**ELECTRONOTES**
1016 Hanshaw Road
Ithaca, NY  14850

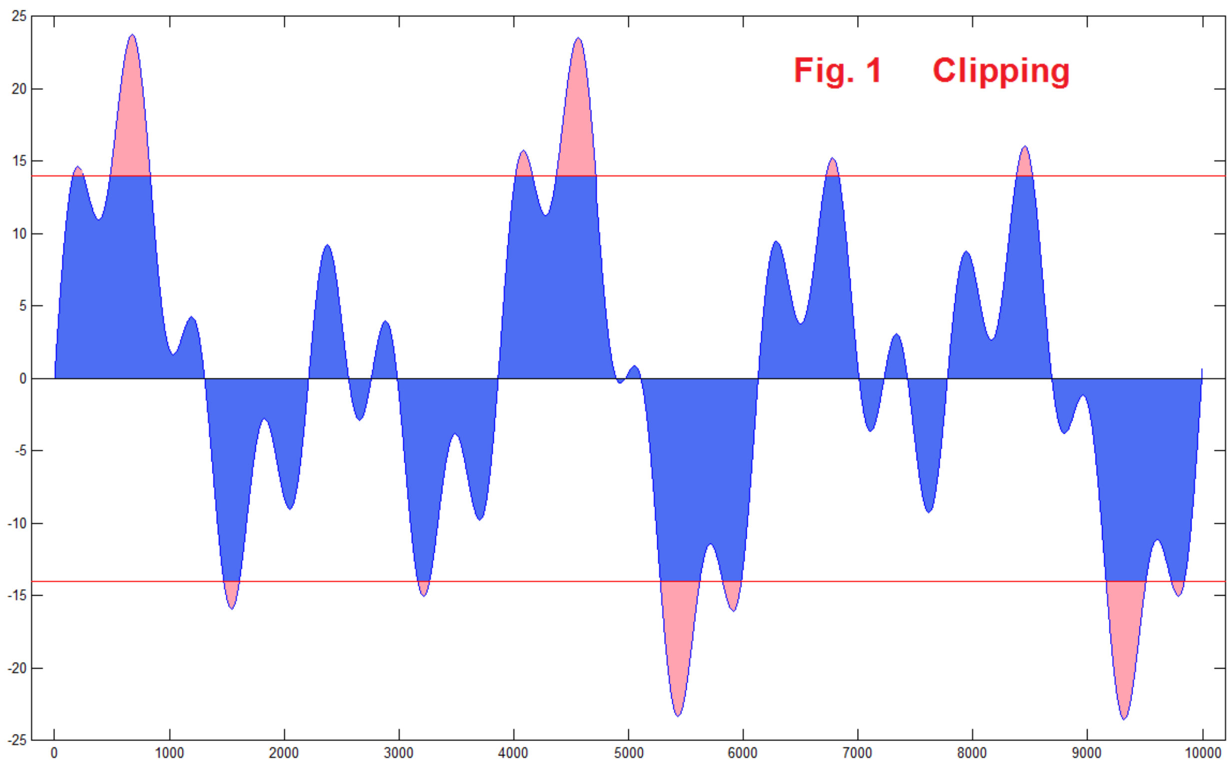
**APPLICATION NOTE NO. 397**

June 15, 2013

# CLIPPING: FRIEND OR FOE

The term "clipping" (sometimes "saturating" or "pegging") means that a voltage, usually a signal, is blocked from taking on its proper value by some limit. Often the limit is well-understood and appreciated as unavoidable even though not what we might call a desirable "feature". The most familiar form is probably clipping of a signal against one or both power supply limits. The signal is simply too large. Instead we get that portion of the signal within the power supply limits, and the portion that is beyond the limits is replaced by the limit. Fig. 1 shows a typical example where three sinusoidal components are added and portions of the sum exceed the limits (±14). Here the ±14 limits are chosen as what might be expected with op-amps powered between ±15 supplies.

