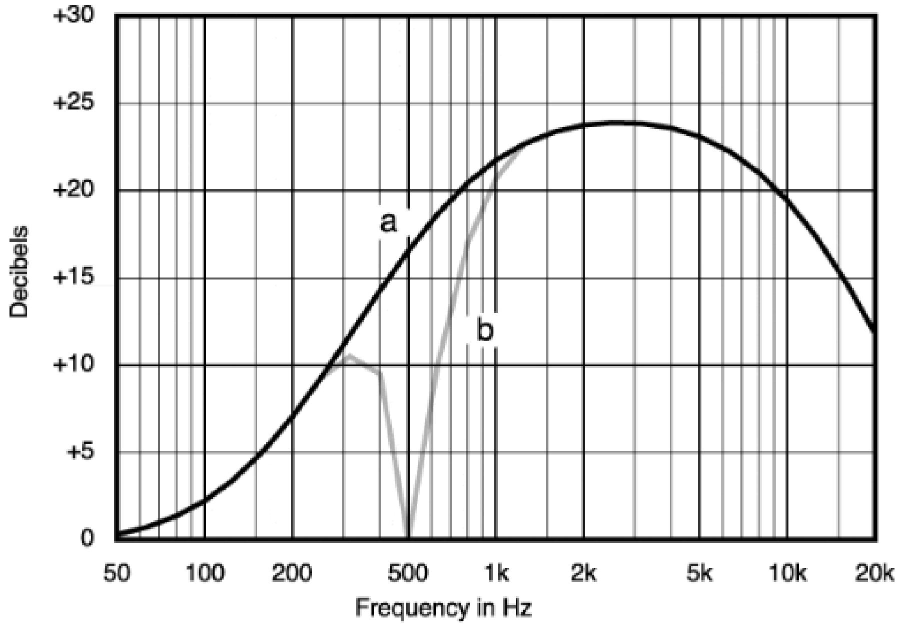


FIGURE 2 BOOST APPLIED IN ENCODER WITH a) NO SIGNAL, b) LOUD 500 Hz SIGNAL

During playback, the cut at that frequency and its neighborhood would be reduced correspondingly, reducing the noise reduction effect there--but would remain unchanged elsewhere (Figure 3). Thus the output noise level would change only at and near the frequency of the loud sound, where masking would conceal the change. At all other areas of the spectrum, the noise would be reduced to an unchanging level. There could be no audible noise modulation.

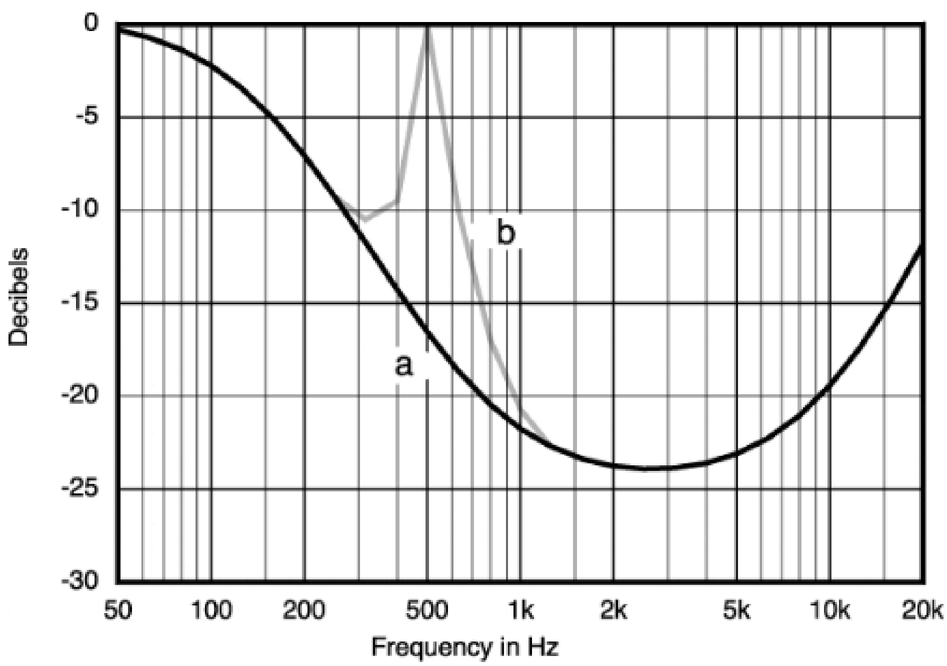


FIGURE 3 CUT APPLIED IN DECODER WITH a) no signal, b) LOUD 500 Hz SIGNAL.

Now that we have specified one vital characteristic of the ideal reduction system, we can move on to the others. But that will have to wait until next time.

That Waterfall Again

It may have occurred to you to ask, isn't otherwise audible noise becoming inaudible in the presence of a loud signal also a form of noise modulation? Yes, it is, but since this is what happens in nature, it is not disturbing. When you stand next to a loud waterfall and someone shouts in your ear, the voice masks the noise of the water in some parts of the spectrum, but you don't perceive this as the water turning on and off

Part III: Dynamics of Noise Reduction Systems

The problem of overshoots

The main purpose of a compressor is to raise the level of quiet signals more than that of loud ones. This means that if signals have been quiet and suddenly get loud, the compressor has to react by reducing the gain. However, because it doesn't "know" in advance that the signal is going to get loud, a simple compressor overshoots, that is, it provides too much gain for the short time it takes to respond after the fact to the sudden increase in input level.

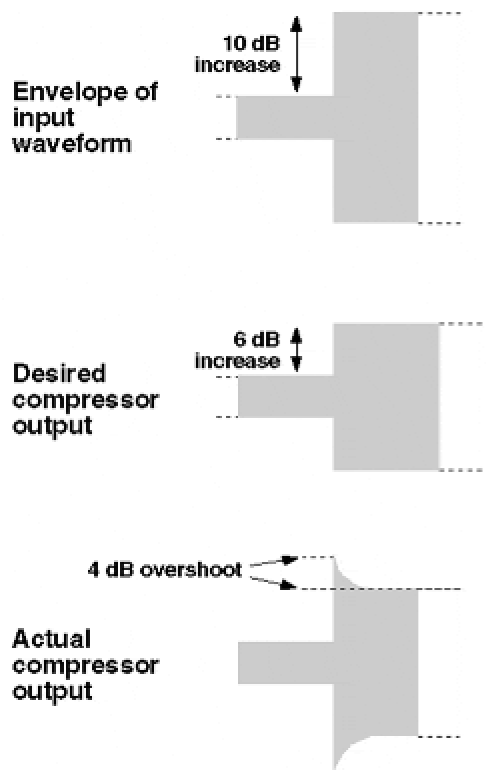


Figure 1. Compressor overshoot.

I show this for an elementary case in Figure 1. At the top is an input signal consisting of a sinewave which suddenly rises 10 dB (about 3x) in level at point A (the figure doesn't show the detail of the waveform, only its envelope). Consider a compressor designed to reduce this change to 6 dB (2x), that is, reduce its gain by 10 minus 6, or 4 dB. At the middle is the output waveform we would like the compressor to provide. However, what we actually get is an overshoot 4 dB in magnitude, as shown at the bottom.

Obviously, overshoots on high-level signals can drive equipment further down the line into overload, therefore producing audible distortion. In general, the size of an overshoot--the amount by which the output level exceeds the desired one--is equal to the amount by which the compressor has to turn down the gain. The duration of the overshoot depends on how long it takes the compressor to react (its so-called attack time).

Compressors as amplitude modulators

At first glance, it might seem that the problem of overshoots could be solved by giving the compressor an attack time so short that overshoot is neglible. To see if this would work, it is necessary to consider a compressor as a sort of amplitude modulator.

When a single sinewave is disturbed in any way, such as by changing its amplitude, it does not remain a single sinewave. It becomes a spectrally more complex signal containing more than one frequency. The more the original sinewave is disturbed, the more extra frequencies, called sidebands, are generated. In addition, the faster the rate of the disturbance, the further away from the original frequency the extra frequencies are.

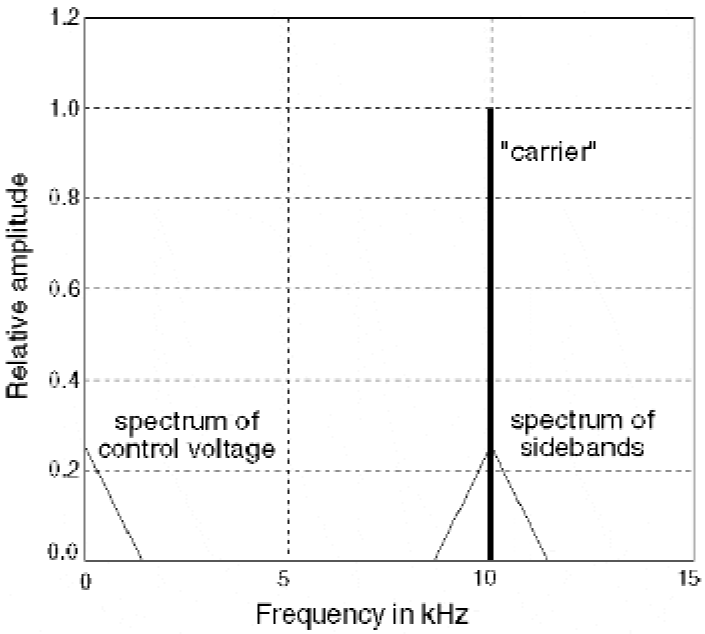


Figure 2. Slow attack time.

This is the principle of AM (amplitude modulation) radio, and it also applies to a compressor. The compressor subjects the audio input signal, which is equivalent to an AM radio carrier, to a changing gain. Because the amplitude of each component frequency of the input signal is changed, sidebands are added to each frequency.

In Figure 2, I have arbitrarily chosen a moment when a gain change results from a compressor's control voltage with the spectrum shown in the bottom left-hand corner. If the input signal contains 10 kHz, the output will contain that 10 kHz plus sidebands having the same shape as the spectrum of the control voltage. Every other input component will have sidebands added in the same way. Specifically, if a gain reduction of 6 dB is called for, the sum of the upper-frequency sidebands will be roughly 25% of the audio output level, and so will the sum of the lower ones. Thus we could say crudely that during the gain change, the compressor gives 50% modulation distortion!

