VORSIS Application Note

AN2008-05

The Magic Behind the Vorsis 31 Band Algorithm

March 2008 - Jeff Keith

600 Industrial Drive New Bern, NC 28562 USA 252-638-7000

www.vorsis.com
techsupport@wheatstone.com

Background

When it comes to discussions about multiband audio processors it quickly becomes apparent how many opinions there are about the right and wrong ways of designing one. The level of passion in some discussions can even rival that when the subject is politics.

When the discussion of 'high band count' audio processors comes up, those like me who are old enough recall the crude implementations that preceded Digital Signal Processing (DSP), instantly remember their unnatural swishy and phasey artifacts. The truth is there is little reason for artifacts like those in an audio processor today *when the crossover has been implemented correctly*.

Why Multiband?

The primary reason for multiband audio processing is to electrically separate different frequency program elements from each other so that, for instance, low frequencies cannot cause inappropriate gain modulation of the higher frequencies.

Let's use the example of the single band (broadband) limiter: when the limiter attacks an energetic bass beat in music the mid and high frequencies go along for the ride resulting in audible 'ducking' of the higher frequencies as the limiter reduces the gain for the lower frequencies. A term was coined for this effect - "spectral gain intermodulation".

If the limiter's time constants are made faster to try to minimize the ducking effect, a new problem emerges – the bass waveform itself modulates the high frequencies. This results in 'intermodulation distortion'… our ears tend to be fairly tolerant of harmonic distortion, but unfortunately the same does not seem to hold true for our tolerance of intermodulation distortion.

Depending on the application, multiband processors having two, three, four, or more bands can efficiently work on the audio spectrum with fewer audible side effects than a broadband device ever could. This is especially true if one of the goals is to increase the average power level (loudness) of the audio passing through the audio processor.

A Look into Multiband's Past

While there were likely broadcast-related multiband audio processing that existed earlier, the first one that *this* author became aware of was back in 1974. As a 21-year old chief engineer at a small market AM/FM I would frequently listen to CKLW on my way home from work. I was fascinated by their air sound – it was loud and clean and completely free of all the artifacts that I had been wrestling with while trying to make our station more 'competitive'.

One afternoon I called CKLW and asked the switchboard operator who their chief engineer was and could I please speak to him. Not only did Ed Buterbaugh accept a call from that young and green engineer, he took the time to speak to him for almost a half hour to talk about radio things in general and explain how he got that "Big 8" sound. (Thank you Ed! You have no idea what you started!)

Over three decades later I still remember Ed's analogy: "You know how bi-amplified speaker systems work?" he asked. "Think what would happen if you put a crossover network in front of a couple of limiters".

The light bulb in my brain instantly went to full brightness… a crossover! Of course! Why didn't I think of that? Since I was already hopelessly hooked on audio processing the rest as they say, is history.

Soon I was building audio crossovers and filling up the transmitter room racks with spare limiters and compressors to see what kind of sound I could put on the dial. Was it crude? You bet! Did it always sound good? Heck no. But the competition across town that was still using the popular CBS Audimax and Volumax combination didn't have a chance.

A year or so later I became aware of Mike Dorrough's three-band DAP310 and we bought one. Mike's genius brought "all-in-one-box" multiband audio processing to the broadcasting marketplace.

The next challenge was figuring out how to make our old plate modulated AM transmitter more faithfully follow the energetic waveforms coming out of our new audio processor, but that's a subject for another day…

How Many Bands?

It's been one of the biggest questions to haunt mankind since the dinosaurs disappeared – how many bands is enough (or too many) in a multiband audio processor? Four? Six? Forty six?

The answer lies mainly in the goal one is trying to achieve and generally the best answer is "as few bands as possible in order to get the job done". But what job? And done how well? And exactly how *is* "done" defined? Maybe it depends on what the definition of "*is*", is...

In <u>very</u> general terms, fewer frequency bands works best for slower functions like compression and leveling when gain reduction greater than 6dB to 10dB per band is required.

As time constants get faster though, the depth of gain reduction that can be used without generating audible and unnatural side effects rapidly decreases. Turning this thought around implies that *more* bands can be used *if* the time constants are faster *and* gain reduction is less deep.

In fact, this is exactly what happens – as the audio spectrum is broken down into more frequency bands, there is energy within each band. Less energy in the bands means that less gain control depth is required in order to control the energy in the bands and that means that faster time constants can be used without generating noticeable (or objectionable) artifacts.

When the number of bands becomes high enough the resulting audio bandwidth becomes so narrow that a point is reached where the human ear can no longer accurately detect what is happening within one band when energy outside that band is *also* present. This bandwidth is typically referred to as the 'critical bandwidth'.

In other words, when the bands are narrow enough, program material that is residing at frequencies *outside* of a band where work is being done helps 'mask' that work. This is the underlying principle behind many perceptual codecs such as MPEG-3 (MP3). Such algorithms hide what they are doing to the audio inside what are called "critical bands" by using program material at nearby frequencies as 'maskers' for what is going on inside the band that is doing the work.