In the figure, the blue portions of the signal are faithfully retained, but the pink portions are clipped off. Obviously this makes a difference. It is true, and perhaps most obvious,



that the peak amplitudes are reduced. ( Here this does not appear as terribly serious, but examples could easily be shown where amplitudes are greatly changed. The large unclipped signal seems to be about ±24, and this is clipped at ±14. But there are few actual peak regions here. If the signal had been say ±50 clipped at ±14, the loss of amplitude would of course be much more dramatic. ) The second thing to note is that the clipping is going to introduce distortion, and the spectrum (originally just three sine waves, by information we have volunteered) will have many more frequency components due to the clipping. It turns out that a broader-banded signal may have an apparent loudness that is noticeably, or even greatly in excess of that of a signal with a simpler spectrum, even for very similar amplitudes. In consequence, the clipped signal may have an apparent loudness greater than that of the unclipped signal.

## Some Tests:

### Clipping Sinewave:

As a first example consider what is an easy experiment or likely a result easily retrieved from memory of just playing around with circuits. Listen to a sine wave and turn up its amplitude until it clips. Right at the point where it starts to clip you hear a sudden change. This is a change in acoustic character, mainly the generation of harmonics, but perceived in part as an increase of overall loudness. It is more likely to wake us up.

### Comparing Waveforms:

A simpler experiment is just to take a music synthesizer and play out the waveforms. Comparing the sine wave to a square, sawtooth, or pulse, all of which are designed for the same amplitude, should convince you that the latter three are “louder” than the sine, at least in the sense that the sine is more “mellow” and the other three have more “bite”. The triangle wave, in comparison is much more similar to the sine. Indeed, the harmonic content of the triangle is much less than the square, saw, or pulse.

Accordingly we recognize that a change to a richer spectral content can be conflated with a perceived change of loudness. This is not an “error” but simply the way the ear-brain works. Sorry if this is troubling to engineers!

### Equalizing Power:

Above we thought of clipping from the viewpoint of amplitude. (We noted intentional design achieving the same amplitudes for various synthesizer waveshapes.) Instead of amplitudes, can we compare waveforms of the same power? Yes, but we need to be a little careful. It is clear enough what the power in a sinewave is. We have only to integrate the sine squared over a full cycle, and divide by the cycle length:

$$P_s = \frac{1}{2\pi}\int_0^{2\pi} sin^2\theta \, d\theta = 1/2 \tag{1}$$

which is the well-known result. Thus the RMS (root of the mean squared) voltage for a sine wave is $\frac{1}{\sqrt{2}}$. Now we consider the ordinary pulse, as it might be obtained by clipping, such that it is some high level (say α) for the first part of a cycle and then abruptly changes to a low level of the opposite sign (thus to –α) for the remainder of the cycle. Thus the magnitude of the pulse is always α and the power is obtained by integrating $\alpha^2$ over a full cycle, say of length T, thus:

$$P_p = \left(\frac{1}{T}\right)\int_0^T \alpha^2 dt = \alpha^2 = 1 \tag{2}$$

where we have set α=1 so that the pulse has amplitude = 1 the same as the sine. This is essentially what we have with a music synthesizer where all waveforms have the same amplitude, and typically the signals range between ±5. So we come to the conclusion, comparing equations (1) and (2), that the sine has half the power of the pulse. So of course it sounds louder – you might say. But we do need to be careful here.