Varying Attack Time

I mentioned above that the faster the rate of disturbance, in this case, the faster the compressor's attack, the further away from the original frequencies the sidebands are. If the attack time is long, say tens of milliseconds, the gain control signal contains only low audio frequencies. Thus each pair of sidebands is close to the corresponding input frequency component, as shown in Figure 2. Remembering from Part 2 of this series that lower-level unwanted components close to wanted ones are masked, slow attack therefore gives distortion that is innocuous (i.e. essentially inaudible).

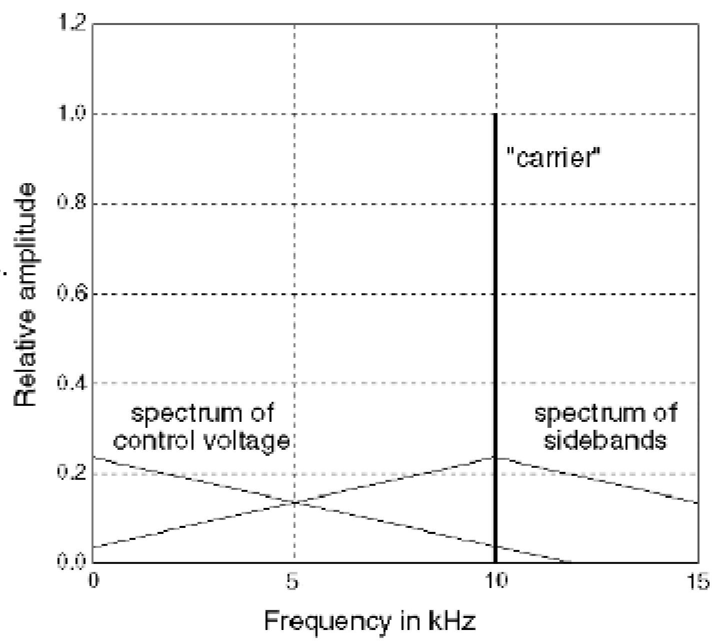


Figure 3. Fast attack time.

However, if the attack is fast, the gain control signal contains high frequencies (Figure 3). Thus the sidebands are spread out far from the audio frequencies of the input signal, and so may well not be masked. Indeed, because the distortion occurs briefly during the gain reduction, and consists of frequency components spread widely across the spectrum, an audible click is usually added to the signal transient (see sidebar). Hence for a simple audio compressor/limiter, the choice of attack time is a compromise between distortion downstream caused by overshoots, and distortion (clicks) from the processor itself.

On the surface, it might seem that with a complementary noise reduction system, the clicks caused by rapid attack would not matter because the subsequent complementary expansion would cancel them out. However, that would require the expander to generate equal sidebands with exactly opposite polarity. If the expander received via the medium precisely what the compressor generated, sidebands and all, then click cancelation might work. Unfortunately, in the real world, the medium will always have imperfections in addition to added noise, such as a bandwidth limitation and/or phase distortion. Thus the expander will receive a slightly modified signal, and will generate slightly different sidebands which cannot cancel those from the compressor. The result will be audible modulation distortion.

No one solution

Such distortion can be avoided by a combination of measures, none of which is sufficient alone. First, overshoots on low- to medium-level signals are not likely to overload anything following the compressor. Therefore we might as well accept them on such signals, and use long attack times (tens of milliseconds) and their attendant freedom from audible modulation distortion.

However, something different must be done for high-level signals, leading us to use program-adapting attack times. Our second measure is to choose a different compression characteristic on high-level signals, one which minimizes overshoots. Remembering that overshoots occur only while the compressor gain is being reduced, we can best accomplish this by fixing its gain for high-level signals. There can be no overshoots where there are no gain changes.

Third, we can confine fast attack to those instants when the resultant modulation distortion will be masked. This requires knowledge of the signal spectrum at those instants, and choosing an attack time, and thus a sideband spectrum, which takes advantage of masking. A high degree of frequency dependence would help immensely here, and fortunately, for other reasons, we have already established frequency dependence as vital to an ideal noise reduction system (as explained in Part 2 of this series).

Switching gain at zero-crossings

Switching the gain at a zero-crossing does not prevent the clicks which result from fast attack times. Yes, it would eliminate a possible discontinuity in the waveform, but there still may be a discontinuity in the slope. The resultant disturbance would be smaller, but still wideband.

You can easily prove this if you have a toneburst generator available which permits varying the switching point. Listen to, say, a 100 Hz burst; it stops and starts with a clearly audible click, even when the burst stops and starts at zero-crossings. The click can only be eliminated by spreading out the beginning and end of the burst over several cycles, such as by feeding it through a bandpass filter, or using a multiplier with a smooth "switching" waveform, that is, by simulating a slow attack.

Next time we'll look at compression characteristics, how to express and interpret them, and which ones satisfy the requirements we have discussed this time.

Part IV: Compression and Expansion Curves

So far in this series we have seen that a good noise reduction system, one which reduces noise without audible side-effects, must have a high degree of frequency dependence and adaptive dynamic characteristics. Before delving any deeper into the design and operation of NR systems, I want this time to explore compression and expansion curves.

Compression and expansion characteristics are illustrated by "transfer curves." As we will see, these curves can provide valuable clues to the effectiveness of an NR system. Unfortunately, they are often not properly understood.

About the graphs

Let me begin by explaining some criteria for all the graphs in this article:

- The horizontal axis represents the input level and the vertical axis the output level of a processor. The levels are expressed in dB with respect to a reference 0 dB at the top right. For NR systems, 0 dB is the maximum level which the noisy medium can pass without distortion.
- The lines on the graphs--the transfer curves themselves--represent the output/input level relationships of various processors. To facilitate explanation, all my examples use straight lines, but in reality such circuits often yield smooth curves without abrupt changes in slope.
- For now we are assuming that these relationships are "steady-state," that is, we are considering moments when the gain has settled after a change in input level (I will relax this condition in a later article).

What they show

Any point on these graphs has two coordinates: an output level and an input level. For example, consider a processor which with an input level of -15 dB yields an output of -10 dB. In Figure 1, this signal condition is represented as point 1 (input = -15, output = -10). The difference between the output and the input levels is (-10)-(-15), or +5, so at this point the device is behaving as an amplifier with a gain of 5 dB.

Let us further say that with this processor, an input of -30 dB would give an output of -25, represented at point 2 (input = -30, output = -25). At this point, too, the device behaves as an amplifier with a gain of 5 dB.

Let us also say that for all inputs between -30 and -15 the processor also provides a gain of 5 dB. The locus of these points is a straight line between 2 and 1; that line is

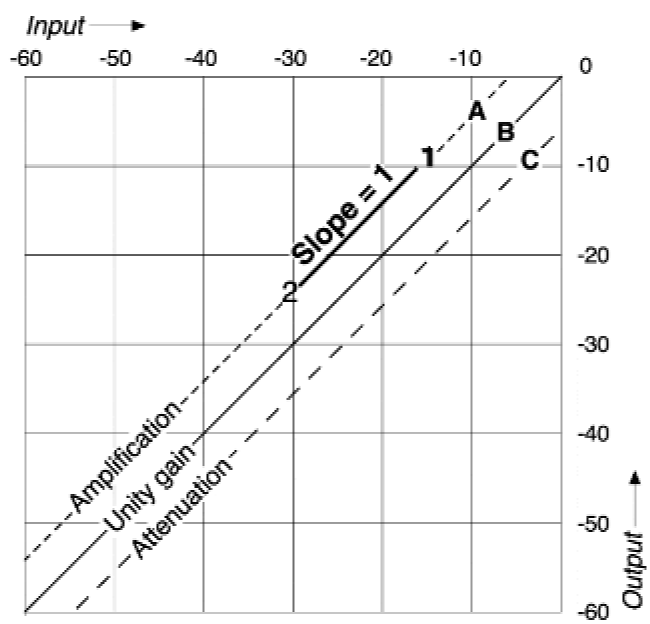


Figure 1: A straight line with a slope of 1 means that a processor acts as a constant-gain amplifier (A), a unity-gain amplifier (B), or a constant-loss attenuator (C).

the "transfer curve." Because any change in input results in an identical change in output, the curve is said to have a slope of 1 (see curve E in Figure 4.1 for further clarification). Should the device act as a constant-gain amplifier for all input levels, we can extend the line on the graph in both directions without limit (line A), if we ignore the inevitable overload in one direction and noise in the other.

Now consider an amplifier with unity gain. Its output level is equal to its input, so all points on its transfer curve have equal x and y coordinates. Thus it would be represented by the straight line B, which also has a slope of 1, passing through the point (0, 0). We use this line as a kind of reference; it represents the characteristic when we bypass the processor.

We can now make some generalizations:

- A processor with a transfer curve consisting of a straight line with a slope of 1 will provide constant gain (amplification), unity gain, or constant loss (attenuation).
- The processor’s function depends upon the curve’s location with respect to the reference line B. If it lies to the "north" of the reference, the processor is a constant-gain amplifier. If it lies on the reference, it is a unity-gain amplifier. If it lies "south" of the reference, it is a constant-loss attenuator (line C).
- In fact, any point north of the reference represents amplification and any point south represents attenuation, by an amount indicated by the vertical spacing from the reference line.

More than one slope

Of course, not all processors are simple amplifiers or attenuators acting the same for all input levels, as a glance at their transfer curves will immediately reveal. If the transfer curve in its entirety is not a straight line with a slope of 1, the processor being described is not a simple amplifier or attenuator with constant gain or loss. However, portions of the curve may be straight and have a slope of 1, indicating that over that range of input levels the processor behaves as a simple amplifier or attenuator. For example, Figure 2 represents a circuit which at input levels below -20 dB behaves as an amplifier with a constant gain of 20 dB. At higher inputs, however, the device no longer amplifies, but limits: it permits no further increase in output level no matter how much the input level increases (I will have more to say about limiters later on).

