James L. Tonne, W4ENE
118 Garnet Lane, Oak Ridge, TN 37830: tonne@comcast.net

Speech Processing: Some New Ideas

James L. Tonne, W4ENE

The author presents some simple but high performance circuitry to enhance transmitted speech signals.

Speech Processing is the term commonly used to describe the kind of signal alteration that affects the speech waveform with the goal in mind of making the transmitted signal as loud as possible given a set of constraints.

The alteration can take the form of frequency response modification by using filters or tone-control techniques. It can involve automatic gain control of the speech signal volume by using a compressor. Or it can involve waveform alteration by using a clipper.

The constraints may be legal (as an example, we are only allowed to use a certain amount of power or occupy a certain bandwidth), or practical (the transmitter can handle only a certain signal level). Most commonly the limiting item will be the transmitter itself.

The Microphone

To design the processor we should first understand the microphone that supplies the signal to it. There are three types of microphones in common use at this time.

One type is the piezo-electric such as the crystal element, a well-known brand name example being the D-104. These devices tend to be fragile in nature.

Another type is the dynamic. It is expensive (because it is a labor-intensive device to construct) but rugged and usually of high quality.

And there is the electret element, a relative newcomer on the scene. These are small, inexpensive and quite satisfactory for this service. They require a pullup resistor and a bias voltage.

For a given speech volume level each of these microphones will deliver roughly the same output signal voltage (fortunately, for our purposes).

Now let’s look at the equivalent circuits of the microphones themselves. Figure 1 shows the equivalent circuit for the D-104 crystal microphone:

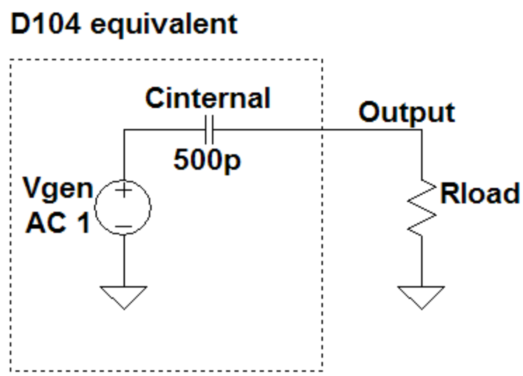

Figure 1 - This approximate equivalent consists of a voltage generator in series with a series capacitor of about 500 pF. This forms a single-pole highpass. As the load resistor goes down in value the low-end frequency response will roll off as seen in this plot:

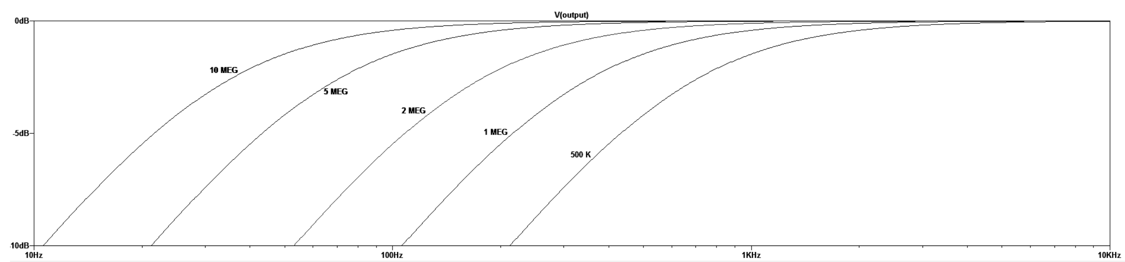

Figure 2 – This series of plots shows how the bass frequency response will roll off in such a type of microphone if the load is reduced (made less than a few megohms). In this case we see the response has dropped about 10 dB at 200 Hz with a 500 kohm load (not at all uncommon). This will vary somewhat from design to design.

Dynamic microphones are also load-sensitive. Figure 3 illustrates the important aspects of the equivalent circuit of a high-impedance dynamic microphone.

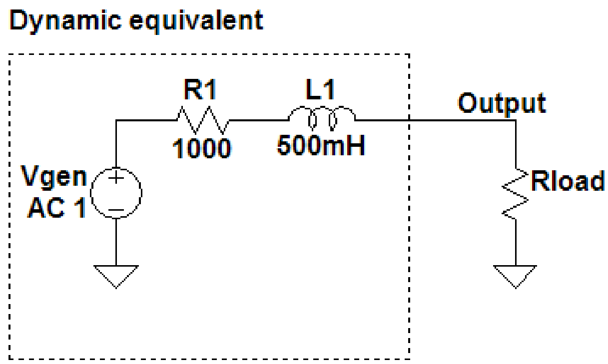

Figure 3 – An approximation of the equivalent circuit for a high-impedance version of a dynamic microphone. This is a single-pole lowpass. As the load impedance is lowered the treble response will be affected as shown in Figure 4. This is due to the series inductance inherent in the device. The signal level at low frequencies will also change because of the resistance internal to that inductance.