**The Fourier Series View:**

The sinewave is taken to have zero mean (zero DC component) while the pulse most assuredly does not (except for the particular case of the pulse known as a square wave). You can't hear the DC component. So how do we allow for that? Let's choose a specific example**:** a pulse of "duty cycle" 1/5, which is high for 1/5 of a cycle and low for the remaining 4/5, and assume the levels of the pulse are ±1. This has a DC value of -0.6 (1/5-1/5-1/5-1/5-1/5), and the fundamental amplitude is 0.74839 (using a standard Fourier Series approach). The value of 0.74839 for the fundamental may seem large, but after all, when the duty cycle becomes 1/2 (a square wave), the fundamental exceeds 1 (becoming 4/π = 1.2732). Fig. 2 shows the Fourier Series reconstruction (blue curve) using 25 components with the coefficients printed to the right of the figure, and we see this is approximating the pulse of 1/5 duty cycle (keep in mind the periodicity). Here we include the DC term of -0.6. Here we also show the ordinary unit amplitude sine wave (red). The real point is perhaps that we also plot (dashed blue) the first sinusoidal term of the Fourier Series of the pulse. As mentioned, its amplitude is 0.74839, below 1, but not that small either. So when we listen to the pulse, we are listening to the red curve, attenuated to 0.74839, PLUS all the harmonics, and this sum sounds louder.

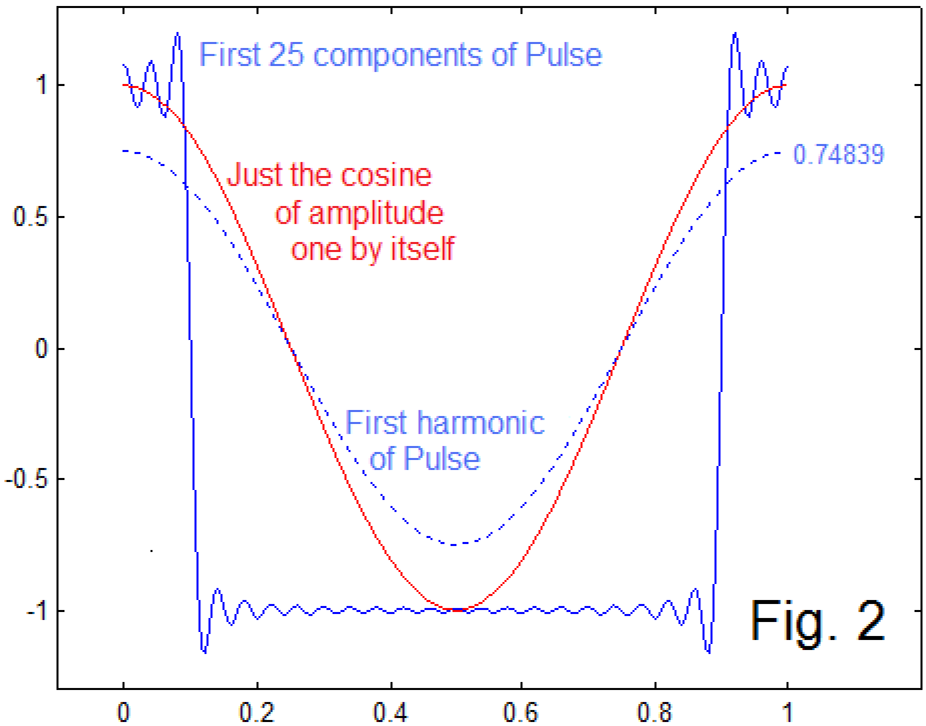


Fig. 2

a(0) = -0.6
a(1) = 0.7484
a(2) = 0.6055
a(3) = 0.4036
a(4) = 0.1871
a(5) = 0.0000
a(6) = -0.1247
a(7) = -0.1730
a(8) = -0.1514
a(9) = -0.0832
a(10) = 0.0000
a(11) = 0.0680
a(12) = 0.1009
.............