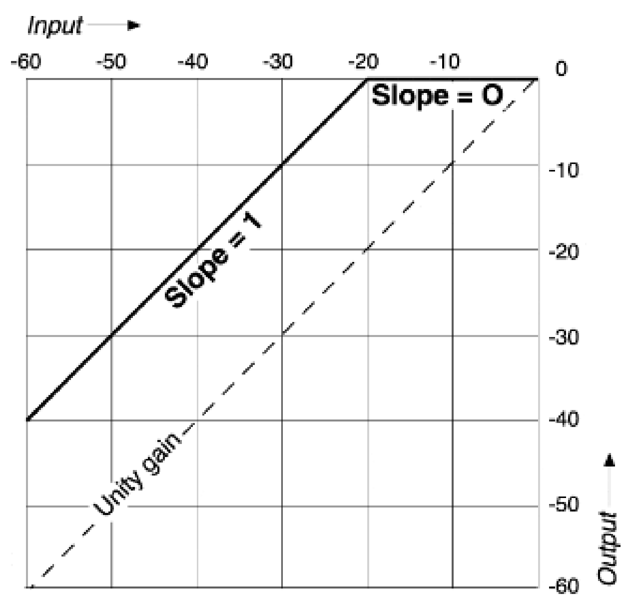


Figure 2: With inputs below -20 dB, this processor acts as an amplifier with a constant gain of 20 dB (slope = 1). With inputs above -20 dB, it acts as a *limiter* (slope = 0).

Slopes of more and less than one

With constant- or unity-gain devices, when the input level changes, the output changes by the same amount. With a compressor, however, when the input level increases, the output increases by a lesser amount. This results in a slope of less than 1 (i.e. less than 45 degrees to the horizontal). An expander does the opposite: when its input level increases, its output increases by a greater amount, resulting in a curve with a slope of greater than one. These characteristics are illustrated in Figure 3.

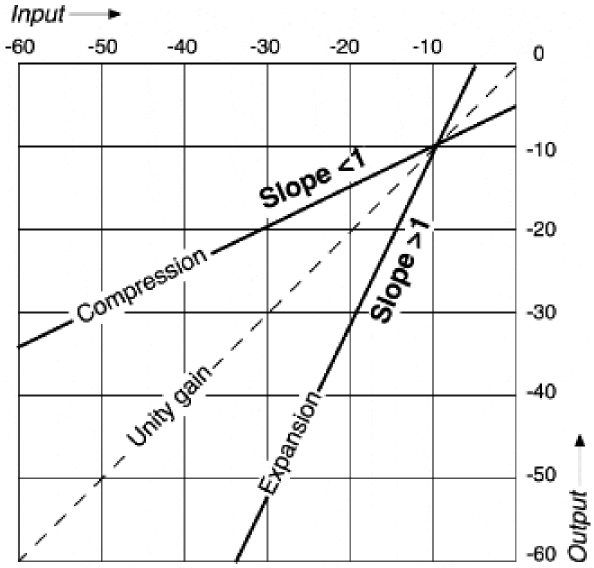


Figure 3: A processor with a slope of less than one is a compressor; with a slope of more than one, an expander.

Figure 4.1 shows the compressor from Figure 3 (curve D), plus a constant-gain amplifier for contrast (curve E). Note that for the amplifier, when the input level changes 1 dB, the output level also changes 1 dB.

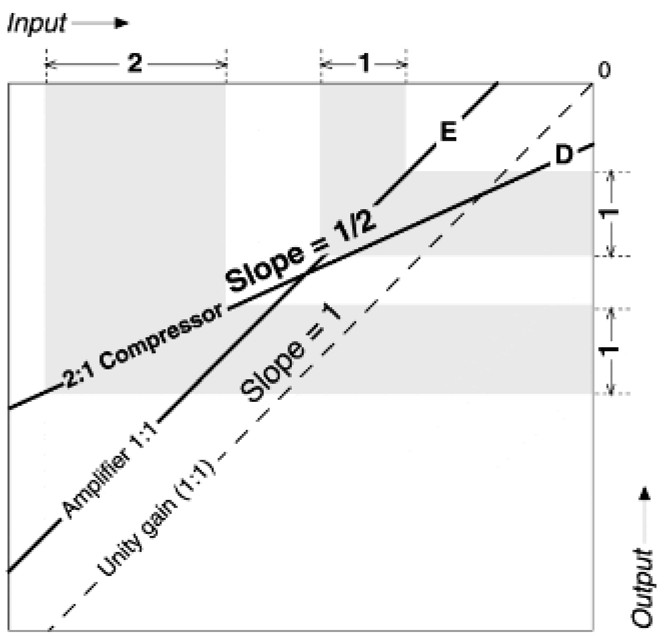


Figure 4.1: A compressor with a slope of 1/2 is said to have a compression ratio of 2:1.

With the compressor, however, it takes a 2 dB change in input level to effect a 1 dB change in output level. This compressor therefore has a slope of 1/2.

Figure 4.2 shows an expander (curve F) which is the inverse of the compressor in Figure 4.1, plus a constant-loss attenuator for contrast (curve G). As with the amplifier, when the attenuator's input level changes 1 dB, so does its output level. But with the expander, a 1 dB change in input level results in a 2 dB change in output level. Therefore the expander has a slope of 2.

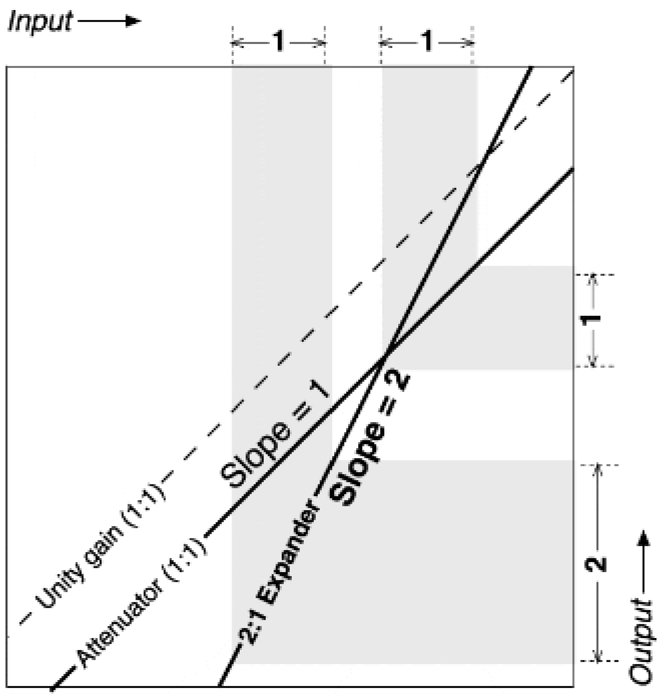


Figure 4.2: An expander with a slope of 2 is said to have an expansion ratio of 2:1.

When compressors and expanders are discussed, reference is often made to their compression and expansion "ratios." With an expander, its expansion ratio at any particular point is simply its slope at that point (that is, how much its output changes for a 1 dB change in input level). Thus the expansion ratio of the expander depicted by curve F in Figure 4.2, which has a slope of 2, is said to be 2:1.

For no good reason that I can think of, a compressor's compression ratio is not the same as its slope (that is, a number less than one), but is its slope's reciprocal. Thus the compressor represented by curve D in Figure 4.1, which has a slope of 1/2, is not said to have compression ratio of 1:2 (as logic would dictate), but rather one of 2:1, the same as the expansion ratio of its complementary expander.

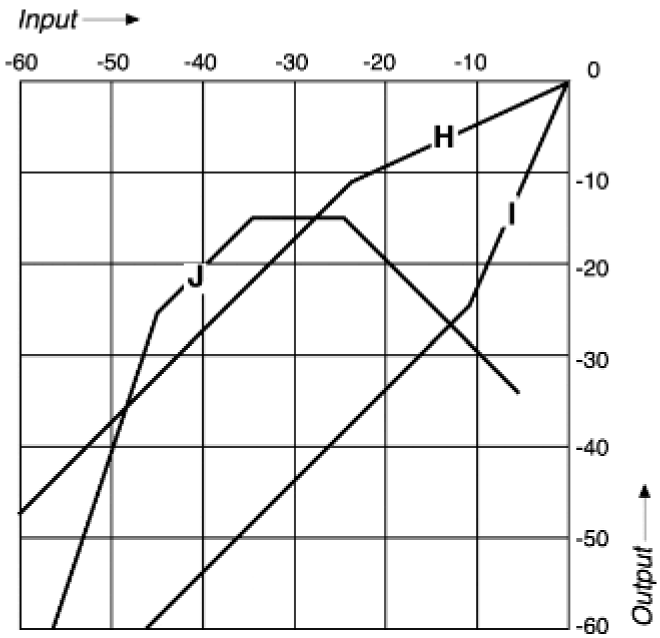


Figure 5: See text.

The value of these curves

If I have done my job properly, you should now be able to look at the transfer curve for any processor and quickly identify what it does. Take a look at the examples in Figure 5. Is it clear that curve H shows a kind of compressor which provides 2:1 compression in the top 25 dB of the dynamic range, and a constant gain of 12.5 below the -25 dB threshold? That curve I is H's complementary expander? And that curve J is a rather fanciful expander which expands with a 2:1 ratio on inputs below -45 dB, provides a constant gain of 20 dB from -45 to -35 dB, limits from -35 to -25 dB, and "re-enters" above -25 dB?

From now on, when someone talks about a "compressor" or an "expander" without further qualification, ask to see a transfer curve. Without it, you won't know much. But with it, you'll quickly glean what you need to know. What you need to know, for example, to assess the usefulness of the processor as a noise reduction system--a topic I shall get to eventually.

Part V: More on Compression and Expansion Curves

In part IV, we first explored transfer curves for compressors and expanders under "steady-state" conditions, and the levels that exist after the gain has settled to its final, constant value. This time I will discuss how a compressor or expander responds to ***changes*** *in input signal level, and how we can predict some aspects of dynamic behavior from transfer curves. Then I shall move away from generalities and back towards noise reduction systems.*

Overshoots again

In part III, I showed that compressors usually produce overshoots on rapid increases in input level, and that trying to eliminate them simply by using very short attack times led to other audible problems. If very short attack times are not the answer, then what *can* we do about overshoots? In trying to answer that question, we must first determine the size of the overshoots we are likely to encounter.

Figure 1 shows the transfer curve of a 2:1 compressor with unity gain at an input level of 0 dB. At fixed input levels of -20 and-10 dB, the output levels would be -10 and -5 dB respectively (the points marked A and B). In other words, gains for inputs of -20 and -10 are 10 and 5 dB respectively, which are shown by the spacing between the compressor transfer curve and the 45-degree reference line representing a fixed gain of 0 dB.

