ELECTRONOTES 216

Newsletter of the Musical Engineering Group

1016 Hanshaw Road, Ithaca, New York 14850

Volume 23, Number 216 July 2013

REVISITING PULSE-WIDTH MODULATION

INTRODUCTION:

Designers of music synthesizers use the term “modulation” in a sense that is usually different from the origins of the term in engineering (in communications), or in music for that matter. In communications, we modulate to make a signal transmittable, and I think without exception, we envision a corresponding demodulation to recover a program signal. (In music, modulation means changing key.) In music synthesizer design, we use modulation techniques to produce a complex spectrum which we then listen to in its modulated form. Hence we have AM, FM, balanced-modulation (“ring” modulation) and even pulse-width modulation (PWM). All these are extremely useful and likely essential. They seems as important as filtering.

In a recent application note, AN-397 [1], for purposes of calculating power, we looked at a pulse train from which we removed any DC bias. The traditional PWM approach, as used in communications, requires this sort of bias (frequencies low relative to the “carrier”) – in fact, this bias IS the signal that is to be demodulated. Our music applications are different.

From the point of view of music synthesis, we are interested in complex and dynamically time-varying spectra. In a few instances, the modulation process may imitate a corresponding physical process for producing sounds, but other than birdcalls, good examples do not jump to mind. Instead, the use of modulation is an efficient way of obtaining an interesting signal that is sufficiently like a traditional acoustic instrument to be of practical use. Nothing magic.

Of the traditional synthesizer waveforms, sine, square, saw, triangle, and pulse, the pulse is relatively unique in having an actual parameter (the so-called “duty cycle” or ratio of time high to time low). A pulse of 50% duty cycle is just a square wave (and has zero DC incidentally).

By changing the duty cycle, the spectrum changes, so we can get an overall effect similar to filtering. On the one hand, we could think of PWM controlled by a simple panel pot to thereby provide a continuous selection of waveshapes by setting a "reference level". But more importantly, the reference level is just a voltage, and can be controlled by another oscillator or an envelope. Thus a dynamically changing spectrum, ongoing throughout an individual musical note, becomes available at low cost (relative to using a voltage-controlled filter - VCF).

So what is the deal about the DC level? After all, we can't even hear DC. Actually in this case we are saying we can't hear the low-frequency such as the controlling envelope. But perhaps it does make a difference when we have a controlling envelope with a sudden onset (as is very often used for a percussive or impulsive sound).

Here we will briefly review PWM and then consider the spectra that result from a DC term. It turns out that removal of the DC is quite trivial. Is it worth doing? Worth an experiment.

When we design a PWM capability, we possibly give it too little thought. The basic ideas are well established [2]. A couple of interesting variations on algebraic/logical waveshapers were published [3, 4] and Lester Ludwig and I wrote a series of three extensive articles on PWM: review and research [5-7]. Here by way of review we need to discuss why there are differences between the use of up-saw, down-saw, and triangle (possibly best?) to drive against the reference. We will compare the triangle to a sawtooth (possibly giving a slight shift to the spectrum), or a sinewave, but certainly not a square! By way of something new, we need to understand that the DC bias has an interesting effect (at least mathematically).