**<u>Equalizing the Power and Removing the DC:</u>**

As we said, the pulse confined between two limits (such as we would get with clipping and as produced by most music synthesizers) has a DC bias, except for the exact square wave. Suppose, for purposes of calculation, that we decide to name a high level of A for the pulse and a low level of B. We then want to choose A and B to meet two conditions (two equations in two unknowns after all). The first condition is that the pulse has no DC component. The second is that the power in the pulse is the same as that of a unity amplitude sine wave (i.e., 1/2). This is easy to solve. First, to get a DC level of 0 we need to have:

$$\frac{A}{5} + \frac{4B}{5} = o \qquad \text{(3a)}$$

That is, for 1/5 of the cycle the voltage is A, and for 4/5 it is B. This solves to:

$$A = -4B \qquad\qquad or\ B = -A/4 \qquad \text{(3b)}$$

The condition on the power is that:

$$\left(\frac{1}{T}\right)\int_0^{T/5} A^2\,dt + \left(\frac{1}{T}\right)\int_{T/5}^{T} B^2\,dt = \frac{1}{2} \qquad \text{(4a)}$$

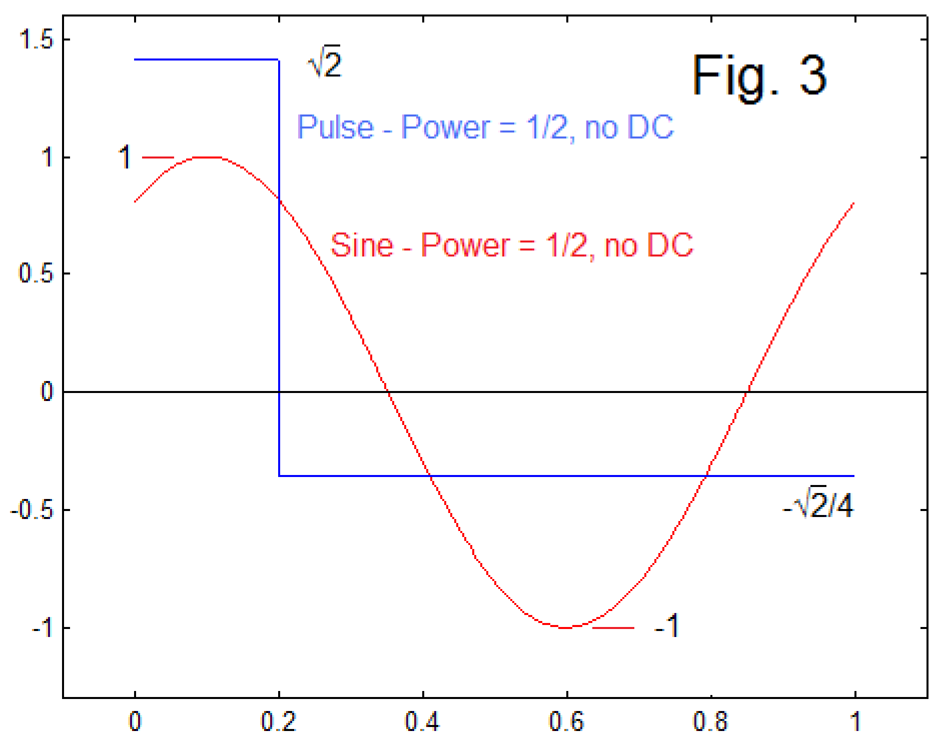
$$\left(\frac{1}{T}\right)\int_0^{T/5} 16B^2\,dt + \left(\frac{1}{T}\right)\int_{T/5}^{T} B^2\,dt = \frac{1}{2} \qquad \text{(4b)}$$