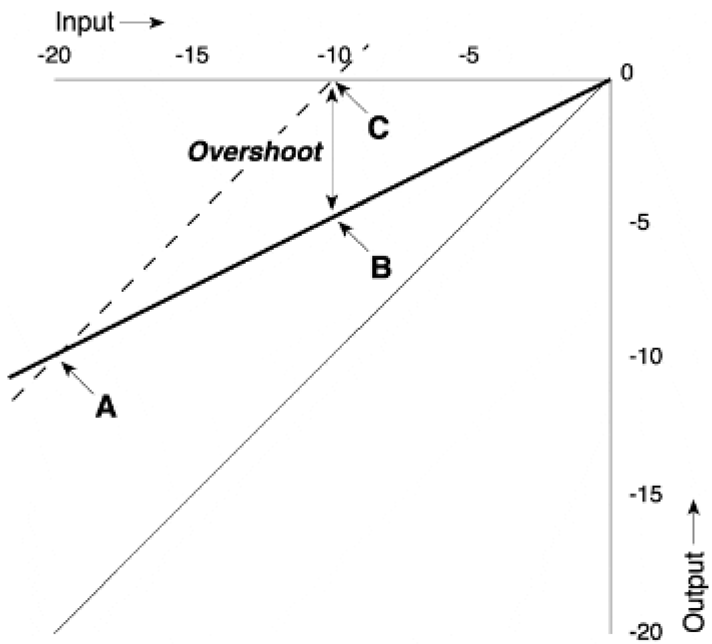


Figure 1: When the input signal changes suddenly from -20 dB to -10 dB, a 2:1 compressor overshoots by 5 dB.

Now let's look at what happens when an input level of -20 dB suddenly *changes* to -10. We know from the previous paragraph that this 10 dB transition requires the compressor to reduce its gain from 10 dB to 5 dB. As we saw in part III, however, doing that will take the compressor a finite time. At first its gain stays at 10 dB, even after the input level increases. The dashed line represents this condition: the input signal has risen to the new input level, -10, but is still subjected to 10 dB gain, thus yielding an output of 0 dB (point C). This is 5 dB above the final value (point B), which will be obtained only after the compressor settles back down to the "normal" transfer curve. Thus we have an overshoot of 5 dB. As we saw in part III, the overshoot is equal to the degree of gain reduction required by a particular increase in input level.

In Figure 1 we also can see a way to estimate the size of an overshoot by examining the steady-state transfer curve. Draw a line (dashed) at 45 degrees through the point on the curve representing the initial input condition, follow the line up to a point directly above the final input condition, and see how far it is above the steady-state curve.

Let's follow this procedure for a different transition in input level, from -15 to -6 as shown by Figure 2. The steady-state output corresponding to -15 is -7.5 (point D). As in Figure 1, draw a line at 45 degrees through this point (dashed). Follow this line up to point F directly above the final steady-state output for the new input of -6 (which is -3, point E). The distance between E and F, 4.5 dB, is the overshoot.

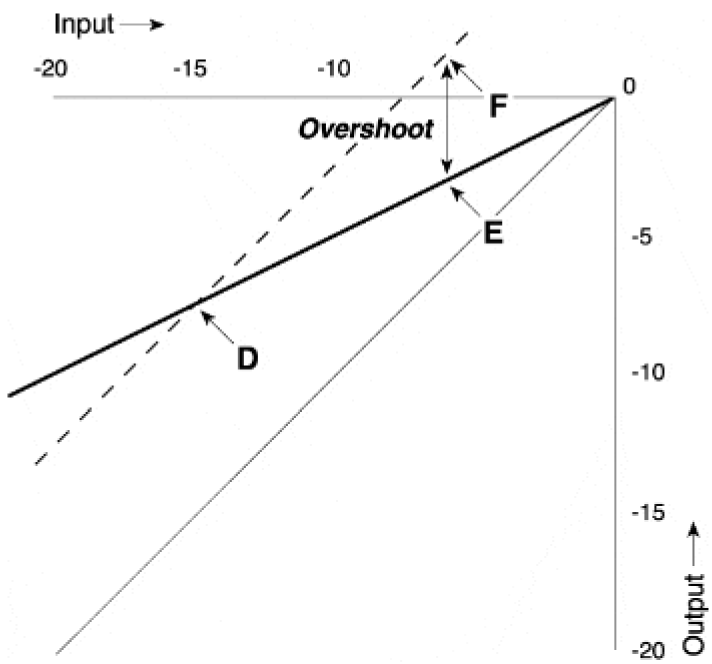


Figure 2: When the input signal changes from -15 dB to -6 dB, the 2:1 compressor overshoots by 4.5 dB.

Note that this time the overshoot goes above 0 dB. If the equipment which follows the compressor overloads at 0 dB, it will distort the compressed signal. Furthermore, if the compressor is the first part of a complementary process such as a noise reduction system, the subsequent expansion will not in fact be complementary, because the compressed signal will not be reproduced accurately. Thus we will get mistracking as well as distortion.

For this compressor having a constant compression ratio of 2:1, we didn't really need this procedure to predict the size of the overshoot. Any increasing transition requires a gain reduction equal to half the size of the level increase, and hence yields an overshoot equal to half the increase. However, the procedure is useful for estimating the size of an overshoot when the compressor has a more complex transfer curve.

Consider for example the compressor in Figure 3. Above an input of -20 it has a slope of 1 and a fixed gain of 0 dB, while below -20 it has a slope of 1/2, representing 2:1 compression. By the way, even though no compression occurs above the -20 threshold, this device as a whole is nevertheless considered a compressor because overall it has an average slope of less than 1.

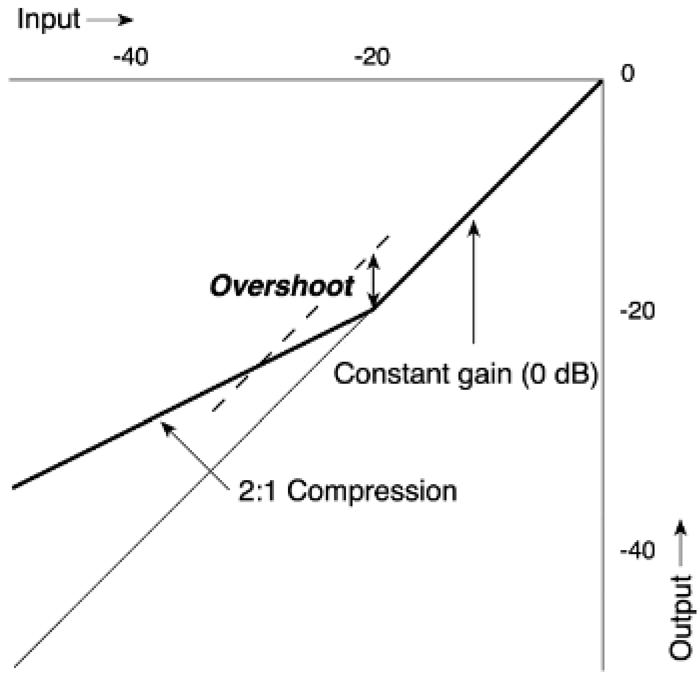


Figure 3: This compressor will overshoot on transitions below -20 dB, but not on transitions above that level.

Following the procedure explained above, for 10 dB input transitions from -30 to -20 and from -10 to 0 we find overshoots of 5 and 0 dB respectively. In other words, for the transition at high levels there is *no* overshoot. The low level transition does give an overshoot (equal to the degree of gain reduction), but it does not get anywhere near an output level which would overload subsequent equipment. In fact for *any* transition occurring within the top 20 dB of the dynamic range, there is no overshoot and no transient overload of the equipment downstream. Thus we have the potential for complementary expansion without distortion or mistracking.

I hinted at this sort of characteristic in part III, and as I suggested there, this sort of compression curve with its fixed gain at high levels is not a complete answer to overshoots, but helps immensely.

Bilinear compression characteristic

Figure 3 shows a compression curve which at high levels is a straight line with a slope of 1, representing constant gain or loss. When the signal lies in this range, the device is linear. Its input and output signals have a constant relationship, the output contains no spectral components (harmonics, intermodulation, etc.) which were not present in the input, and any change in input level gives an exactly equal change in output level.

Back in part II I discussed the ideal noise reduction system. One of its properties was that for low-level signals which give little or no masking of noise, the encoder/compressor must have a constant gain. This would result in a complementary expander/decoder with a constant loss, thus providing a constant noise reduction effect and yielding a constant noise level. For the encoder this property is shown in Figure 4. At low levels, the transfer curve is a straight line with a slope of 1 (constant gain), converting to compression above a predetermined threshold (-40 dB in this example). Below that threshold the system is linear.*

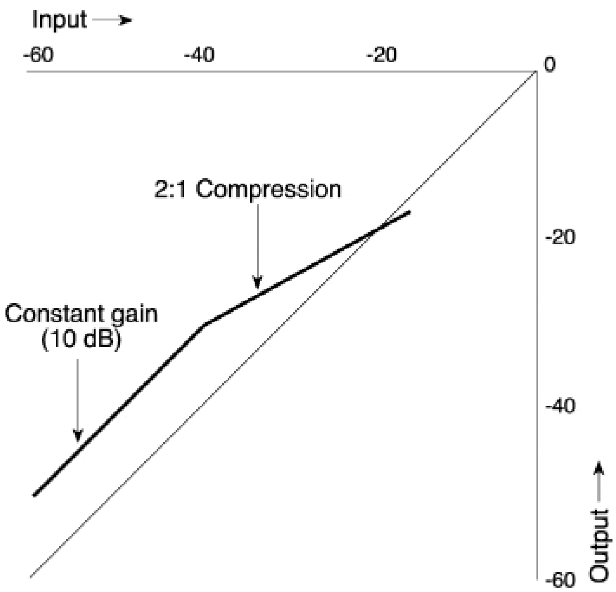


Figure 4: This compressor will overshoot on transitions between -40 and -20 dB, but not on transitions below -40.

If Figure 3 shows ideal treatment of high-level signals, and Figure 4 of low-level signals, why not combine them? That would yield the device depicted in Figure 5, which has what is known as a *bilinear* characteristic, obviously because there are two regions in which the device is linear. All actual compression (that is, gain-changing) is confined to an inter-mediate region. All the Dolby noise reduction systems are based on this bilinear characteristic. Indeed, it is difficult to imagine a satisfactory NR system–one free of audible noise modulation and transient distortion–which used any other shape of compression curve.