Barking… sort of

The approach taken to create our 31-band limiter is loosely related to the Bark Scale and we've created a filter bank with enough bands to be able to distribute the dynamics control work needing to be done in a way that makes it less audible to the ear. Compared to other, less complex approaches, the work needing to be done in each band of simpler five and six band limiters is by comparison, *quite* large.

Our technique enables dynamics work such as limiting and clipping to be better hidden "acoustically" from the ear because:

- The amount of audio energy per band is less and therefore the work needed to be done per band is necessarily less, and;

- Program material residing in adjacent bands that are not undergoing limiting or other modification helps mask the work being done in bands that *are* doing work.

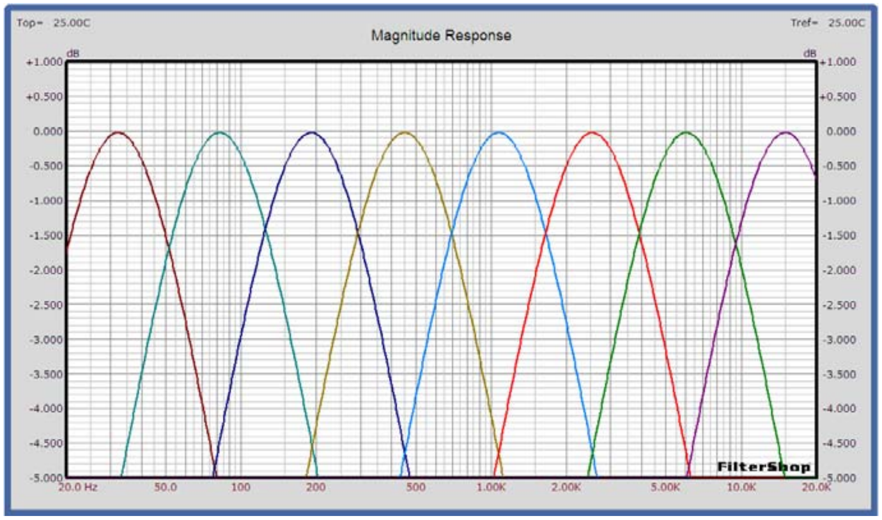
In fact, for most program material the work done by the 31-band algorithm is not audible to the ear as limiting and/or clipping. Instead, because the average energy within a band is being increased the ear perceives a loudness increase instead of the distortion that may have been related to creating that loudness increase.

Comparing Apples and Grapes

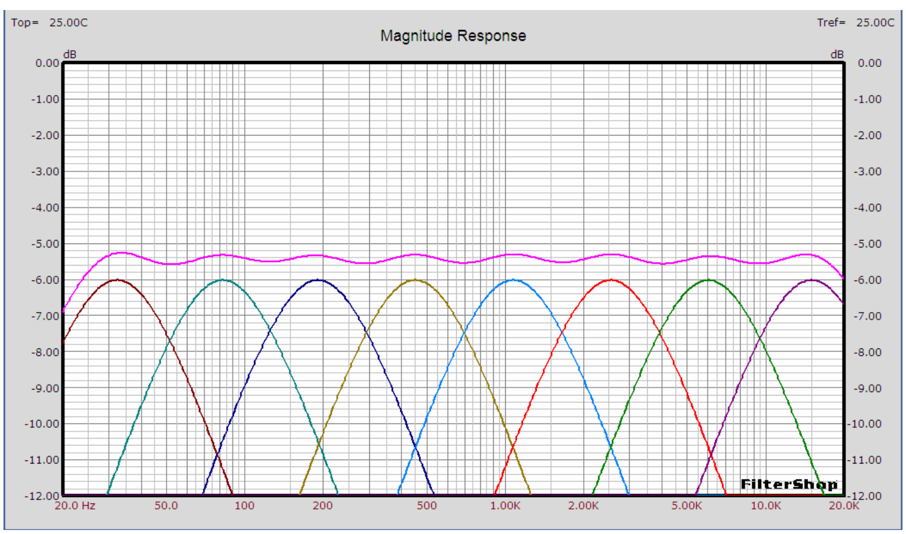
We will be the first to admit that audio processors on the market using a smaller number of bands have an advantage over our approach, and that advantage is that they require a much less complex crossover network. They also require far fewer level detectors, multipliers, time constant computations, DSP cycles, and they *certainly* have fewer user controls. But that's where their advantage ends!

The biggest reason why our approach sounds so much better on the air is that when one of our limiter bands is in limiting, program material that is present in adjacent bands and below their limit thresholds is not arbitrarily reduced in level. The program content that *is* inside the band that is in limiting is so close to the frequency causing the limiting that it falls beneath the masking curve and would likely not be heard by the human ear anyway even if the band were *not* in limiting!

Just for fun, let's analyze an eight band algorithm to see how it might behave with real program material - for ease of plotting we'll use a single sine wave tone stimulus. Below is a graphic plot of how the eight individual band's responses might appear.