ONE WAVEFORM HARD-CLIPPING AGAINST A REFERENCE LEVEL

Fig. 1

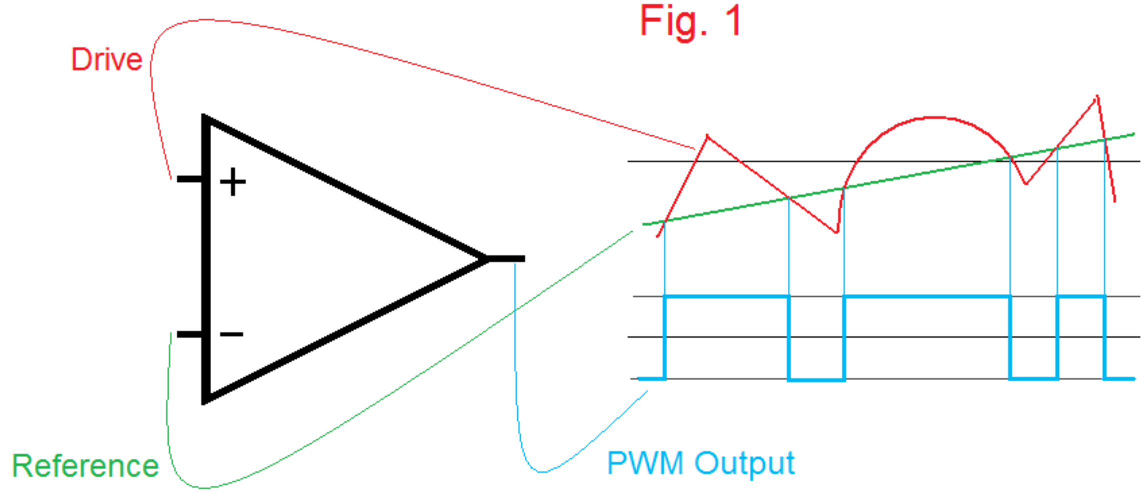


Fig. 1 reminds us of just how simple the PWM process is – just a comparator or open-loop op-amp. Here we show a "Drive" voltage and a "Reference" voltage although these are largely interchangeable. When the drive voltage exceeds the reference level, the output of the comparator is high, otherwise it is low. In many VCOs that have PWM capability (probably all of them) we often think of the drive as one of our core waveforms such as a triangle or saw. The reference could be as simple as a pot employed as a voltage divider, but it is usually itself obtained by an op-amp summer that may adjust a weighted sum of an initial voltage from a voltage divider pot, an envelope, and perhaps another oscillator output.

In the case where the reference is just an envelope (with an offset), where the envelope is a note-by-note event, perhaps taking 1/10 second to several seconds to evolve, the reference is relatively stationary and has little effect on the spectrum <u>OTHER THAN</u> what happens when we change pulse width. We know that a pulse of duty cycle d=1/b is composed of all harmonics, but the b^{th} harmonic and multiples of the b^{th} harmonic are nulls. In general, b will not be an integer, but there is still a mathematical null in the spectral envelope at b, and adjacent harmonics are expected to be weak. The mode of spectral analysis here is just the Fourier Series.

Thus PWM can occur in an envelope-like mode, much as FM could be used to give a "chirp" of pitch. The result is quite similar to filtering with a VCF which is responding to an envelope. This is common practice. Because we have a preference (in general – the "marketplace" being the boss) for individual notes of fixed pitch, we seldom use FM with an envelope to "sweep" a VCO during an individual note. Instead FM is used to dynamically <u>vary the spectrum as an individual note progresses</u>. In this regard, FM and PWM are used in a similar manner <u>when an envelope is employed</u>.

A NON-ENVELOPE REFERENCE SIGNAL

While less used, we can do PWM with a periodic signal (instead of an envelope) at audio frequencies, much as we use FM or balanced-modulation. We use the PWM to set a base spectrum – essentially a new waveform (often this new waveform is not harmonic). In this case, unlike the envelope, the reference signal is rapidly changing and it not to be treated quasi-stationary.

In our previous work on PWM, particularly reference [6] from which several figures will be reproduced below, we looked at this case. Fig. 2 [6] shows a drive waveform being a sawtooth of frequency 8f while the reference is a sawtooth of frequency f. Both are thought of as being audible in their own rights. For example, a 150 Hz saw and a 1200 Hz saw would be such a set. Here it is important to note that we have chosen specifically a small integer ratio of 8:1. This makes it possible to analyze the results by Fourier Series, the full-length waveforms in