Thus the first step we took towards discovering ways to minimize overshoots–determining their size–has led to a means of keeping them from happening in the first place. By fixing the compressor's gain at 0 dB at high signal levels where no noise reduction is needed, we prevent overshoots from occurring–and prevent distortion and mistracking. By fixing gain at some predetermined amount at low signal levels, we also prevent overshoots–and provide an ideal noise reduction property.

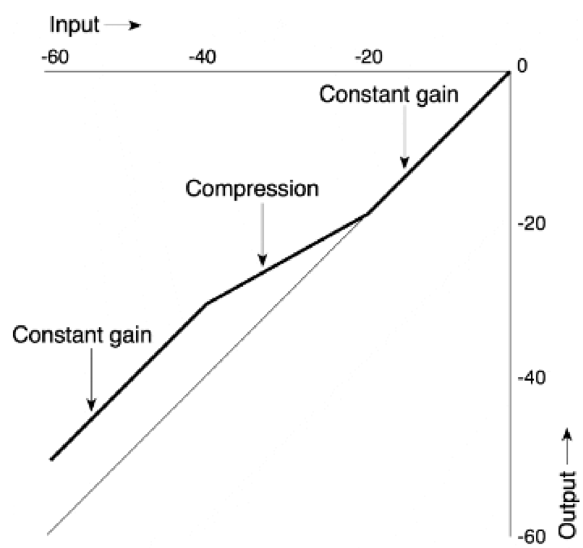


Figure 5: The bilinear compression characteristic.

But discovering the ideal compression characteristic and the other desirable properties of the processors making up a noise reduction system is one thing. Realizing them is quite another. Next time I shall therefore start to talk about the practicalities.

*For this example I have again shown 2:1 compression because it is a practical figure and makes the arithmetic easy, but there is nothing magic about it. Furthermore, practical systems change more gradually from constant gain operation to compression than shown here, resulting in smooth curves that make it difficult to tell just where the transition from one mode to the other occurs.

A Word About Frequency Dependency

As we discussed in Part II, in the ideal noise reduction system compression would occur only in those parts of the spectrum where noise is not masked naturally by music. Thus in real-life NR systems the compression characteristics described in this article may be frequency dependent, that is, apply only in some parts of the spectrum and not in others. In Dolby B-type NR, for example, bilinear compression occurs at higher frequencies, somewhat less at middle frequencies, and no compression at all at lower frequencies.

Part VI: Realizing the Bilinear Compression Characteristic

In the previous article (Part V) in this series, I derived the bilinear compression characteristic so highly desirable for a noise reduction system (Figure 1). This time I shall start by showing a convenient way of realizing this characteristic.

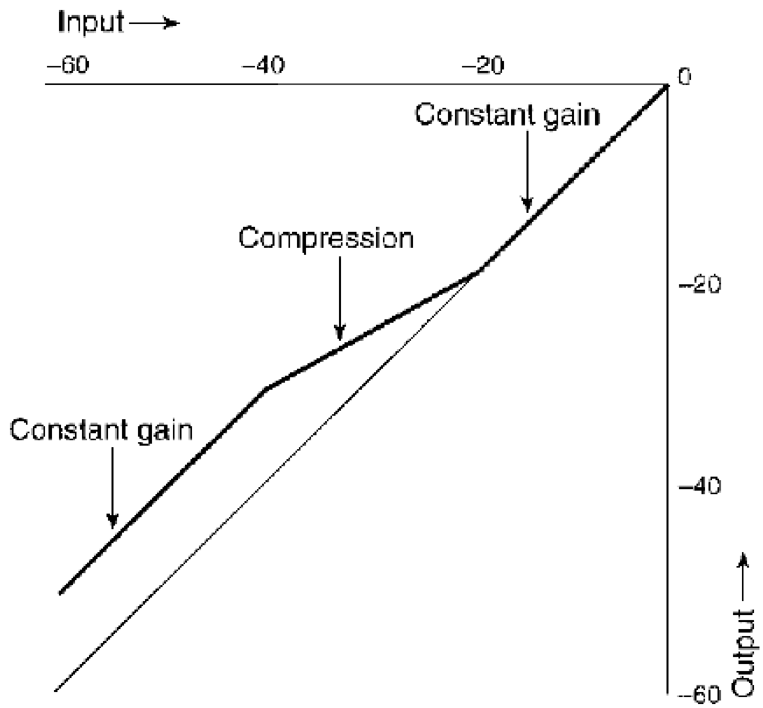


Figure 1: The bilinear compression characteristic.

The two-path approach

The first thing to note about the bilinear characteristic is that the compression turns into unity gain at high levels. We can bring this about in a noise reduction system by having a straight-through path with unity gain, the "main path" shown in Figure 2a. The desired compression can be provided by adding the output of another parallel path, the "further path." This path contains a compressor or limiter so designed that when input level is high, its output is negligible compared with that of the main path, whose unity gain therefore predominates.

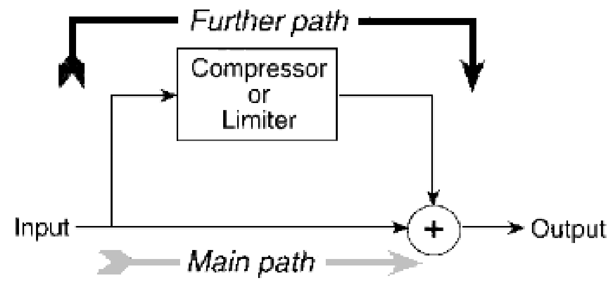


Figure 2A: Two-path compression configuration

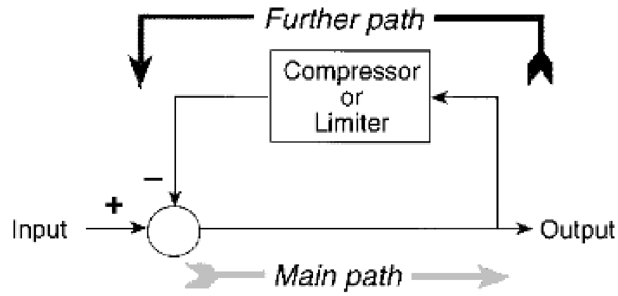


Figure 2B: Two-path expansion configuration

If the maximum gain of the compressor is defined, then the maximum gain of the combination of main and further paths for other signal levels is also defined. For example, looking at Figure 3, suppose that below a known input threshold level ("starting point") the further path has a gain of 6 dB. 6 dB is a voltage gain of 2, so for 1 unit of input, we get 2 units from the further path plus 1 unit of output from the main path, for a total output of 3 units. Thus, below threshold we have an overall voltage gain of 3 (9.5 dB), represented on the compression graph as a straight line at degrees 9.5 dB "north" of the unity gain line.

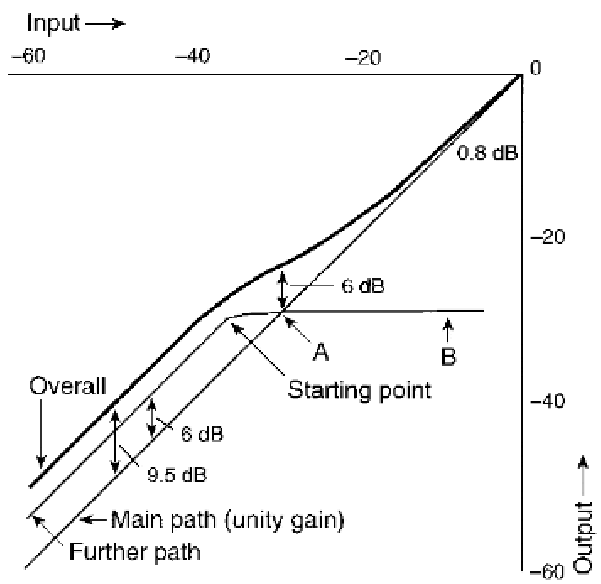


Figure 2: Overall gain for any input level can be determined by adding the gain of the main path to that of the further path.

Once we reach the starting point, the gain of the further path falls with increasing input level. Consider the point labeled A, where the further path gain has fallen from 2 to 1. The overall gain must now be 1 (main path) plus 1 (further path) equals 2 (6 dB). Consider point B, where the further path gain has fallen to 1/10; the overall gain is now 1 + 1/10 = 1.1 (0.8 dB), very nearly the desired high-level condition of unity gain.

In addition to giving us the desired bilinear compression characteristic, this two-path configuration also is easily converted into the complementary expander as shown in Figure 2b. The further path is now fed from the output instead of the input, and is subtracted from the input instead of being added to the output. Note that if the compressor and expander are indeed complementary, the signal conditions in the further path are the same during compression and expansion, a necessary condition for complementarity. By the way, for the rest of this article, whatever I have to say about a bilinear compressor applies equally to its complementary expander.

Implementing a two-path NR system

Clearly if every input level gives a known gain in the further path, we can calculate the resultant overall compression characteristic. Conversely, if we know what overall characteristic we would like, we can calculate the desired shape of the compression of the further path. The precise shape of the final overall compression curve, however, is not very critical. Provided it has the desired gains below the starting point and above the finishing point, does not have an excessive compression ratio between those points, and is reproducible, we can accept whatever shape convenient circuitry will give us.

Take, for example, Dolby A-type NR (Figure 4). Its further-path compressor looks somewhat like a limiter over a range of levels above the low-level threshold. But at still higher levels, it becomes re-entrant; that is, any further increase in input level gives a decrease in output. The resultant overall compression characteristic has a fairly well-defined starting point, but a rather vague finishing point, although the gain does reduce to unity as the input level gets within 15 or 20 dB of the maximum. Most important is that this shape is readily reproducible.