$$\frac{16B^2}{5} + \frac{4B^2}{5} = \frac{1}{2} \qquad \text{(4c)}$$

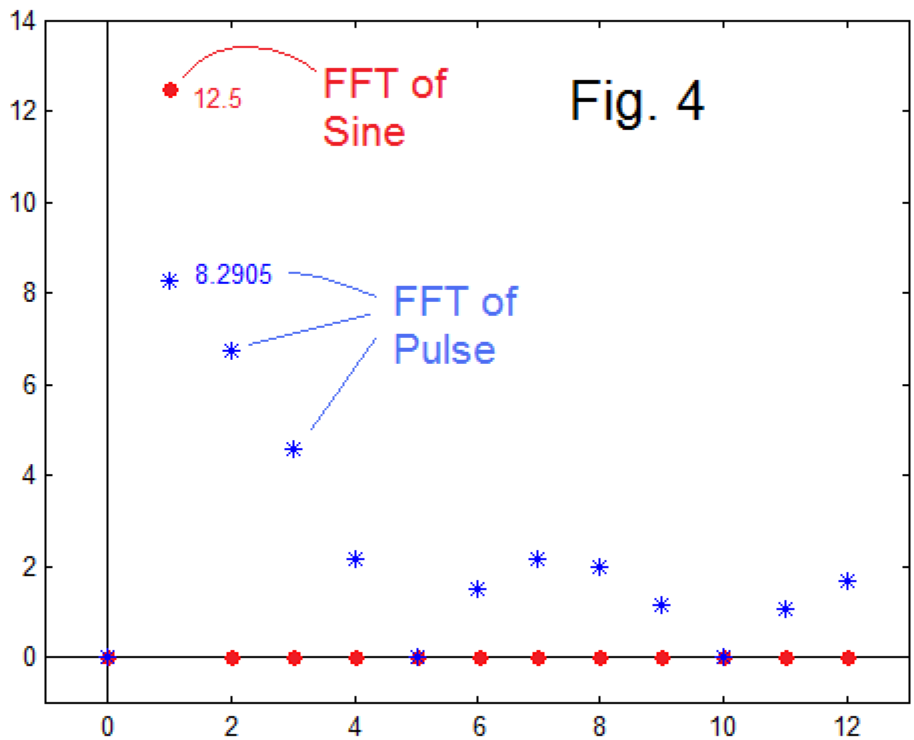
$$B = -\sqrt{2}/4 \qquad \text{(choosing the negative root)} \qquad \text{(5a)}$$

$$A = -4B = \sqrt{2} \qquad \text{(5b)}$$

[Note that we didn't really need the integrals of (4a) and (4b) since we just have rectangles, and could use the areas directly (4c).] In either case, (5a) and (5b) are the answers. This is drawn in Fig. 3



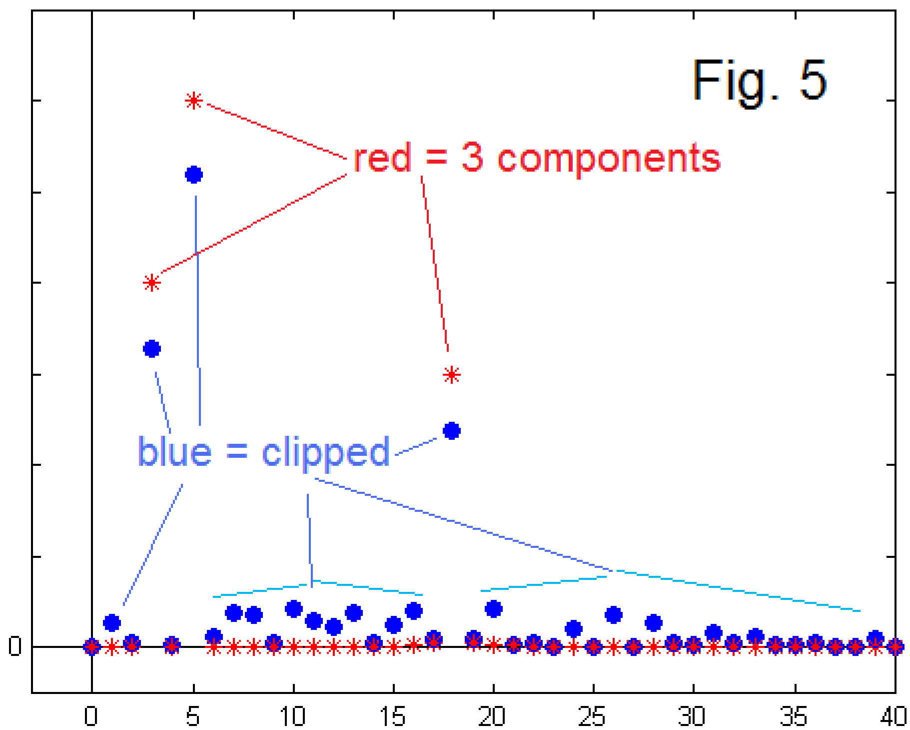
In Fig. 3 we have adjusted the phase, relative to Fig. 2 for a slightly different view. Note that we have chosen for both figures to plot a sinusoidal waveform of amplitude 1. The pulse in Fig. 3 has an amplitude of $5\sqrt{2}/4 = 1.7678$ and is thus smaller than the pulse in Fig. 2, which has amplitude of 2. Accordingly, the pulse that gives the same power as that of the sine wave is smaller than the full-amplitude pulse, and would be less loud. None the less, subjectively it may be louder, or have more "bite". We should note that the fundamental of the pulse in Fig. 3 (an attenuation of the red curve – not shown) is of course smaller than that of Fig. 2. Its amplitude should be 1.7678 x 0.74839 = 1.3230 which would be 0.6615 times the amplitude of the sine wave. A Fourier Series analysis should confirm this.