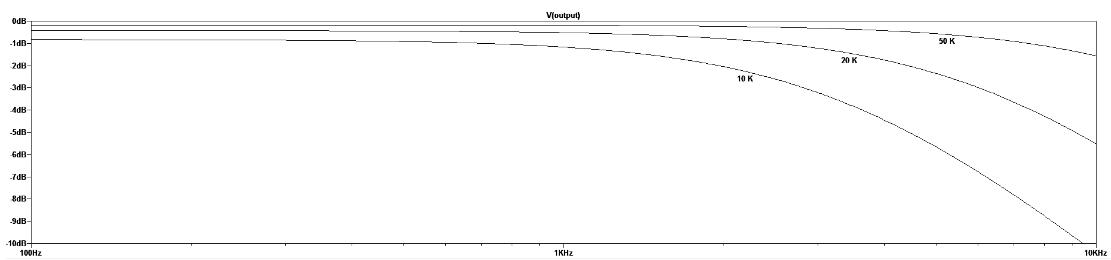

Figure 4 – Illustrating the treble rolloff seen in a dynamic microphone when a finite load is placed on its output. The degree of rolloff will be dependent on the inductance in the microphone. This of course will vary from design to design.

Electret microphones are also load-sensitive but their equivalent circuit has no components that are frequency-sensitive. As a result only the output signal amplitude varies when the load varies..

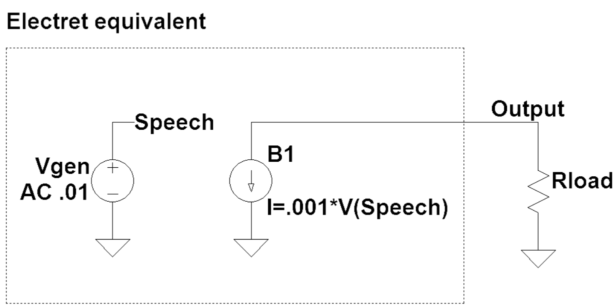

Figure 5 – The basic equivalent circuit for an electret element has no frequency-sensitive parts. The voice waveform generates a current, converted to a voltage by the following (external) load resistor. As this resistor value is increased, the audio signal level will also increase. In practice there are limits on the value of the value of the load resistor.

In most cases a value of between 2k and 10k is used for the load resistor. It is connected to a supply voltage of +5 to +9 volts in a "pullup" manner.

The Preamplifier

The output level from a microphone is generally of the order of 10 millivolts, too low to operate the usual manipulating circuitry. For this reason the signal is first applied to a preamplifier. That part of the proposed new design is presented in Figure 6.

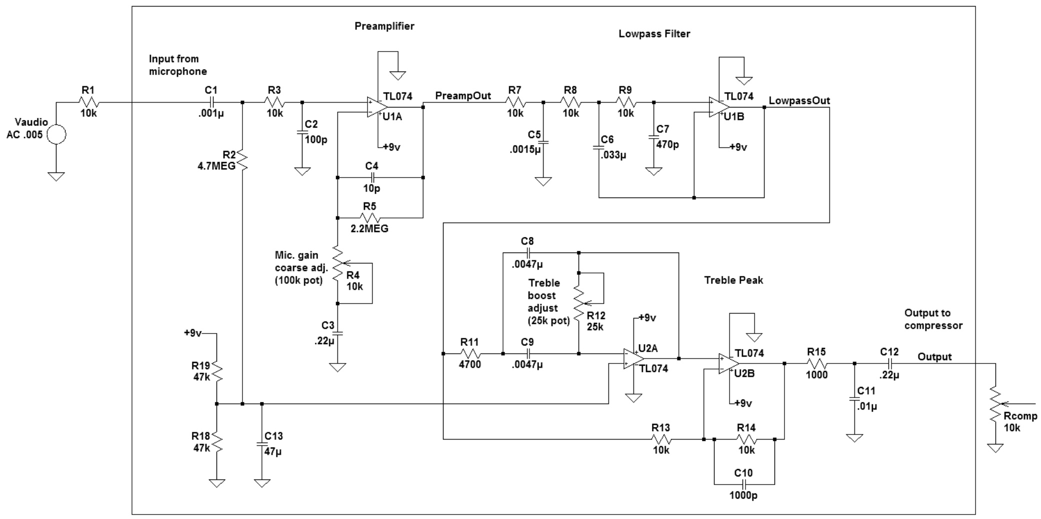

Figure 6 – This entire block of circuitry is being referred to as the preamplifier. The first amplifying stage in this schematic is the actual preamplifier stage. The input impedance is set by the 4.7 megohm resistor (assuming a CMOS or equivalent opamp), high enough that any of the previously-mentioned microphones can be accommodated. Next is a lowpass filter with a modest rise in its response prior to cutoff. Following the lowpass is a treble-peaking stage.