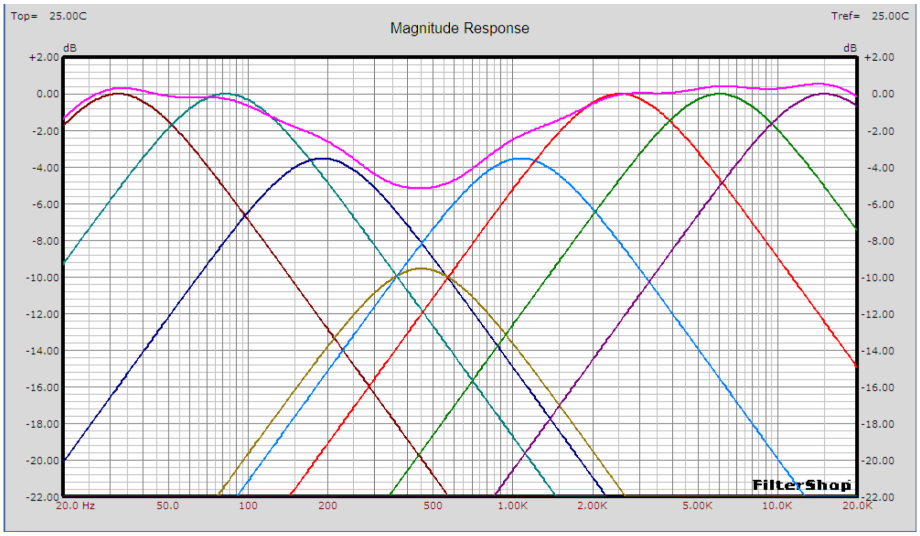


Below is a plot of the recombined response of those eight bands when there is no gain reduction occurring due to limiting.



The recombined frequency response (shown by the top-most wavy line) is not perfectly flat but it's still pretty good. But this is the least of the problem!

Next, let's look at the graphic below depicting the recombined frequency response after band four has been tasked to do several decibels of limiting at its 440Hz center frequency.



Allow us clarify a few details about the above graphic:

1. Band four is the band that's been asked to do several decibels of limiting.

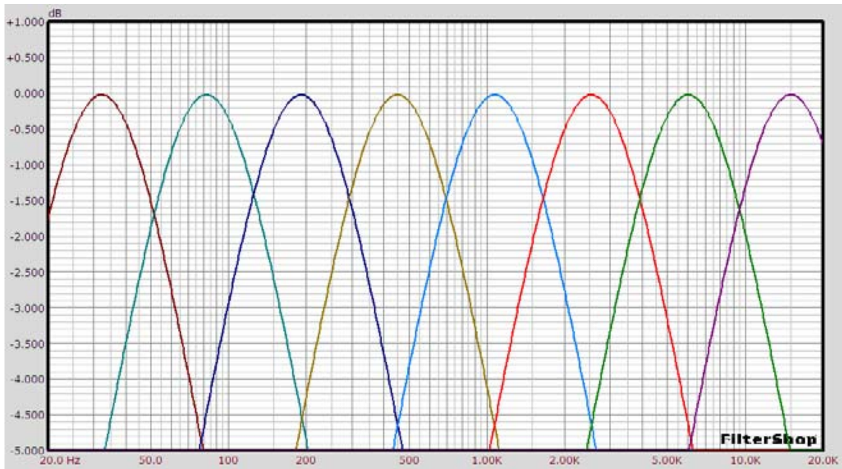
2. Crossover leakage is causing bands three and five to *also* reduce *their* gains by several dB.

3. The response at the center of band four is nearly 5dB down in response to commands from the limiter control circuitry.

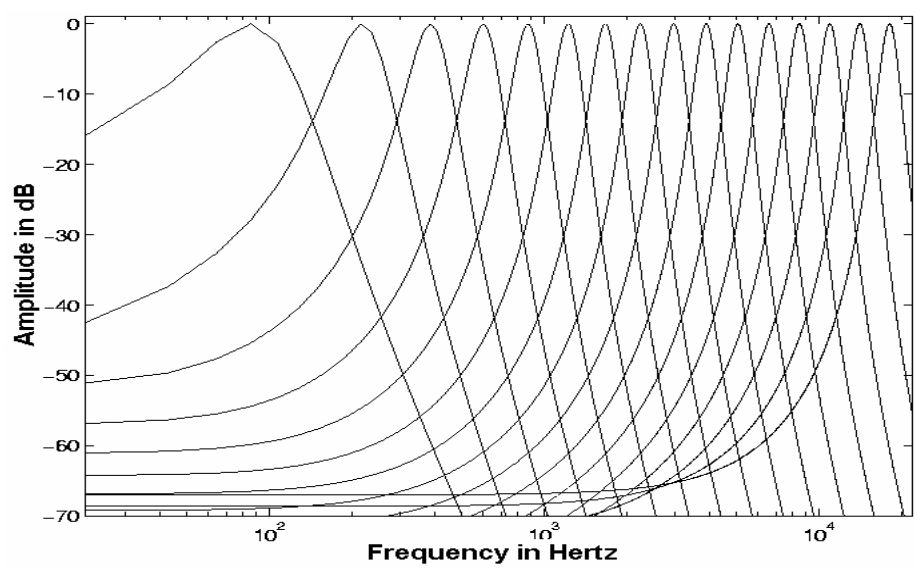
4. Even though there is little of the band four signal in bands three and five, they are *still* unnecessarily reducing their gains! In fact, the response at 220Hz and 880Hz (half of and twice the center frequency of band four) is down about 3dB! This is half power, or equivalent to <u>half loudness</u> at those frequencies.