Fig. 2

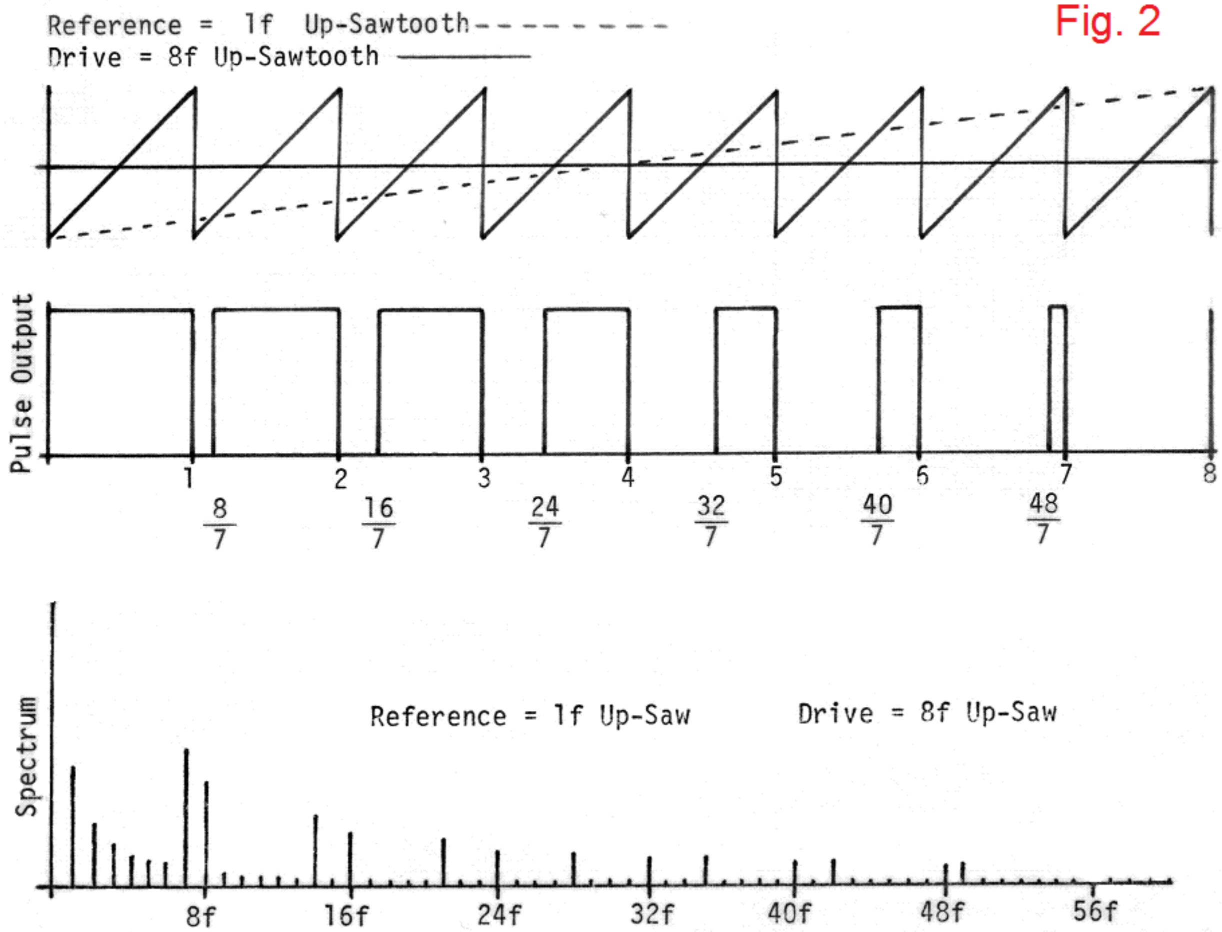
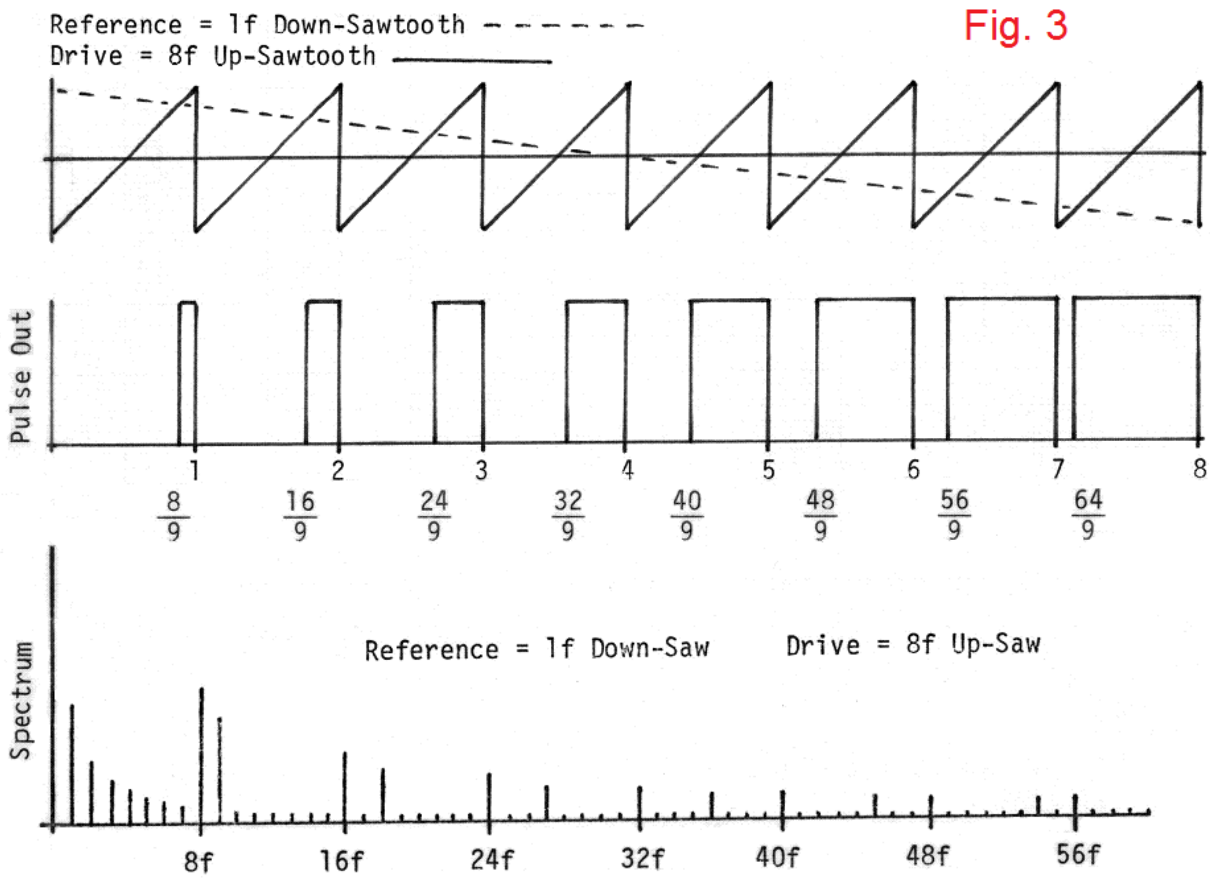