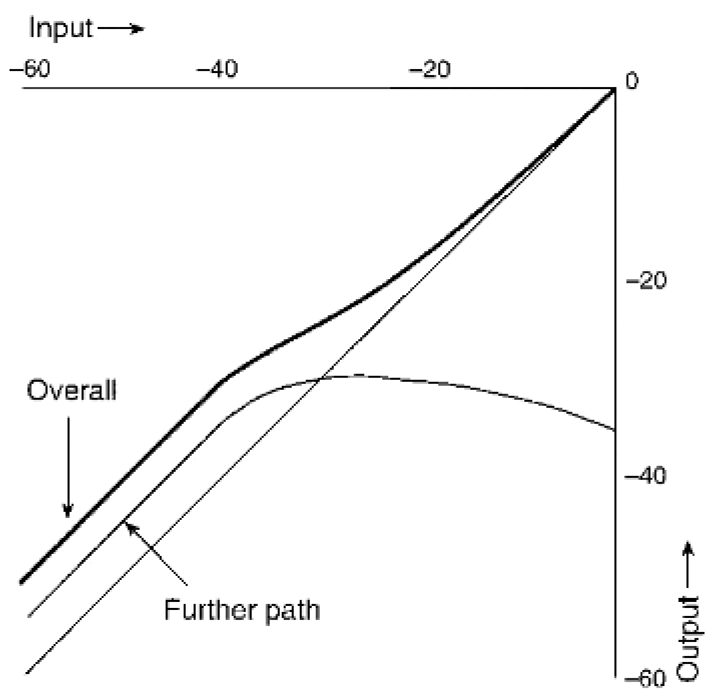


Figure 4: Dolby A-type NR compression characteristic.

One of the virtues of this two-path configuration is that the maximum voltage output of the compressor in the further path is never more than a small fraction of the maximum voltage at the output or input of the processor. This permits the use of simple and relatively inexpensive components. This was vital in 1965 when A-type NR was first introduced; VCAs did not exist, nor did ICs of a quality appropriate for audio. With modern circuitry we can achieve just about any compression shape we might want; but, even today, with a circuit like Dolby SR needing ten "compressors" in its further path, it would be far too expensive for its market if high-quality VCAs were necessary.

Dolby noise reduction at last

Well, there it is–I have finally introduced a Dolby noise reduction system into this series! The time has finally come–next time, to be precise–to show how the theory I have presented so far was applied to the development of A-type NR, our first system. I will also contrast it with other systems, which, although introduced after A-type NR, have not added anything in principle to what was known in the 1950s.

But before ending this installment, I would like to explain the philosophy behind what we do.

Without knowing much about audio or the properties of human hearing, any competent engineer could design a compander which would work satisfactorily most of the time, and you might listen to music for a long time before something came along to reveal the limitations of the system. At Dolby Laboratories, however, it is our philosophy to design systems which are as nearly free of audible side-effects on as much real program material as can be achieved for their cost and complexity. Our experience allows us to predict the weakest aspects of a particular NR system and devise the cruelest tests in order to assess acceptance levels. Extensive listening, therefore, is an important part of the design process.

In designing a practical system there are many factors to be weighed, and a combination of theoretical knowledge and listening tests allows a realistic choice of parameters. Twenty-five years ago, for example, this meant limiting our objective to 10 dB of noise reduction. However attractive, significantly more NR–say 20 dB–was impractical at the time without unacceptable impairment of the original signal. Precisely how we went about achieving that first 10 dB of noise reduction free of artifacts such as noise modulation will be our first order of business next time.

Part VII: Putting It Together

*I promised in Part VI [*Dolby News, *Vol. 4, No. 2*] that the time had come to show how the theory presented in this series so far was applied to Dolby A-type NR, our first system. And so it has. But first, so that you get a sense of the background from which A-type NR emerged, lets look briefly again at what had gone on before (and has since!).*

The simplest NR systems are generally wide-band companders, usually with a fixed 2:1 compression ratio and fixed (level-independent) dynamics. These grew out of early (1930s) work on telephone communications, where side effects such as noise modulation, transient distortion and mild mistracking impair intelligibility very little, and certainly not as much as the unreduced noise. Even ten years ago, transatlantic calls via submarine cables often showed clearly audible artifacts of companders.

As we saw in Part I [*Dolby News*, Vol. 1, No. 1], wide-band companders inherently give a degree of noise reduction which varies from moment to moment with the signal level. The gain changes in the expander occur uniformly across the whole audio spectrum, so if the signal demands *x* dB increase in expander gain, there will be precisely *x* dB increase in wide-band noise. There are bound to be times when the signal does not mask the changes in noise, so audible noise modulation is inevitable. This is usually most apparent when the signal is confined to the low end of the spectrum, revealing noise modulation at the high end.

In an attempt to reduce this problem, most compander systems use fairly aggressive pre-emphasis, boosting high frequencies in the compressor (and the inverse in the expander), with additional frequency shaping in the control path to reduce the likelihood of high-frequency overload. Such pre-emphasis does indeed reduce the audibility of high-frequency noise modulation on low-frequency signals. But you don't get something for nothing: susceptibility to *low-frequency* noise modulation in response to high-frequency signals is now greatly increased.

As described in Part III [*Dolby News*, Vol. 2 , No. 2], constant-slope compressors (those with level-independent compression ratios) yield overshoots whenever the signal level increases abruptly. These overshoots can exceed the overload point of the noisy medium, leading to transient distortion and mistracking in the expander, which is no longer receiving the correct information to reconstruct the signal.

One of the virtues claimed for constant-slope systems is that there is no need for "calibration," that is, alignment so that the level entering the expander is the same as that leaving the compressor. However, putting this the other way round, a system which claims not to need calibration *must* be a constant slope device–and therefore *must* give high-level overshoots!

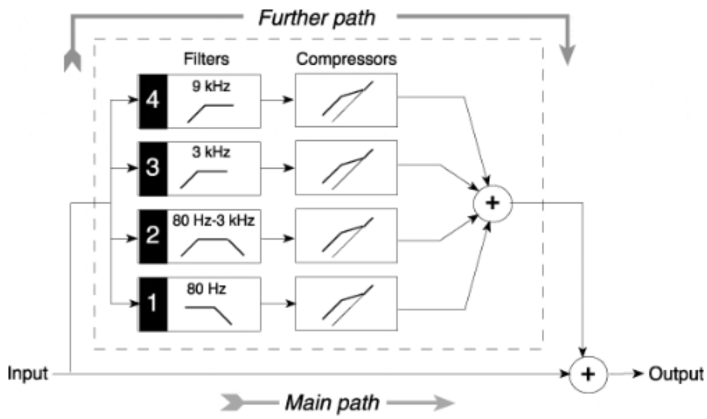
That is quite enough about these elementary systems for the present; they are still in widespread (and audible!) use in the so-called "hi-fi" sound of consumer video recorders and in analog network radio distribution via satellite.

Dolby A-type noise reduction

As I discussed in Part II [*Dolby News*, Vol. 1, No, 2], to be free from noise modulation a system must provide constant NR *wherever in the spectrum the signal does not provide masking*. The first system designed with this aim was Dolby A-type NR, introduced to a suspicious market in 1965. The market was suspicious because over the preceding decade or so several noise reduction systems had been introduced which audibly impaired the music, always with noise modulation, and sometimes with transient distortion as well.

In A-type NR, the audio spectrum is split up into four bands, each subject to independent bilinear compression/expansion, using of course the two-path configuration described in Part VI. The block diagram shows the compressor (encoder); the complementary decoder can be derived (as also described in Part VI) by placing the further path in the feedback with subtraction from, instead of addition to, the main path.

The overall low-level gain/loss in each band is a fixed 10 dB, therefore giving 10 dB of noise reduction. In the presence of signal in a band above a predetermined threshold (the starting point), the noise reduction in that band is reduced. But the signal and rising noise are not far apart in frequency, so the rise is masked. Moreover, because the compression characteristic in each band has that starting point, rather than extending down indefinitely as it would in a constant-slope device, signals in other bands have to be at very high levels before they affect the gain. Thus the degree of NR in one band is substantially independent of the signal in other bands–precisely the definition of a system free from audible noise modulation. Finally, because the compression characteristic is bilinear with a gain of 1 over the top 20 dB or so of dynamic range, very few high-level overshoots can occur.



Dolby A-type compressor

Calibration

A bilinear compression characteristic implies a need for level alignment to ensure that the expansion curve lines up with the compression. To help in this calibration procedure, we define a reference point on the transfer curves, called Dolby level. In the compressor of A-type NR, a tone generator delivers a tone at this reference level which is frequency modulated to be easily identifiable by ear as "Dolby tone." In the expander you adjust the input level until a level indicator (a meter or a string of light-emitting diodes) shows that the received Dolby tone is at the required level.

The reference point could have been anywhere on the compression curve. For example, we might have chosen the level at which the compressor introduces 6 dB of gain (and the expander 6 dB of loss). However, that would be a comparatively low level, and would not register satisfactorily on program level meters (VU or PPM). Instead, Ray Dolby chose a level in the same region as normal line-up tone, where the gain/loss of the process is very small. During manufacture, the checks on compression curves are performed with test signals at levels referred to this Dolby level.

Incidentally, since the overall gain of the compressor for high-level signals is unity, introducing A-type NR doesn't require any change in operating or maximum program level. Furthermore, the degree of NR, 10 dB, is determined by the shape of the bilinear compression characteristics (how far apart the starting and finishing points are), not by the signal levels within the NR circuits, nor by the placement of the Dolby reference level with respect to program levels. This contrasts with the behavior of constant-slope devices, whose NR effect depends critically on the program level.

You may well ask, why 10 dB of noise reduction, or why not 15 or 20 dB? The answer is that 10 dB was possible without audible side-effects at a price the market could bear, and also was enough to improve 15 ips stereo recording from the dominant noise source to a rarely significant one. More NR with the same degree of immunity to side-effects and using the same techniques would have required many more bands–and a much higher price.

Although its main application has been in professional analog tape recording, A-type NR was designed as a general purpose system, equally useful for land-lines or any audio medium flawed by low-level broad spectrum noise. I shall probably return to some of the subtleties of A-type later, but those wanting to know now should consult Ray Dolby's classic paper, "An Audio Noise Reduction System" (*JAES*, October 1967), which is available from Dolby Laboratories.

The next step

The early years of A-type NR coincided with the spread of the first convenient audio recording system for use in the home, the Philips compact cassette, whose performance left much to be desired (open-reel recording was confined to the technically minded). Tests showed that many of the mechanical problems of tape running at the slow speed of 1-7/8 inches per second were soluble, and that the frequency response could be extended by appropriate design of tape heads and associated circuitry. However, these improvements were not worthwhile as long as the signal to noise ratio was in the region of 50 dB, markedly inferior to that of a good LP pressing. Extending the response from the 6 or 7 kHz of the original recorders merely made the noise more disturbing, particularly when the tracks were split into two for stereo.