Question: With this kind of behavior, what happens to program content that is in bands three and five? It's been rendered either completely inaudible (because of "masking" from bands two and six) or has been reduced in amplitude (to half loudness!) to where it's probably no longer even noticed by some listeners!

Let's look back one more time at the plot of the eight band limiter's filter bank.

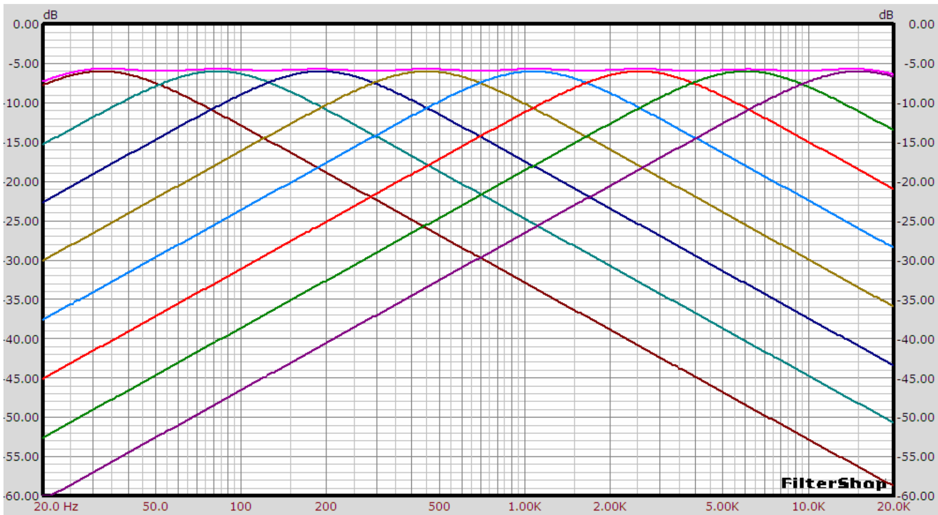


Compare this to the plot below – which is an example of the type of filter bank used in the AP2000 (limited here to just sixteen bands to keep the graphic image clearer).



There are some important things to notice in the two graphics. One is that the resolution, or ability to discern individual frequencies in program material, is a few orders of magnitude better in the Vorsis design.

The second thing to notice is the <u>difference in vertical scales</u> – we only plotted the upper 6dB of the eight band crossover response while we plotted 70dB for the 16 band Vorsis example. To get an idea of how much inter-band ‘crosstalk’ there is in the simplistic crossover used in the eight band example, see the plot below which has now been rescaled to show about the same dynamic range as the Vorsis plot.



The most important thing to recognize is how <u>broad</u> the crossovers actually are in the eight band example – and the typical four or five band crossover *could* be even broader depending on its design.

Of course a five or six band crossover can be made from higher order filters than we’ve used in our example. And while doing so will reduce the amount of inter-band interaction we’ve just been discussing, it won’t nearly be reduced to the level of our Vorsis implementation.