The lowpass filter right after the preamplifier stage can to some extent define the transmitted bandwidth. It can also prevent the following circuitry from having to respond to those signal components that will eventually be discarded anyway in a final post-clipping lowpass filter. For example sibilant audio frequencies (those in the "hissing" region) can cause gain reduction in an audio AGC system and so reduce the overall volume of a signal. It is best to eliminate such mischief-making treble components prior to further processing. This filter also offers a nice opportunity to provide a frequency-response enhancement by manipulating the component values. The response of this filter by itself is shown in Figure 7.

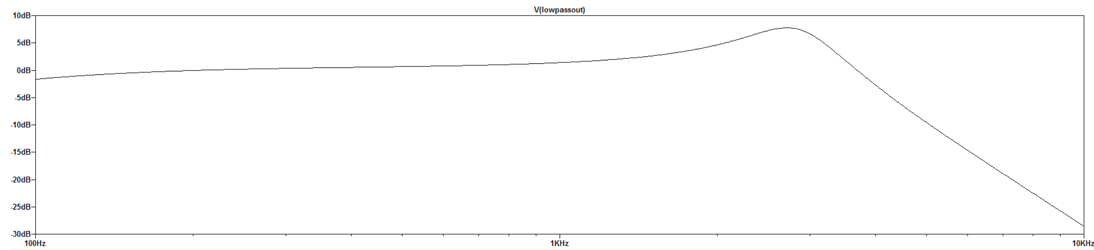

Figure 7 – The response plot of the lowpass filter following the preamplifier stage after component values have been changed to have a response peak. The response is down about 25 dB at 10 kHz even after modifying part values to provide a 6 dB peak in the 2500 Hz region.

Following the lowpass filter is a treble peaking circuit. This is far more efficient circuit for this purpose than the common “tone control” type of circuit. As the boost control is adjusted we see the response family of curves seen in Figure 8.

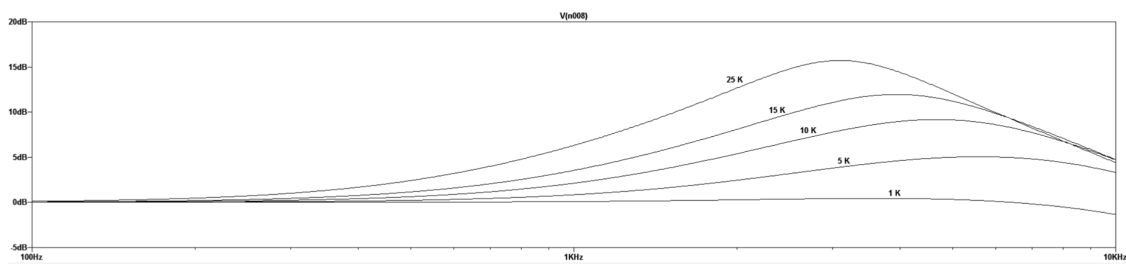

Figure 8 – Family of response curves for the treble peaking circuit alone . This set of curves illustrates the set of responses available from the peaking circuit as its adjustment control is set to 1k, 5k, 10k, 15k and 25k ohms.

The responses of the preamplifier alone, then of the following lowpass filter and finally of the of treble peaking stage are all shown together in Figure 9.

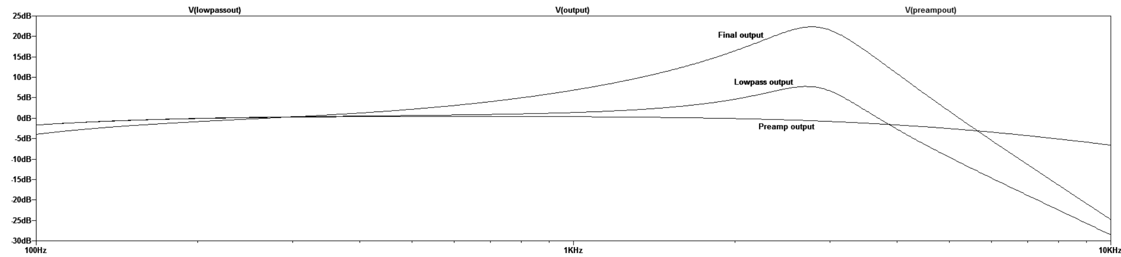

Figure 9 – Family of curves showing the response plots for the input preamplifier itself (bottom), the output of the following lowpass filter (middle), and the final output of this block of circuitry, the output of the treble peaking circuit. For this series the treble boost was set to maximum.

Because "peaking" techniques have been used instead of the older "step" (tone-control) methods, significant amplitude boost can be achieved in the important treble region while at the same time restricting the signal bandwidth.

At this point in the discussion the signal frequency response of the system has been optimized for communications purposes and its level has been built up in amplitude so that it may be applied to the next item in the chain of processing: the compressor.