Some in the audio industry (but probably not those actually working in the field of tape development) imagined that improvements in magnetic materials would lead to huge improvements in S/N ratio without the need for electronic processing. In practice, the improvements in all the years since have amounted only to 5 or 6 dB (tapes have improved far more in their mechanical properties and consistency than in noise). Therefore it was apparent to those in the know that only with a practical, inexpensive NR system might the cassette take off in a big way.

While applying A-type NR to the cassette gave encouraging results, its complexity made it far too expensive for consumer products. New ways had to be found to apply the general principles by now established (and which I have been writing about all this time!). Simple calculations suggested that using the components then available, in particular all discrete transistors, a processor could use only one band; any more would be too expensive.

A-type NR had been designed for a variety of media, most with broad-spectrum noise. Cassette noise, however, is of a known spectrum which is subjectively weighted heavily towards the high frequencies. Thus as a first step we experimented with using A-type NR with only one high-frequency band in operation. The tests were not promising, however. With a band similar to A-type band 3, 3 kHz and upwards, the system worked without audible ill-effects, but the degree of noise reduction was disappointing. If the band was made wider, say from 1 kHz upwards, the noise reduction effect was greater, but noise modulation was audible: now the signal within the band did not necessarily mask the change in noise in the band. In other words, widening the band caused the system to behave more like a wide-band compander, with its associated and inevitable noise modulation.

Next time I will explore the solution to this problem. If you suspect it might have something to do with the next letter in the alphabet...you're right.

Part VIII: The Sliding Band

Towards the end of Part VII [Dolby News, Vol. 5, No. 1, Spring 1994], in which I discussed primarily the first Dolby noise reduction system, A-type, I introduced the topic of the original Philips compact cassette. I explained that its potential was limited by a signal-to-noise ratio markedly inferior to that of the best LP record pressings of the day, and that cassette noise, unlike that of professional open-reel tape, is pre-dominantly at high frequencies. While this characteristic pointed the way towards an economical consumer NR system using only one band, we found that a single, fixed high-frequency band did not give satisfactory results after all.

Single-band noise reduction: the early experiments

With a fixed NR band from, say, 3 kHz upwards (as in A-type's band 3), signals below 3 kHz don't modulate noise above 3 kHz. In addition, while signals higher than 3 kHz and above the low-level threshold (the "starting point") *do* affect the noise above 3 kHz, we found that at the same time they mask that effect. The result was effective hf noise reduction with essentially no audible noise modulation. However, obtrusive noise in the 1 to 3 kHz region still remained on low-level signals, or in the absence of signals.

We next tried extending the NR band down to 1 kHz, which greatly improved the apparent degree of noise reduction under low-level signal conditions. Equally important, what noise remained had an audibly smooth spectrum, thereby minimizing its subjective impact. Unfortunately, with this 1 kHz NR band, above-threshold signals in the 1 kHz region modulated all the noise above 1 kHz–disturbingly so in the 6-8 kHz region.

Figure 1 illustrates the behavior of such an expander, one offering 10 dB of NR from nominally 1 kHz upwards, in response to an input sine wave at levels ranging from well below threshold to well above. Clearly, the gain it provides at 10 kHz, and hence its NR effect in that region, is a function of the level of the sine wave. If that sine wave is at 1 kHz, it provides little masking at 10 kHz–the perfect recipe for audible noise modulation.

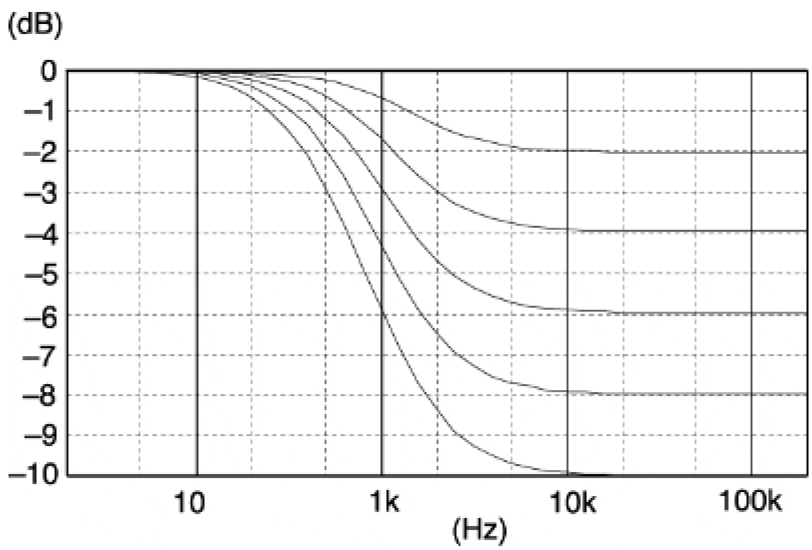


Figure 1 Response of fixed band decoder to increasing level of input signal.

So the question was, how could we provide NR in a band extending upwards from, say, 1 kHz without signals at the lower end of the band modulating noise at the upper end?

Sliding band noise reduction: what finally worked

The answer was right there in the experiments I have just described, wherein we tried fixed bands starting at different frequencies. The 1 kHz band was satisfactory as long as signals above threshold were not present in the region of 1 kHz. When above-threshold signals *were* present in that region, the 3 kHz band operated satisfactorily (it didn't provide NR in the 1 kHz region, but any noise there was effectively masked by the signals them-selves).

In other words, what was needed was indeed a single band system–but one which could *adapt to changes in the spectral content of the signal*. With low-level signals, this signal-adapting band would offer satisfactory NR at 1 kHz and above. But when high-level (i.e. above-threshold) signals came along in the 1 kHz region, instead of allowing them to modulate hf noise, the band-determining filter would slide up out of their way, leaving them unaffected and thereby maintaining constant hf hiss reduction.

A picture here is worth lots of words. Figure 2 shows a family of response curves for what we now had come to call a *sliding band* expander or decoder. It consists of a high-frequency downward shelf of substantially constant shape and height (10 dB is this example) which can move (slide) along the frequency axis. Clearly this response will provide high-frequency noise reduction, reaching down to where the cut-off frequency lies at the moment (i.e. which curve of the family shown applies at the moment).

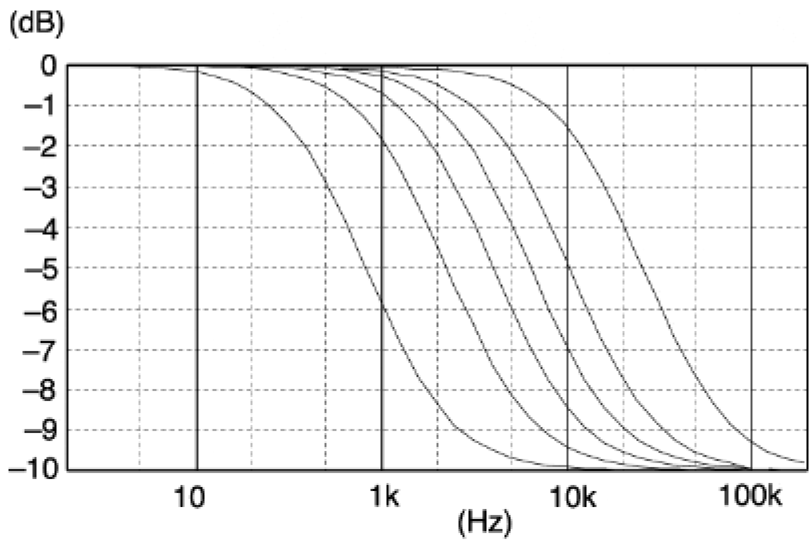


Figure 2 Response of sliding band decoder to increasing level of input signal.

In the presence of high-level signals, control circuitry moves the filter so that those signals fall in the unity-gain part of the curve, where little or no attenuation occurs. Off course, no NR is obtained in this region as a result, but we don't need it; the high-level signals themselves mask the noise. More importantly, notice that to accommodate a high-level 1 kHz signal the band has to slide up out of its way (the third curve in Figure 2), providing only, say, 1 dB attenuation at 1 kHz . This means that the degree of NR in the 10 kHz region changes only by a very small amount. In other words, we have essentially eliminated noise modulation.

The NR *en*coder, of course, has to have reciprocal characteristics: a 10 dB shelf upwards, with the same sliding behavior (Figure 3). This has an important further advantage: a high-level signal entering the encoder causes the band to slide up until there is little or no gain applied, preventing any further boosting of those signals that could cause tape overload.

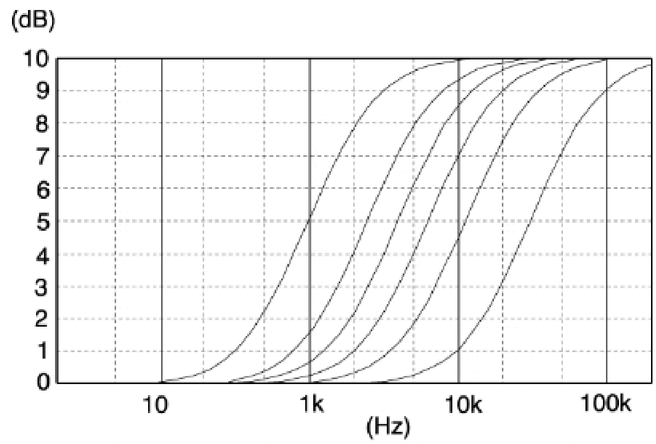


Figure 3 Response of sliding band decoder.

Now you might well say that what I'm describing is really variable pre- and de-emphasis, not compression and expansion. Well, it certainly is variable emphasis, but consider the behavior of an encoder when fed with, for example, a 5 kHz sinewave at ever increasing level (ignore what happens to the gain at other frequencies, and just take a "cross-section" of Figure 3 at 5 kHz, indicated by the dotted line). Below threshold, the input signal is boosted 10 dB. Above a predetermined threshold, the filter slides upwards and the gain at 5 kHz falls progressively. When the filter has slid far enough, the gain at 5 kHz falls to unity. The plot of output level versus input level looks like Figure 4. Does this look familiar? It should to those of you who have been following this series; it is the classic, highly desirable bi-linear compression curve I described at considerable length in Part VI [Dolby News, Vol. 4, No. 2, Fall 1993]. Hence, with respect to any one frequency, a sliding-band NR encoder *is* a compressor, and a sliding-band decoder an expander.