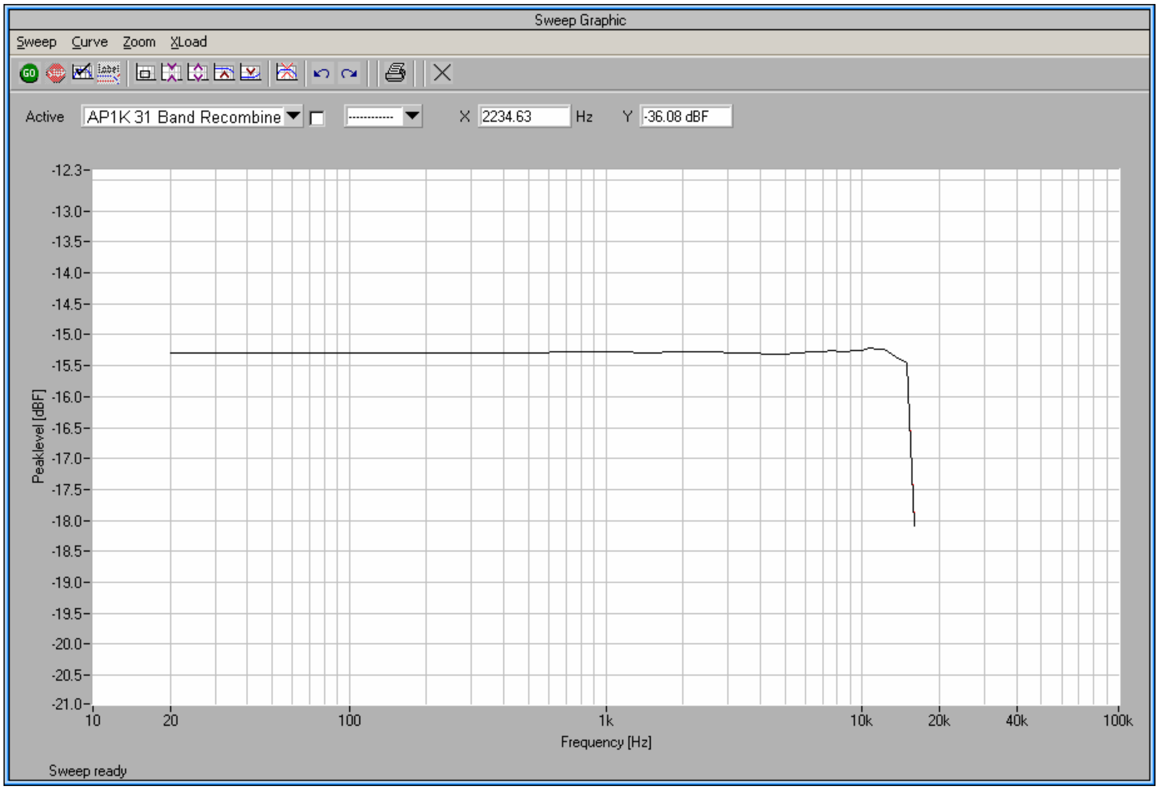
While such broad crossovers may work perfectly fine for a slower time constant AGC or compressor (which is how we use them in the some Vorsis products), in order to operate with fewer audible side effects fast operating limiters require much narrower and more selective crossovers in order for their work to be hidden from the human ear.

Research in the Vorsis lab has revealed that if the bands are numerous enough and narrow enough (and behave well during recombination) we can hide *nearly all* of the dynamics errors and distortion artifacts generated by the processing below the threshold of human hearing.

In summary, the benefit of limiters working in so many bands is an increase in audio clarity accompanied by an increase in perceived loudness accompanied by at least partial masking of any processing artifacts.

31 Band Recombination

The implementation we used in the AP2000 was painstakingly tuned to provide near perfect recombining both below the threshold of limiting and during limiting by one or more bands. The response plot shown below is a plot graph of recombined frequency response of our 31 band section when swept with sine wave tone below the limit threshold.



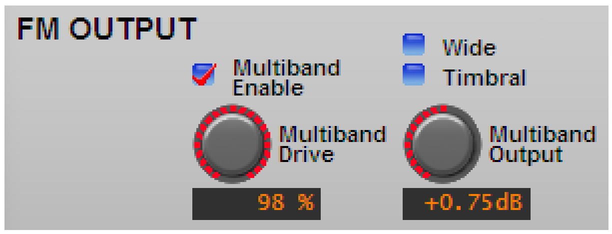
As can be seen from the plot, other than some expected amplitude ripple just prior to the abrupt cutoff of the 15 kHz FM low pass filter the recombined frequency response is virtually flat. And like the frequency response, the phase linearity is also flat with no abrupt departures from linear phase.

In order to achieve such perfect recombination under all possible program conditions we found it necessary to do away with “Band Output Mix” controls for each of the 31 band limiters. This choice has proven correct in actual field use since the resolution of the 31 band section is so high that individual control of each band’s output is simply unnecessary.

31 Band Operating Modes

Neglecting for a moment the usual adjustment related to limiting and clipping functions, there are three special operating modes for the 31 band limiter section; ***Discrete*** (the default), ***Wide***, and ***Timbral***.

Note: ***Discrete Mode*** is enabled when both ***Wide*** and ***Timbral*** are not checked!



The operation of each mode is briefly described as follows:

Discrete Mode – In Discrete Mode each of the 31 bands operates entirely on its own with no interaction from information in adjacent bands. This is the Default operating mode in many of the factory presets.

Wide Mode – In Wide Mode the 31 band section operates as a pseudo-ten band limiter. How it works is like this: the second band of the 31 band section is paired with the first and third bands and the overall control signal is derived from the greatest of the three signals. This creates a ‘band one’.

The next group of three limiters creates a ‘band two’) by pairing the fifth band with the fourth and sixth bands. This sequence repeats across the audio spectrum until all ten ‘bands’ have been configured.