Fig. 2 being perfectly periodic with frequency f. This is for convenience. The result if the ratio were 8.0345:1 would sound similar [see also note at end]. Nothing aurally dramatic need happen when something numerically dramatic happens. (Just as nothing aurally dramatic happens with through-zero FM at the instants the modulation goes through zero.) In setting an exact 8:1 ratio we have simply chosen to observe the situation through a convenient Fourier Series window rather than strain to see through a wall.

So what do we have in Fig. 2? At the top we have the drive sawtooth (solid) and the reference sawtooth (dashed). In the middle, we have the resulting pulse – the output being high whenever the drive exceeds the reference. These waveshapes repeat periodically: we have noted that the exact 8:1 ratio was only for ease of computation. The pulse waveform in the middle is thus periodic with period 8, and can be analyzed by Fourier Series, which is greatly simplified for the pulse [8]. The bottom of Fig. 2 shows the spectrum as obtained by the Fourier Series. There are a lot of components. We kind of expected to see the harmonics 8f, 16f, 24f, etc. which we might attribute to the 8f drive sawtooth. Likewise, there seems to be a series f, 2f, 3f, etc. which must have come from f reference sawtooth. But where did the strong series at 7f, 14f, 21f, etc., a difference frequency, come from?



The first thing anyone is likely to do when encountering some strange result is to change things around a bit to see what difference that makes. Here we make the reference wave a down sawtooth rather than an up sawtooth, and the result is shown in Fig. 3 [6]. Things look somewhat similar, but there is an important difference. We still have harmonics of 8f, and the harmonics of f, but now the "extra" strong harmonics apparently come from 9f, 18f, 27f, etc., which is a sum frequency, in contrast to Fig. 2.

In ones sense, the mathematics is telling us the whole truth. But it is always interesting to see if we can understand things in more than one way – indeed in as many ways as possible. So when we see the spectrum leaning down in Fig. 2, and up in Fig. 3, we might well look at the pulse waveform more closely. Let's consider what we will call the "center of mass" of the rectangular pulses. Despite the fact that only one side actually moves (the one on the slope of the driving waveform, the centers move in a consistent manner, further apart in Fig. 2, and closer together in Fig. 3. Indeed, Fig. 3 has eight pulses in the full cycle that has only seven in Fig. 2. Another analogy is perhaps to look at it like a Doppler shift. And this shift resets at a frequency f. So you might wonder which sawtooth relationship we should use. Perhaps neither, or is there even an audible difference [see again note at end]?