The Compressor

Next in a cascade of circuits in the processor is a compressor, the purpose of which is to compress the dynamic range of the speech signal – bring up weak voice passages and lower the volume of strong voice passages. Ideally this will be done in some low distortion manner on a syllable by syllable basis to maximize the transmitted signal volume. In this processor it is used to control the signal level into a following clipper stage. In this way, the effectiveness of the clipper is clearly enhanced. Or the clipper can be used only to "catch" those transients that escape the compressor.

The schematic of the compressor with its associated metering circuitry is shown in Figure 10.

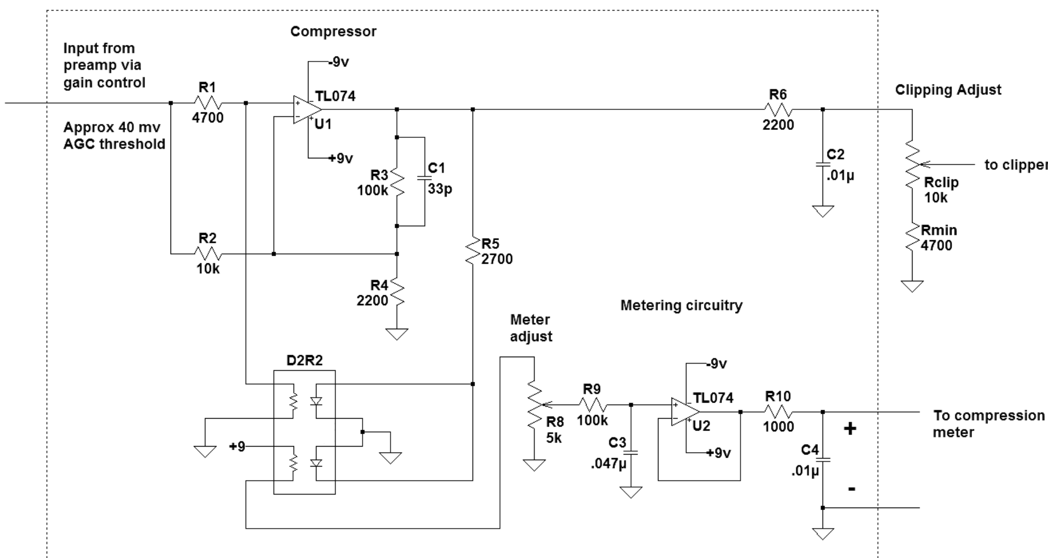

Figure 10 – Schematic of the compressor with its metering circuit.

The signal from the first (preamplifier) block is applied to the input of this circuit at the input side of R1. On the output side of R1to ground is a second resistor in the form of a photoconductor. Changing the resistance (or conductance, depending on one's viewpoint) of that element will change the amplitude of the signal entering the opamp. The audio signal appearing at the output of the opamp is applied to a pair of LEDs via a resistor (R6). The LEDs are arranged so that either polarity of audio will cause the photoconductor to conduct, provided the audio signal amplitude is sufficient to excite the LEDs. The circuit as so far described is quite workable but will be found in practice to have a limited control range. That range is greatly extended by adding a single resistor (shown here as R2) making the circuit function in a bridge-like manner.

It will be noticed that there are neither attack time nor release time components (rectifying diodes, storage capacitors, resistors and the like) in this circuit. Photoconductors have intrinsic time constants of their own, both when illumination is applied and when it is removed. This circuit is taking full advantage of that oddity. Plain audio signals are applied to the pair of LEDs (via R6) which illuminate the photoconductor. As a result of this minimalist circuit configuration, the release time (and to a lesser extent the attack time) is provided by that characteristic of the photoconductor operating here in conjunction with R1. R1 was set to yield the proper recovery time for the purpose at hand. R2 is to be twice the value of R1.

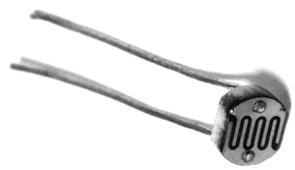

Figure 11 – Photo of the small cadmium sulfide photoconductor cell of the type used in this compressor circuit. It has a diameter of about six millimeters.

Fortunately it appears that these little photoconductors seem to be satisfactorily similar brand to brand, unit to unit, at least for our purposes. The LEDs are mounted, presumably on an etched circuit board, so that they face upward at the cell as illustrated in Figure 12. The photoconductive area of the cell faces downward toward the pair of LEDs. Chip type LEDs may be used as well.

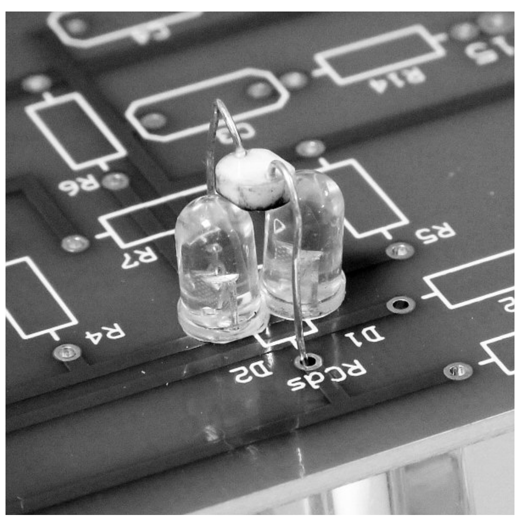

Figure 12 – Showing how the photoconductive cell faces downward toward the pair of LEDs.