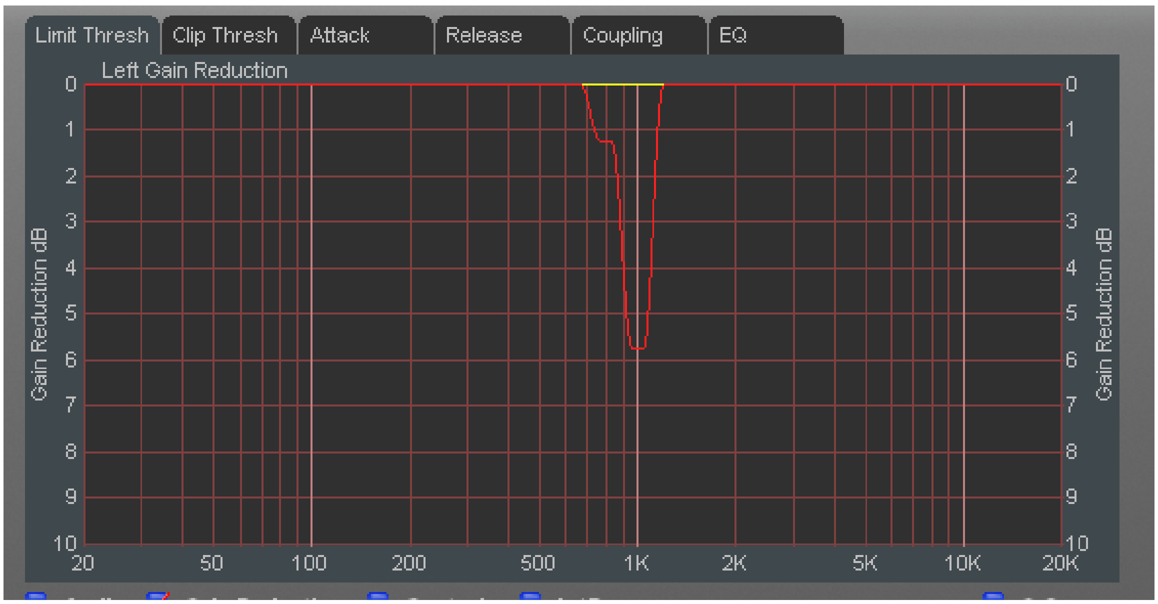
Timbral Mode – In Timbral mode FFT (Fast Fourier Transform) calculations are utilized in order to predict the second and third harmonics of a signal that is above the limit threshold in a band. Equal gain reduction is then created in the bands where those calculated harmonics fall.

Each of the bands has its own limiting threshold control as well as controls for adjusting the individual attack and release times.

You can think of the behavior of the 31 band when it is running in Timbral mode as getting all of the benefits of a broad band limiter while leaving behind every single one of its nasty habits.

In the following pages we'll explore how the 31 band section behaves with certain forms of non-program stimulus (tones - which makes it far easier to demonstrate it in a paper than with real program material!).

Below is a graphic of the 31 band section when it is operating in "**Discrete**" Mode and fed with a 1 kHz tone at a level 6dB above the threshold of limiting. Exactly as expected, the 1 kHz band is undergoing 6dB of limiting and there is no limiting occurring at any other frequency.

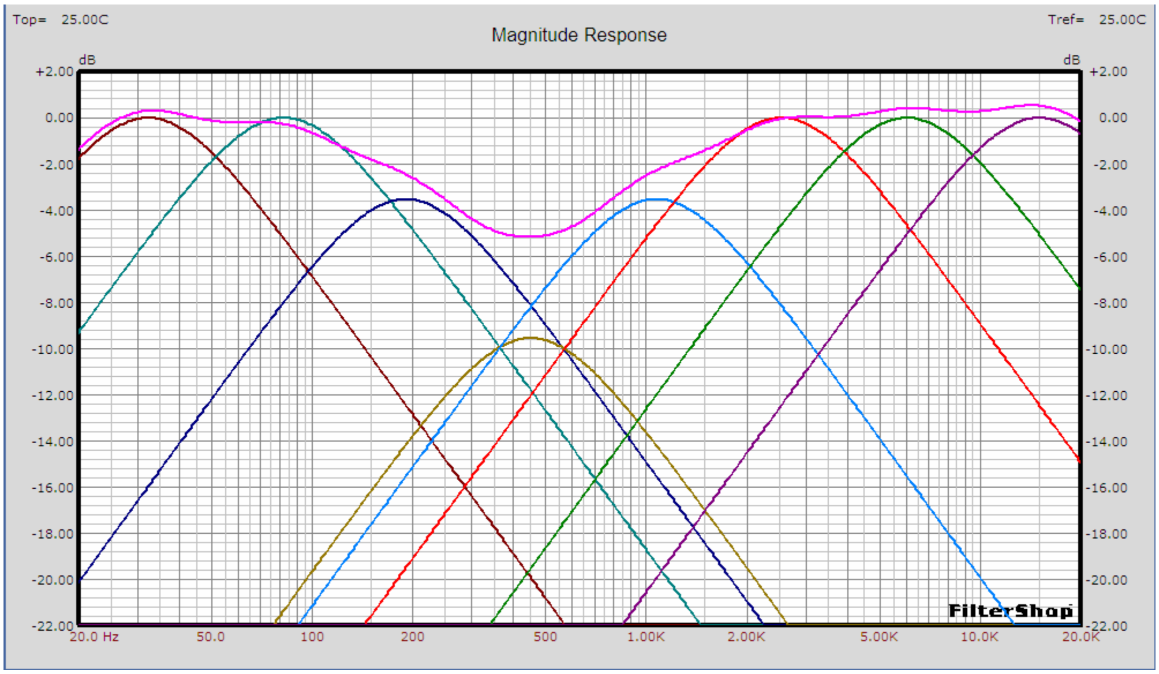


Because there is no limiting at any frequency other than what is required by the 1 kHz tone, any program material that would be near 1 kHz but not high enough to cause its own limiting activity ***would remain at its original amplitude***!