I did a listening test using Matlab. This meant putting in the pulse trains as ones and zeros for a full cycle and repeating the cycle many times. In playing back, we need to adjust for the difference sequence lengths (Fig. 2 is length 56, Fig. 3 is length 72) by scaling the sample playback rates. Result – the two sound different, although they do seem to have have the same pitch. They have different timbres. The pitch, by the way is <u>clearly f, not 8f.</u> Pretty much what we would have guessed. [Again- see notes at end.]

It is of course possible to avoid the shifting by driving with a triangle instead of one of the two sawtooth choices. This is shown in Fig. 4 [6]. <u>Perhaps we should make this as the default choice</u>.

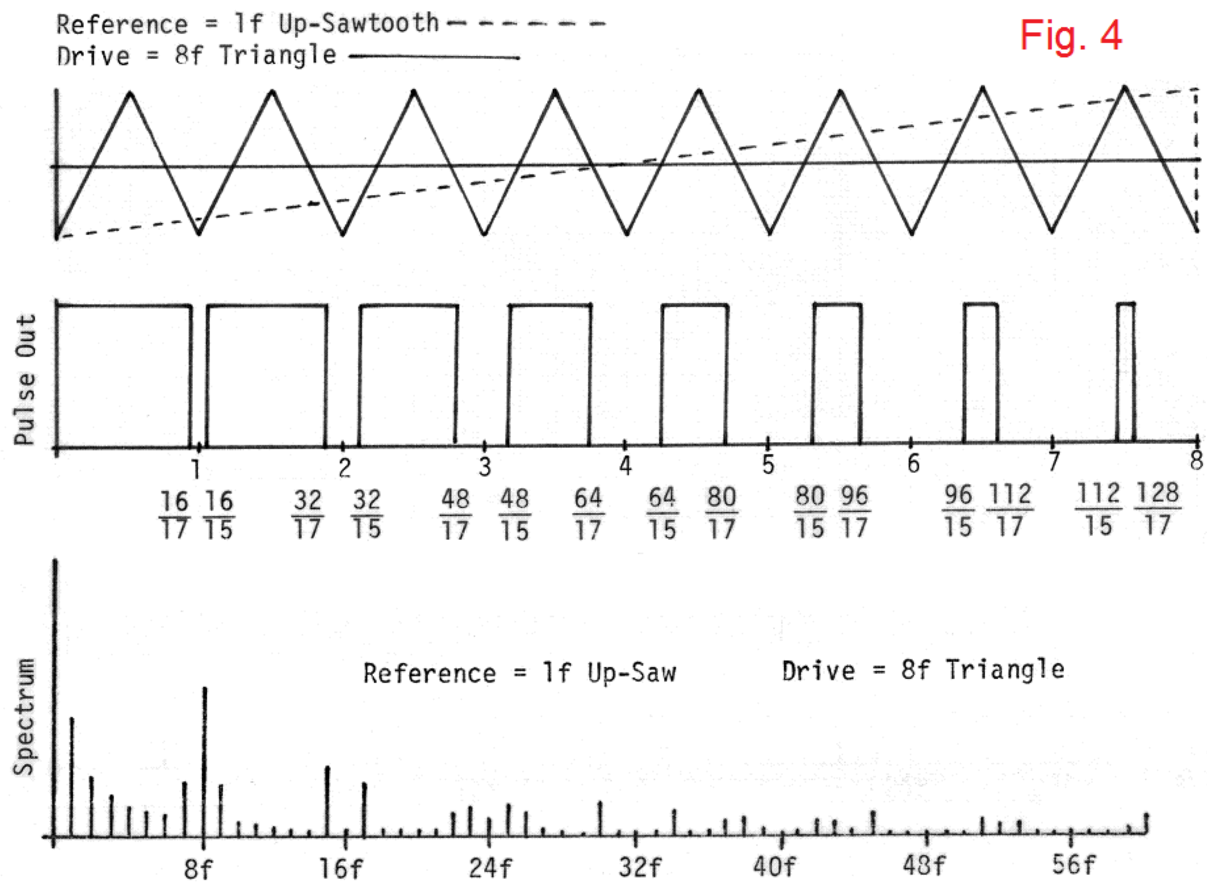
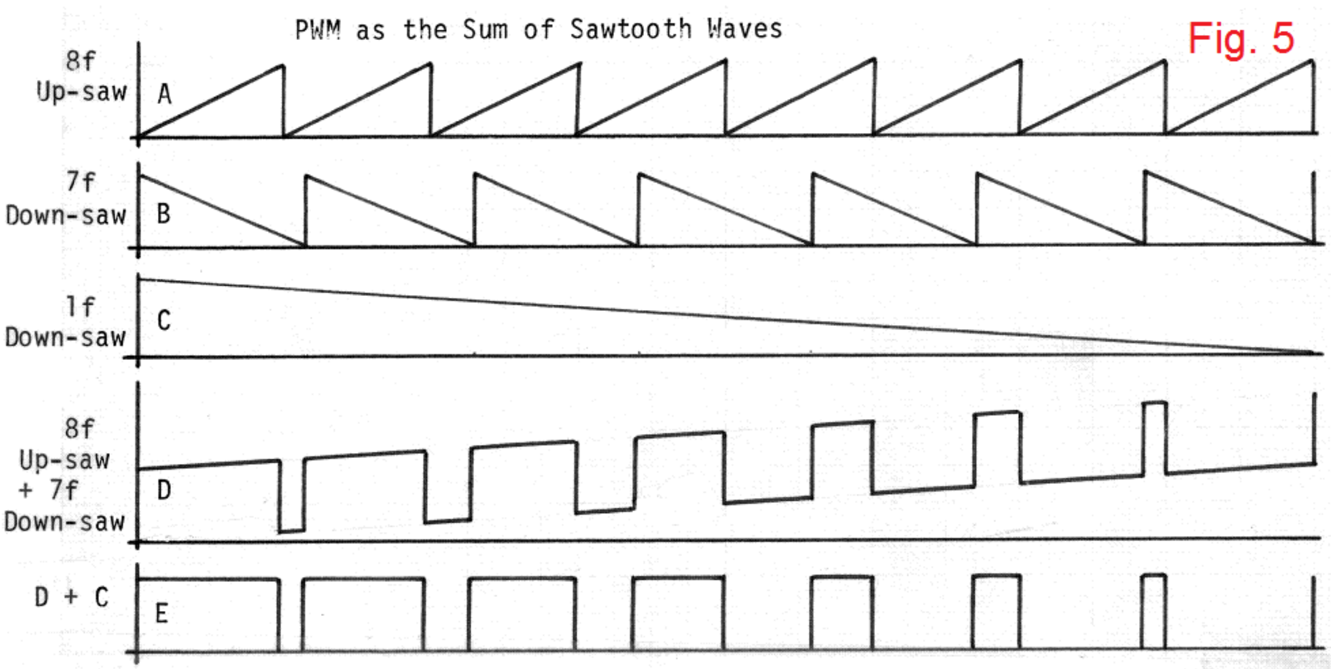