As shown in the schematic (but not in that particular photograph) a second cell mounted in a similar fashion is used to monitor system activity. The conductivity of that second cell will increase with increasing compression, providing an increasing voltage at the metering calibration control (shown as R8 in the schematic). The wiper of this control connects to the metering buffer IC (U2). The buffer drives an analog meter movement with a slightly nonlinear scale calibration as illustrated in Figure 13.

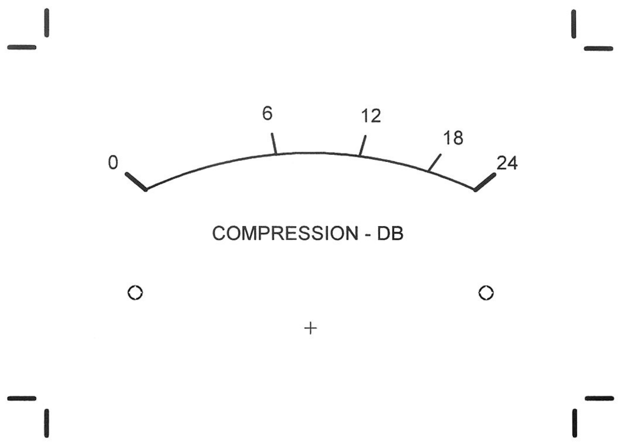

Figure 13 – Compression meter scale

To test the compressor we apply a tone burst seen in Figure 14.

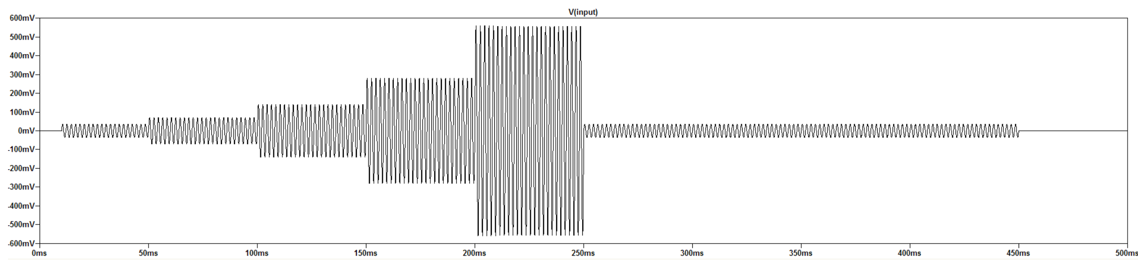

Figure 14 – Appearance of tone burst as applied to compressor input

This burst is a 500 Hz sinusoid with a peak amplitude of 35 millivolts starting at 10 milliseconds. At 50 ms the amplitude is doubled to 70 mv. At 100 ms the amplitude is again doubled, and again at 150 ms and at 200 ms. Finally at 250 ms the signal level is returned to its original level of 35 mv. At 450 ms the level goes to zero. From a practical viewpoint, we will consider the first level to be AGC threshold and each increase after that is a 6 dB increase.

The appearance of the signal at the output of the compressor with the burst applied to the input is shown in Figure 15.

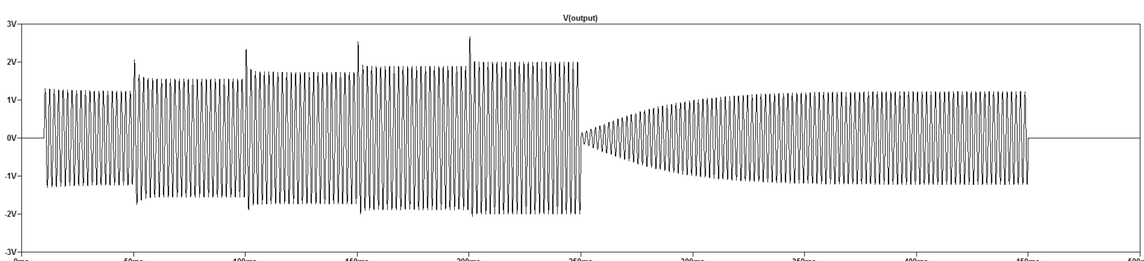

Figure 15 – Tone burst as it appears at the compressor output

To be observed is that the output signal amplitude increases somewhat when the input signal level increases, but to a far lesser extent. The dynamic range of the signal has been compressed. The time that it takes to accomplish this is about one cycle or less of the 500 Hz signal; the attack (reaction) time is about one millisecond or less. When the output signal is reduced from its final level back to its original low level the recovery (gain increase) time can be seen to be about 70 milliseconds, quite optimum for speech.