Because the dynamics effects of whatever work is being done within such a narrow band is essentially inaudible to the ear, we can get some nice acoustical benefits of that work (such as increased loudness) with few of the downsides that fewer bands would necessarily have to have, just by design.

What this means is that signals near frequencies requiring limiting remain audible instead of being unnecessarily reduced in level and the result is a louder, more transparent, and more detailed sound on the air.

Now refer back to the original 8-band limiter plot (shown again below). Here, program energy near the signal undergoing limiting would be reduced by half or more of its loudness and could be acoustically masked by program material that would not otherwise be able to mask it.



The 31 band algorithm does not have this fault. In fact many Vorsis customers have reported hearing details and instruments in familiar program material that they were never able to hear on any radio station before. But why should this be? There are two primary reasons:

- The first is that our five band AGC crossover is a 48dB/octave (yes, it's an 8th order filter bank!) design. It has been very carefully tuned for the best phase behavior that we know how to create. This results in a flatter frequency response without the unnecessary peaks or dips that cause other processors to either expose too much of, or lose detail in, a particular part of the audio spectrum.

- The other reason is that because the 31 band does not limit adjacent material when it needs to limit a particular sound, material that would have been masked or attenuated by other audio processors is actually *enhanced* because it is not being limited unnecessarily!

More about Discrete Mode

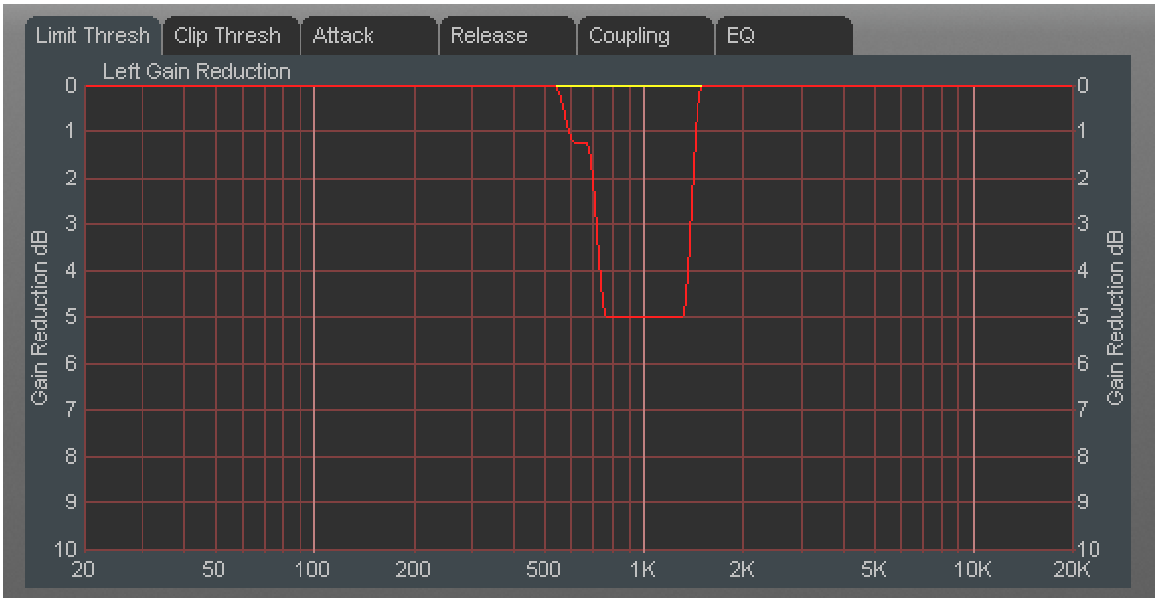
The Discrete Mode is the most transparent of the four operating modes because peak limiting occurs only at a frequency where the limit threshold is being exceeded.

One caveat of this mode is that because limiting occurs only at a frequency where the limit threshold is exceeded, material containing stringed and brass instruments might not sound as good as one of the other limiter modes if absolute faithfulness of the sound of those instruments is important.

This is because Discrete mode will perform limiting at such instruments' fundamental frequencies but *not* at its harmonics. When this happens, instruments can sound overly bright, metallic, or harsh because their harmonics have been left at a level which is now higher than it was originally and perhaps even higher now than the fundamental, depending on how much limiting is occurring at the fundamental frequency.

For program formats where the above behavior might be an issue we recommend the Timbral mode (described after the discussion about the Wide mode).

In the next graphic we'll show the same 1kHz stimulus condition but this time the 31 band operating mode has been switched to '**Wide**' instead of 'Discrete' as before.



Notice how the range of the frequencies undergoing limiting is wider than in the first example? This is because there are now <u>three</u> bands in limiting even though the 1 kHz tone is present at the center of <u>only</u> the middle band.