Fig. 4

Notice here that we see in the spectrum "sidebands" about multiples of 8f. And there is no shift of the "center of mass" in the time-domain view.   Here if we rush ahead to a listening test, we encounter unexpected problems. The two sawtooth cases had transition times that were all multiples of 1/7 (Fig. 2) or of 1/9 (Fig. 3). In Fig. 4, the triangle case, the times have multiples sometimes of 1/15 and sometimes of 1/17, so all are multiples of 1/255! This is too hard to program, and if we approximate, we can't be sure that if we do hear something different whether it is due to using the triangle, or of settling for the approximation. Here is a good place for an analog test with someone with an outstanding ear.



Continuing with our notions that we should try to explain things in as many ways as possible (see reference [6] for additional examples) we offer an alternative explanation of Fig. 2 in terms of a sum rather than a PWM process. Looking at the spectrum in Fig. 2 there seems to be an 8f sawtooth (harmonics 8f, 16f, 24f, etc.). There also seems to be a 7f sawtooth (harmonics 7f, 14f, 21f, etc.), and an f sawtooth (harmonics f, 2f, 3f, etc.). Fig. 5 [6] shows how the pulse waveform of Fig. 2 seems to be a properly weighted sum of these three sawtooth frequencies. This now becomes a matter of drawing the situation in the time domain. It may seem strange that a modulation process as radical as PWM (clipping) can result in a simple sum. However, every time we derive a spectrum as we did here, we of course end up with a sum (we list the components that are added together – that's what a spectrum <u>is</u> after all). [The balanced modulation trig identity is another, simpler, good example: a product being the sum of components with sum and difference frequencies.]