A 24 dB input level variation has been compressed down to about a 3 dB output level variation. This can also be seen as an input vs output plot in Figure 16.

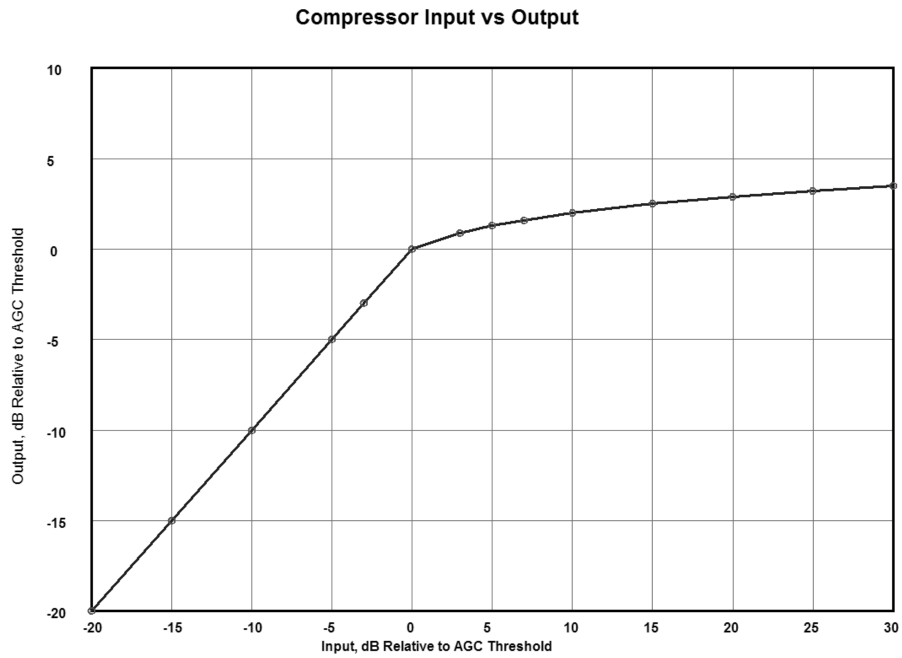

Figure 16 – Compressor input vs output shown in plot form

The output signal from the compressor is rather uniform in level but for further processing it may be applied to a clipper followed by a lowpass filter.

The Clipper

The final stage in this processor is a clipper followed by a modified lowpass filter. This block is shown in Figure 17.

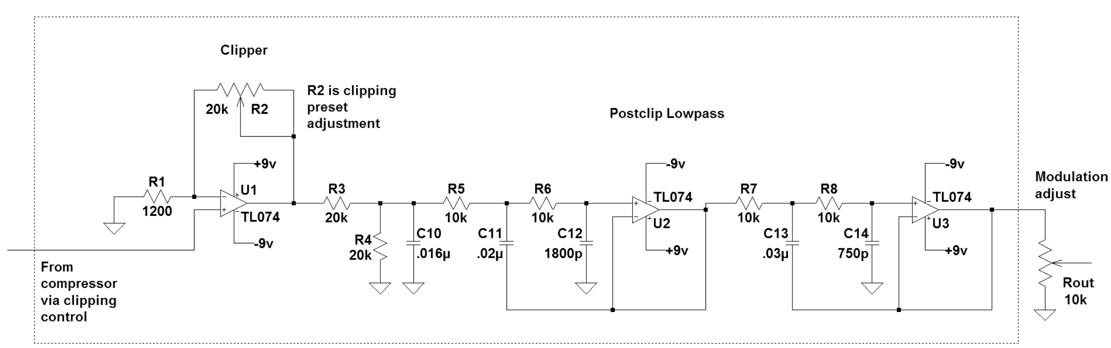

Figure 17 – Clipper and its lowpass filter

The compressor tends to be peak-sensitive. Rephrased, it will respond to the peak instantaneous value of a complex (for example speech) waveform. This makes it admirably suited for single-sideband systems, which cannot handle the flat-topped (clipped) waveforms exiting from a clipper. The use of a clipper should be restricted to amplitude modulation systems or for use as an overshoot-protection method in SSB systems.

In this processor an opamp is used as the clipper. The output of the opamp does not need to go rail-to-rail but it must be of a type that clips (overloads) symmetrically.

Clippers must always be followed by a significant lowpass filter to reduce the amplitude of the unwanted harmonic components generated by the clipping process. Clipping also generates intermodulation components which are even more audibly disturbing than the harmonic components. Unfortunately these intermodulation components cannot be removed.

The spectral output of a clipper handling a 500 Hz sinusoid with 10 dB of clipping is shown in Figure 18.