This operating mode sounds different in texture from the Discrete mode because wider bands are now involved in the limiting process. In fact, this is essentially the same behavior as the algorithm we use in our FM5, FM-10 HD, AM-10 HD, and VP-8 audio processors.

This operating mode may sound a bit more natural to announcers who are listening to themselves off-air than the Discrete and Timbral modes – we are simply not accustomed to hearing our voices with narrow pieces of spectrum at the most prominent signal frequencies reduced in amplitude!

On the air the effect is subtle with most announcer voice. Those listening to the station may never even notice that such manipulation of the announcer's voice is even taking place.

The next graphic shows the same 1 kHz stimulus condition as the first example, but this time we've switched the 31 band section to its "**Timbral**" mode.

Notice how the 1 kHz band is still in 6dB of limiting but now there is also 6dB of gain reduction occurring in the 2 kHz and 3 kHz bands? This is precisely what the Timbral Mode does – it calculates the harmonics of the primary signal requiring limiting and causes *equal* gain reduction at that frequency as well as its second *and* third harmonics.

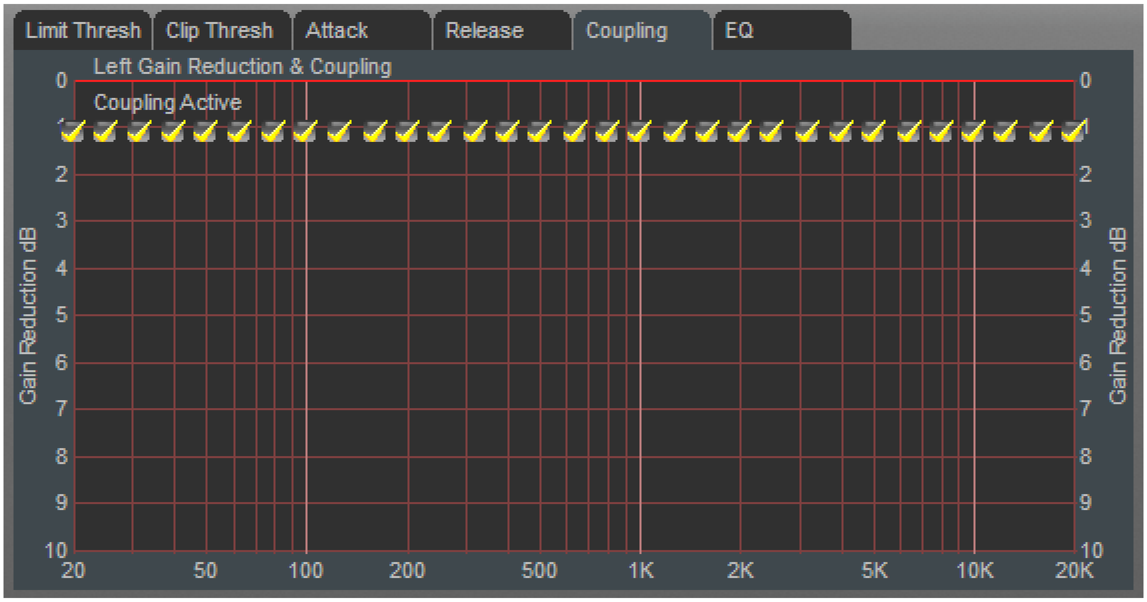
A Special Note about the Timbral Mode's 'Sound'

Timbral Mode is the most natural sounding of the three modes to process stringed and brass instruments or other complex material because harmonics are not emphasized, but rather are limited by the same ratio as the fundamentals. In essence it has some of the most desirable features of a broadband limiter but has <u>none</u> of the negatives.

We recommend the Timbral operating mode when the processor is used in formats where listeners with formal musical training might be sensitive to familiar instruments processed by 31 discrete bands. Because the Timbral mode preserves important harmonic amplitude information, those listeners may find the sound more natural than if the 31 bands were operating in the more competitive "Discrete" mode.

To Couple or Not to Couple?

The 31 band operating modes have yet one more feature that can be useful for certain types of program material. In the 31 band screen there is a tab labeled “Coupling”. Let’s see what it does...



The Coupling controls in the 31 band structure don’t couple the control signals from one band to another as is common for the term “coupling”, the controls ‘hook’ or ‘unhook’ a band from the Wide or Timbral mode when it is running. You cannot unhook a band from Discrete mode!

Normally all bands are “coupled” to the Wide or Timbral algorithm when enabled. Unfortunately we cannot give you a good an example of why you might want to uncouple bands from the algorithms because we’ve yet to find a benefit from having bands uncoupled from the algorithm. So, why are the coupling controls even provided? The simple answer is “because we can”.

If you have questions about operating the 31 band limiter section or have suggestions for how we might improve it, please feel free to let us know!

Wheatstone Corporation
600 Industrial Drive, New Bern, North Carolina 28562
252-638-7000
www.vorsis.com
techsupport@wheatstone.com