THE DC BIAS-FREE PULSE

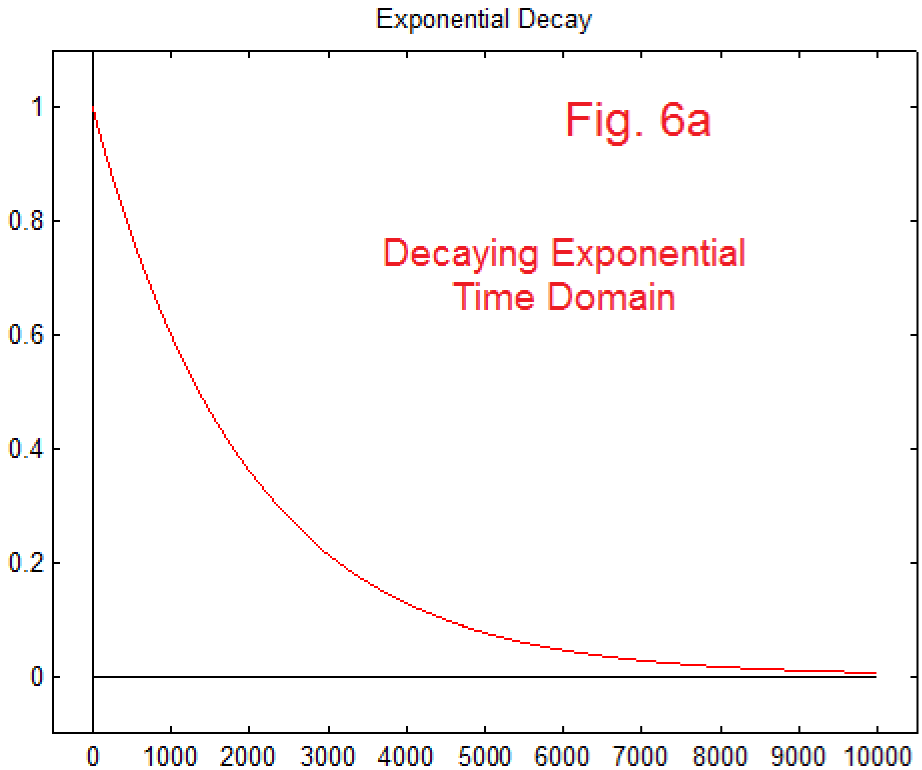
One of the things that happens when you do research and attempt to write about it simultaneously (or just think about what you are likely to write) is that the real world soon insists on taking over. The great Isaac Asimov is reported to have said: 'The most exciting phrase to hear in science, the one that heralds new discoveries, is not "Eureka" but "That's funny...".'. Niels Bohr was also a fan of arriving at a contradiction as an indication that progress was about to be made. And so on.

Here in looking at the effect of the DC bias in PWM I was prepared to write how there would be an audible difference, a tentative conclusion based on what I supposed (correctly) would be seen in a Fourier transform view. A few lines of Matlab code added to that which does the spectral computations allow us to play the waveforms. Not only did I not hear an improvement with the unbiased PWM, but essentially, no difference at all. I then convinced myself that the difference should be at least audible if I used a lower pitch. Indeed, you can just barely hear a difference if you do this. But there is not much difference, and sadly, if you ask which one you like better, it could well be the DC biased cases. Time for a new story. In essence, we have to keep in mind that the Fourier transform (FT) is very far from the last word when it comes to saying how the ear and the brain operate. For example, time-constants matter. But all this really means was that there was a more important fact to understand. But first, the math.

The FT is linear, so it is clear that the FT of a DC biased pulse is the sum of the transform of a DC signal (a sharp delta function) plus the transform of the unbiased PWM. We just add in the DC contribution. But at this point we become concerned that the actual signals we are interested in are not DC and pulse trains that go on forever. Instead, they begin and they end on a time scale of perhaps 20 milliseconds to several seconds. That is, the components that in theory extend from $-\infty$ to $+\infty$ need to be time-enveloped. Specifically here, we are looking at an instantaneous start and an exponential decay, as is appropriate for such musical examples as percussive or plucked sounds. Thus enveloped, the DC component has the exact same shape as the envelope itself (a constant, time-multiplied by the exponential envelope). The contribution of the pulse-train is the time-multiplication of the pulse-train by the exponential envelope (thus we look to convolution in the frequency domain). Thus the FT of the exponential envelope is central. Yet, keep in mind that we are not saying this is what we hear. It is just a step closer to the truth than the FT of the infinite duration signals.