Here it is quite convenient to use an FFT rather than a Fourier Series, which we <u>expect</u> to give <u>almost the same result</u>. We will use a discrete 25 point version of the pulse, and of the sinewave of Fig. 3, and look at the magnitudes of the first 13 points of the FFTs, as shown in Fig. 4. The spectrum of the sine is just the single spike at k=1 of the FFT, and the magnitude is 12.5 (also same at k=25 not shown). The pulse has the periodic sinc structure evidenced by the blue stars of Fig. 4 (compare also to data at right of Fig. 2). Further, as is expected for a pulse of 1/5 duty cycle, every $5^{th}$ harmonic is missing. Note that both spectra are 0 at k=0, verifying that we have no DC. The key result to see if we have correctly reasoned through all this is the ratio at k=1 which is 8.2905/12.5 =0.66324, nearly the same as 0.6615 obtained above. So it worked.

**Back to the Clipped Signal:**

We have taken a detailed side trip to examine a simple sinewave and a simple pulse. However, we started with a more or less arbitrary combination of three sinewaves in Fig. 1. We have now come to expect to find additional frequency components as the result of clipping. Further, we expect additional frequency components will likely give a subjective impression of a "louder" signal. So here, what does the FFT of the signals (clipped and unclipped) of Fig. 1 look like? This is shown in Fig. 5.




The unclipped signal in Fig. 1 is basically three sinusoidal components, and these are seen as the three larger valued red stars, with the remainder of the red stars around zero representing the lack of additional components. By clipping, we are removing portions of the signal. This has two effects. Note first that the frequencies represented by the three main red stars are still there as blue dots somewhat below the red stars. This is what is retained. The major change is the cluster of blue dots somewhat above the horizontal axis replacing the red stars that are virtually on the axis. In this sense, this cluster is the clipped portion. It is broadbanded material. We might well call it distortion.

**Designing Within Limits:**

In describing clipping as "Friend or Foe" in the title of this AN, I am suggesting a design within realistic limits. If we were designing a garage for a vehicle, we would recognize that the abode had the purpose of housing the vehicle safely within limits, but for the most part, also making use of available resources. Most garages (sadly) are not that much more capacious than is necessary for the primary function. In the case of the design of circuits to handle signals, this is part of the S/N (Signal-to-Noise) consideration.

It is a joke among electrical engineers that one way to improves S/N, in lieu of getting down to a more difficult approach of reducing noise, is to increases the signal. This means in general, to make the amplitude levels as near to what the system supply voltages (and sometimes, the speed capabilities of active components) will allow, at as many points within the circuitry as possible. If you have supplies of ±15 volts, you should perhaps try to have signals of levels ±10. If your supplies are ±12, you may well settle for ±5 signals. Why not supplies of ±300 volts? Your signals then might well be ±250.