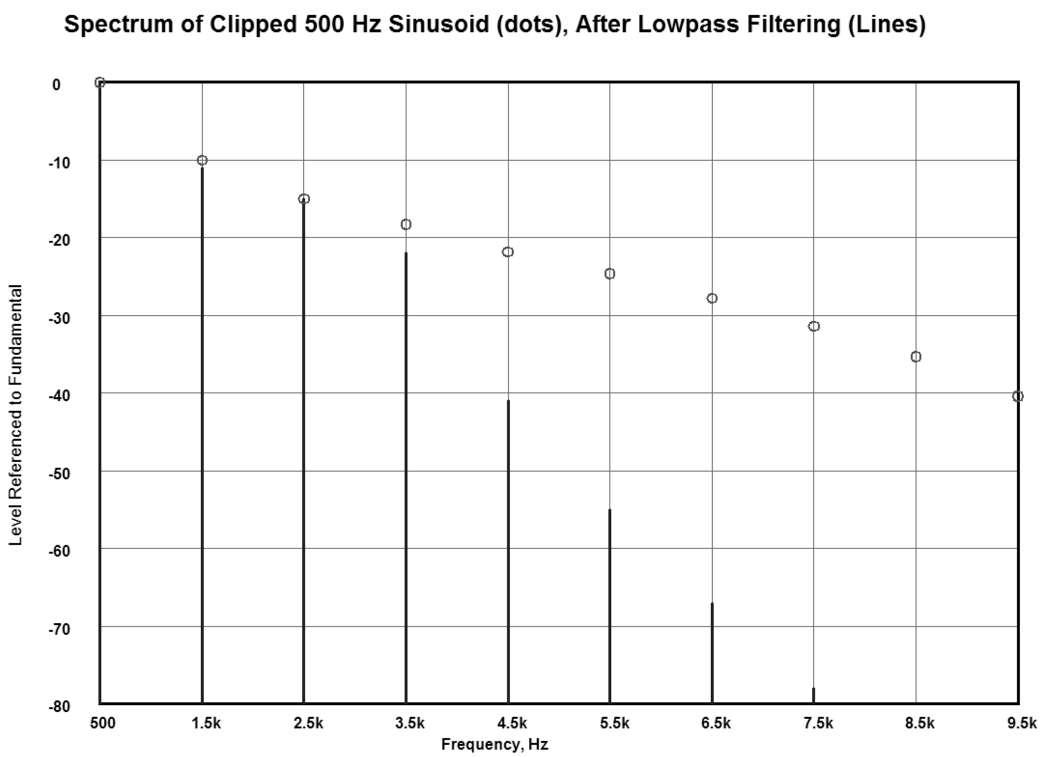

Figure 18 – Spectrum of output of a clipped 500 Hz sinusoid (dots) and after the post-clipping lowpass filter (lines)

As can be seen, even 10 dB of clipping can result in a significant amount of harmonic content. This will result in a transmitted signal with a needlessly wide occupied

bandwidth. A lowpass filter following the clipper can reduce the level of those components to tolerable levels. Ideally this filter should have a reasonably flat response in the passband and have a sharp cutoff. But simply placing a lowpass filter in this location can introduce another problem. It can overshoot on severe transients such as a squarewave (which can result from the clipping process). This overshoot is shown in Figure 19.

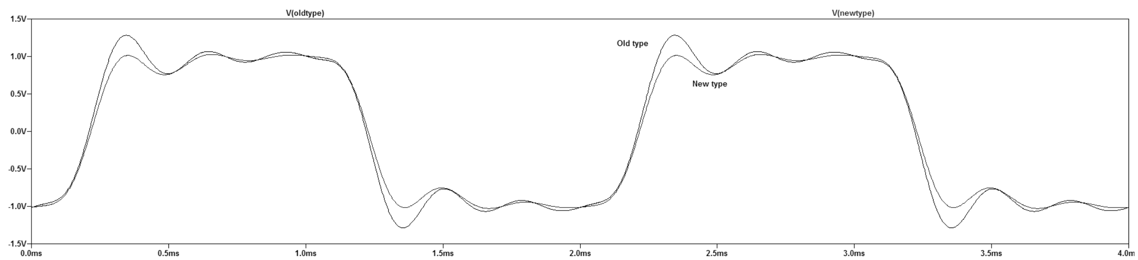

Figure 19 – Lowpass filter output waveforms, old and new types as they respond to an applied squarewave

The usual sharp cutoff lowpass filter following a clipper would have overshoot on its output requiring the modulation level to be reduced because of that overshoot. But this new design does not have that overshoot. It has only what might be termed "undershoots."

There is a tradeoff in in this new filter design. The frequency response is not flat. The output level is reduced somewhat at the upper audio frequencies. This is shown in Figure 20.