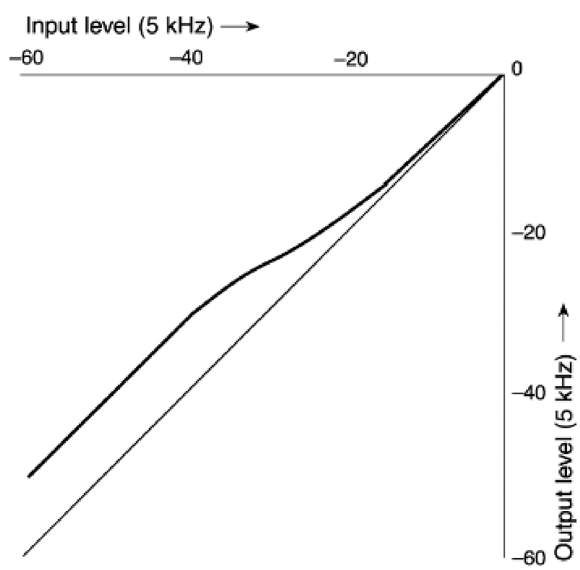
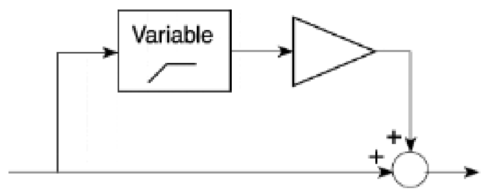


Figure 4 Output vs. Input level.

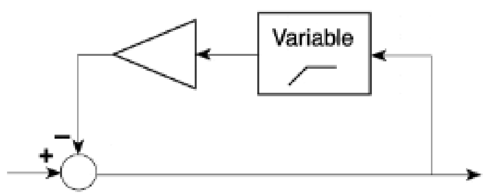
The birth of Dolby B-type

As you will have guessed by now, what I have been describing is well-known and ubiquitous Dolby B-type noise reduction, a sliding-band system offering 10 dB of NR from around 1 kHz upwards. It incorporates all the features discussed in previous parts of this series, including the dual-path configuration.

You might expect that sliding a shelving response along the frequency axis would require moving two filter frequencies, and so two voltage-controlled amplifiers. However, as shown in Figure 5, we put a variable high-pass filter in the further path which, in the tradition of A-type, is summed with a unity-gain main path. The overall response is therefore inherently a shelf of constant shape and height, whose upper break-point is the frequency of the variable high-pass filter, and whose lower break-point is below the upper one by whatever it takes to give the desired height of shelf. Thus in practice only one voltage-controlled device is necessary, which in the early days, before custom linear ICs were economic, was a junction field-effect transistor used as a voltage-controlled resistor.



Sliding-band encoder



Sliding-band decoder

Figure 5 Dolby B-type noise reduction block diagram.

What next?

For many years, throughout the world, fixed-multiple-band Dolby noise reduction (A-type) faithfully served audio professionals, while sliding-single-band Dolby noise reduction (first B-type, then C-type) dealt effectively with the high-frequency hiss on consumer formats such as the audio cassette. Then in the early 1980s, having re-examined the pros and cons of fixed- and sliding-band noise reduction, Ray Dolby devised a way of combining the best of both approaches in the same powerful system, Dolby SR (spectral recording). Recognizing that not everyone would want or be able to convert to the digital systems then coming into their own, SR was introduced in 1986 for professional analog applications, followed a few years later by S-type noise reduction for the audio cassette. These synergistic combinations of our two NR technologies will be the subject of a future article in this series.

Addendum

In the first section of Part VII, the previous episode of this continuing series on noise reduction [Dolby News, Vol. 5, No. 1, Spring 1994], I referred to the use of aggressive pre-emphasis and de-emphasis in wide-band companders to reduce the audibility of the modulation of high-frequency noise by low-frequency signals. I omitted to point out that while emphasis reduces the absolute level of the noise, it does not alter the degree by which it is modulated. If the noise in the presence of a particular low-frequency signal rose 10 dB without emphasis, it still rises by 10 dB with it (even though the level from which it rises may be lower). Thus the emphasis doesn't solve the noise modulation problem inherent in wide-band companders, and indeed it introduces other problems.

Part IX: Action Substitution

In Parts VII and VIII of this continuing series I discussed noise reduction systems employing fixed and sliding bands. A fixed-band system applies variable gain within a fixed, predetermined portion of the audio spectrum, while a sliding-band system applies a fixed gain/loss within a variable portion.

Each approach has its virtues and vices. For simplicity, the following discussion treats encoders only, and assumes corresponding decoders with reciprocal characteristics. It also addresses high-frequency bands, but the principles can be extended to low- or middle- frequency bands. Of course I am discussing Dolby systems, and hence bilinear transfer functions (part V) and dual-path configurations (part VI) are assumed.

Fixed band

Figure 1a shows the response of a fixed high-frequency band like that in A-type NR. Below threshold, as shown by the solid line, the full effect accrues, in this case 10 dB of gain (and thus 10 dB of noise reduction upon decoding). The dashed line shows what happens in the presence of a dominant input signal within the band that is strong enough to demand 2 dB less train. The gain is reduced by 2 dB not only at the frequency of the dominant signal as intended, but across the *entire* high-frequency band as well. During decoding, therefore, the noise reduction effect across the band is only 8 dB, rather than the full 10 dB; put another way, the (reduced) noise rises 2 dB. At least the degree of gain reduction is the same both below and above the dominant signal, which as we shall see is an advantage.

Sliding band

Figure 1b shows the response of a sliding high-frequency band like that of B-type NR. Below threshold (solid line) it responds just like the fixed band, and in the presence of the dominant signal (dashed line), the gain at its frequency is reduced by 2 dB as with the fixed band. The gains at other frequencies within the band, however, are reduced by differing amounts. Below the dominant signal the gain reduction is more than 2 dB, while above it the gain reduction is less than 2 dB. The complementary decoder, therefore, gives less than 8 dB noise reduction below that frequency, but more than 8 above it. In other words, the (reduced) noise level at the decoder output will rise more than 2 dB at frequencies below the dominant signal, but less than 2 dB at those above it.

Thus in the presence of signals, the sliding band is better than the fixed band at reducing noise above the frequency of a dominant signal – but worse than the fixed band *below* that frequency.

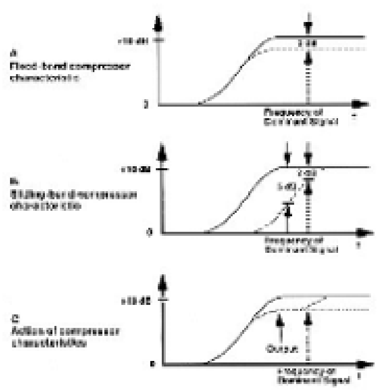


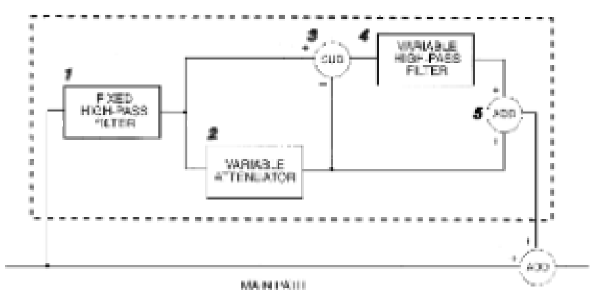
Figure 1
Types of compressor characteristics

Action substitution
Appropriately enough, it was Ray Dolby who discovered how to realize the advantages of fixed and sliding bands while avoiding their disadvantages. He found that fixed and sliding bands covering the same region of the spectrum could be combined to provide what he called "action substitution." While "action" simply refers to gain or loss that can change, as opposed to passive and fixed filtering or amplification, 44 substitution" takes some explanation.

Consider a fixed band and a sliding band arranged so that the effective gain of the pair, and ultimately the NR, is that of whichever band would provide the greater gain on its own (Figure 1c). At the frequency of the dominant signal we still get 2 dB of gain reduction (8 dB of NR). But below that frequency we get 8 dB of NR as with a pure fixed band, while above it we get increasing NR up to the full 10 dB like a pure sliding band. Below threshold, of course, the gain of the combination is the same as that of each of the bands individually; we can think of the sliding band hiding under the fixed one, or vice versa.

When the sliding band moves up to provide 2 dB gain reduction at the frequency of the dominant signal, the presence of the fixed band is no longer hidden, and it steps in or *substitutes* for the sliding band to restore some of the NR lost as the latter slides up. By the same token, as the fixed band reduces the gain by 2 dB across the whole band, the presence of the sliding band is revealed; it then substitutes for the fixed band to restore some of the NR lost above the frequency of the dominant signal.

In other words, at each frequency there are now *two* mechanisms which can provide gain, one standing in for the other as necessary to keep the gain constantly high. Now you may remember that the ideal NR system would apply gain (and loss upon decoding) that changes only at frequencies where the level is high enough to provide masking. This is what Ray Dolby calls the principle of least treatment (see [part II]() of this series)–and clearly a pair of fixed and sliding bands operating with action substitution can better approximate the ideal than either type alone.



[Figure 2]()
Dual-path compressor employing action substitution

Practical realization
How can action substitution be realized? One way is to perform one action (e.g. fixed-band compression) and then perform the other (e.g. sliding filtering) on what is taken away by the first. Figure 2 shows a dual-path compressor configured along these lines. The main path across the bottom has no processing but is summed with the further path containing both gain change (variable attenuation), and a sliding high-pass filter. For simplicity, control circuitry is not shown.

In this example, the further path starts with a fixed high-pass filter (1) to determine a high-frequency band. This feeds a variable attenuator (2), giving us a fixed band. The output of this fixed band is subtracted from the attenuator input. When there is no attenuation, this subtraction yields no signal. But when there is attenuation (i.e. a signal within the band is above the threshold), the output of the subtractor is the signal that has been "removed" by the first action's attenuation. This subtracted signal is then fed to a variable sliding filter (4), the second action, the result being added to the fixed band output (5) to form the output of the further path.

The variable attenuator and sliding filter can be interchanged with exactly the same results; the order shown is preferable because it demands less of the VCAs or FETs used to implement the variable elements.

Real systems

Action substitution is one of the key, inventions that made Dolby SR (professional) and the more recent S-type noise reduction (consumer) possible. However, even the most flexible adjustment of frequency response is useless if you cannot control it appropriately. Therefore next time we will consider another key development, modulation control.