Indeed, why not. Well it will come as a surprise to some reading this that voltage levels of 300 volts (600 volts differential) can give you quite a jolt. (Many old-timers will even be cautious to place fingers across the full 30 volts of a ±15 system.) Old tube circuits typically had several power supplies, including one called the "B+" supply which could be as little as 90 volts or as much as 450 volts, or so. This was the "plate" supply, originally supplied by what was by tradition called the "B Battery". (The A Battery was something like 6 or 12 volts, which heated the filaments or heaters.) To complete this reminiscence, I recall radio repairmen in the army who were comfortable using a "wet finger" (literally) to trace the B+ voltages. Much faster than using a meter. (A "wet finger" repairman was consequentially a traditional term for someone who repaired on instinct, and the term thus used was not related at all to testing the direction of a prevailing wind.) It was apparently possible to develop immunity! The worst I ever injured myself with electronics was when I accidently touched B+. I was not injured by the shock itself, but my hand jerked away fast and ran into the sharp edge of a metal chassis box.

Tube electronics needed the high voltages, but semiconductors do not need it or "want" it. Happy days! But we still try to use most of whatever supply range we choose for S/N reasons. So while clipping is always a possibility, and can usually be anticipated, it is a foe, at first blush. However, if we are using signals that are pushing the limits, clipping can be seen as a friend; as any accidental increase will soon be automatically limited. Your sound system, for example, can't jump up too loud if you are normally operating relatively close to clipping. As suggested above, the signal may face distortion, but it may be possible to assure that the signals can't get too big.

**Considerations For Audio:**

So clipping of signals within the confines of power supply limits provides some assurance in the same sense that knowing your car is in the garage assures you of its whereabouts. We can also ask, within a specific area of application, if clipping is useful or at least interesting. In particular, we find interesting examples in music and audio.

Most obviously, the musical traditions in popular music, particularly electronic guitar "effects", show a wide variety of clipping-related and other non-linear processes which provide a richer, or at least more musically marketable sound. Here one has to begin with the acceptance of the fact that <u>something that looks like an engineering disaster produces something that practitioners find musically valid</u>. This is not really that unusual a phenomena when engineering is actually applied. So there are numerous effects devices such as "fuzz boxes" (essentially intentional clipping) that enjoy widespread use.

Perhaps in a related case we find examples or at least claims that some sort of mild non-linearity enriches either live sounds or synthesized sounds. Those who prefer "tube power amplifiers" often claim such subtle effects, and attempt (usually unsuccessfully) to characterize the effect by terms (words only) such as "warm".

In musical synthesizers we have a special case where the instruments are designed by electrical engineers (degreed or not) who are heavily schooled in linear circuits and notions that a circuit should perform as designed and <u>inside limitations known to them</u>. But not always do engineering specifications tell the whole story. One example is the Moog Four-Pole-ladder voltage-controlled filter, often said to have some almost mystical preferred sound which some attribute to the non-linearities of the control device (the dynamic resistance of the base/emitter terminals of the controlling transistors). Another is the reported <u>failure of a new version</u> of a famous commercial phasing effects device when the voltage-controlled resistors (formed from FETs - with non-linear properties) were replaced by much more linear OTA's. In my own experience, in testing a VCF with voltage-controlled Q, I was delighted to <u>hear</u> a very good result, but disappointed to see on the scope that it was the result of clipping. Backing off the amplitude to prevent the clipping, I could not get anything approaching the preferred sound.

Above we have tried to present a basis for understanding the results of clipping as encountered naturally and as intentionally used for effects. One particular finding is the fact that the added harmonics caused by clipping can result in an <u>apparent increase in loudness</u> (despite being <u>amplitude limited by clipping</u>). Certainly we understand that this, in conjunction with any actual jump to self-oscillation of a filter (with its accompanying "ugly" waveshapes), can result in an annoying aural outcome overall.