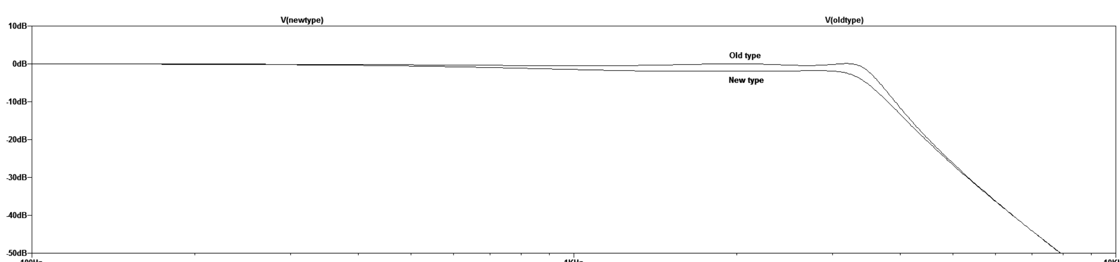

Figure 20 – Comparison of the old and new post-clipping lowpass filter responses

Both the old designs and this new design roll off at 30 dB per octave and are down 50 dB at 8 kHz. The new design has a small step in its passband response with the result that the overshoots are negated. The lower audio frequencies are transmitted at full level and so loudness is maximized.

The clipper with its lowpass filter can be used in an SSB transmitter as an overshoot-protection measure. It can be used much more aggressively in an AM system to alter the speech waveform to enhance loudness provided the transmitter can handle the clipped waveform. In each case the post-clip lowpass filter can be used to limit the transmit bandwidth. This is especially valuable in phasing type SSB generators for bandwidth control.

The System

Each part of the processor has been examined. Now they are put together as shown in Figure 21.

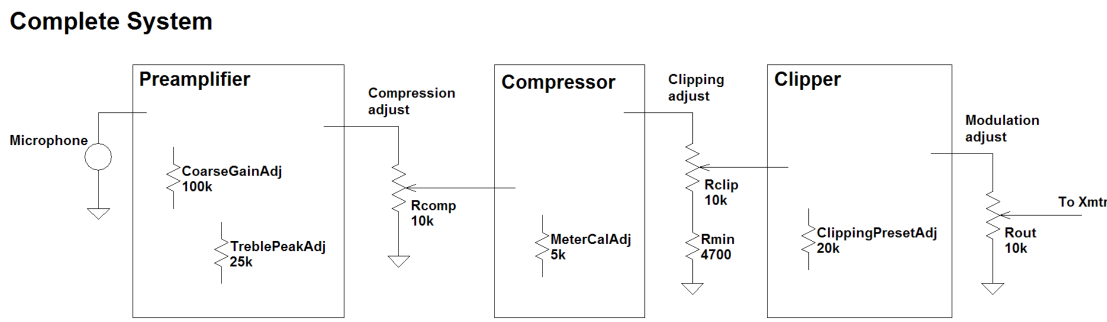

Figure 21 – The complete processor with all three blocks of circuitry connected

As a reminder, the LED and photoconductor assembly must be shielded from ambient light. If initial adjustment and trouble-shooting is being carried out on the usual workbench then that small assembly must be shielded from light or else the room lighting must be turned down.

Speaking into the microphone should yield several volts peak audio at the output of the preamp module. Set the internal coarse gain adjustment (R4 on the preamp schematic) to achieve this level. As the Compression control is advanced with normal speech, a point will be reached where the LEDs will begin to flicker. R4 may be reset as needed so the Compression control is in a desired setting. The output of the compressor will now be relatively constant with varying speech input levels. The internal treble peaking adjustment will have to await on-air testing although monitoring the compressor (or system) output with an amplifier and headphones can prove helpful at this point.

Set the Clipping control to its minimum setting. With a sinusoidal signal input (whistle or better yet an audio oscillator) set the clipper's internal clipping preset control (R2 on the clipper schematic) so the clipper is on the verge of clipping. When the system is set this way, the Clipping control will allow the user to adjust the degree of clipping from 0 to 10 dB of clipping. This is set by the ratio of the values of resistance of the clipping potentiometer and the resistor from its bottom end to ground (3 to 1, Rclip and Rmin in the overall schematic).

To adjust the meterscale, apply a sinusoid to the input and slowly increase its amplitude until compression is observed. Increase the amplitude 27 dB. Set the internal meter calibration adjustment control for a full-scale reading of 24. If you have no intention of using that much AGC activity of course feel free to calibrate the scale as desired.

In operation, it will be found that the treble boost will be especially useful during contests, along with 12 dB of compression. A few dB of clipping can be used in AM if the transmitter is capable of handling it. Always monitor the transmitter using an oscilloscope when using speech processing.

The grandfather of this device was designed by the author and published in *QST*, September 1956. This design seriously outperforms that unit but is tiny by comparison. This article describes an upgrade of the design presented to the Oak Ridge Amateur Radio Club, an ARRL-affiliated club.

James L. Tonne holds the Amateur Extra class call sign W4ENE and was first licensed in 1951. His current Amateur Radio interests are largely focused on speech processing and computer-aided circuit design He has written several articles for QEX and QST and was a major contributor to the RF and Filters chapter in the ARRL Handbook. He is the author of the Tonne Software package on the CD accompanying the ARRL Handbook and included as part of the downloadable package available on the ARRL website.