THE DECAYING EXPONENTIAL

Fig. 6a shows a decaying exponential envelope as 10,000 time samples from 0 to 9,999 and generated in Matlab as t=0:9999; e=0.95.^(t/100). This is the same envelope waveform that we get by very rapidly putting a charge on a parallel RC and letting it decay. Note the instantaneous startup at t=0. This envelope is easily generated and also occurs naturally as the envelope of a decaying sinusoid (a resonator). This is very familiar.



While we are interested in the time waveform of Fig. 6a, we also need to look at the corresponding frequency domain description; that is - some notion of a FT. The simplest way to obtain this is to take the FFT of the time sequence. Fig. 6b shows the most interesting portion of this FFT. Here the time sequence is still length 10,000 so the frequencies are 0.0001 apart. We have 10,000 frequency points (half of them unique) which is a tremendous amount of resolution. Thus we show only 61 of them on either side of 0. Here we are plotting only the magnitude, which is what we are mostly interested in. It is important to remember that there is a negative side here (as shown), as we will soon be taking about convolution. From Fig. 6b, it is clear that the spectrum already falls off rapidly.

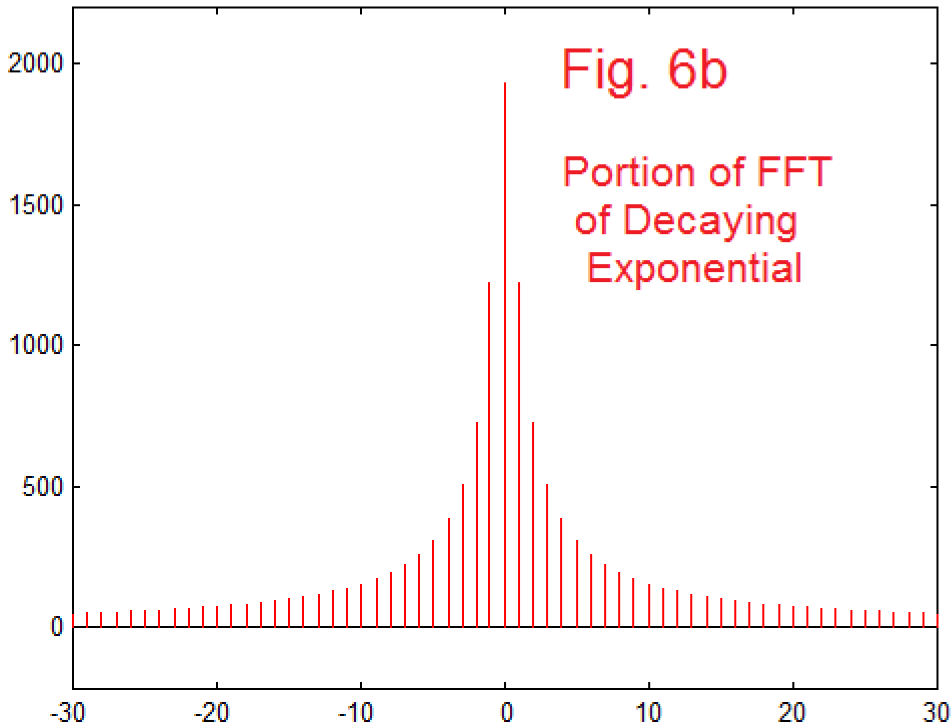


Fig. 6b is correct and tells us the main story. Still it is simply a numerical calculation from a computed time sequence. It is of interest to consider the analytic form of the FT of the decaying exponential function. This we get by using the Fourier integral transform:

$$X(f) = \int_{-\infty}^{\infty} f(t)e^{-j2\pi ft}dt = \int_{0}^{\infty} e^{-\beta t}\, e^{-j2\pi ft}dt = \frac{1}{\beta + j2\pi f} \qquad (1)$$

which was easy enough. We do have to figure out what β is here, as we have represented the exponential sequence in Matlab in terms of $g^{\alpha t}$ (where α=0.01 and g=0.95) instead of in terms of $e^{-\beta t}$. (Why is this the hardest part to get right!). Thus we want

$$g^{\alpha t} = e^{-\beta t} \qquad (2)$$

Taking the natural log of both sides:

$$\alpha t \ln(g) = -\beta t \qquad (3a)$$

or

$$\beta = -\alpha \ln(g) \qquad (3b)$$

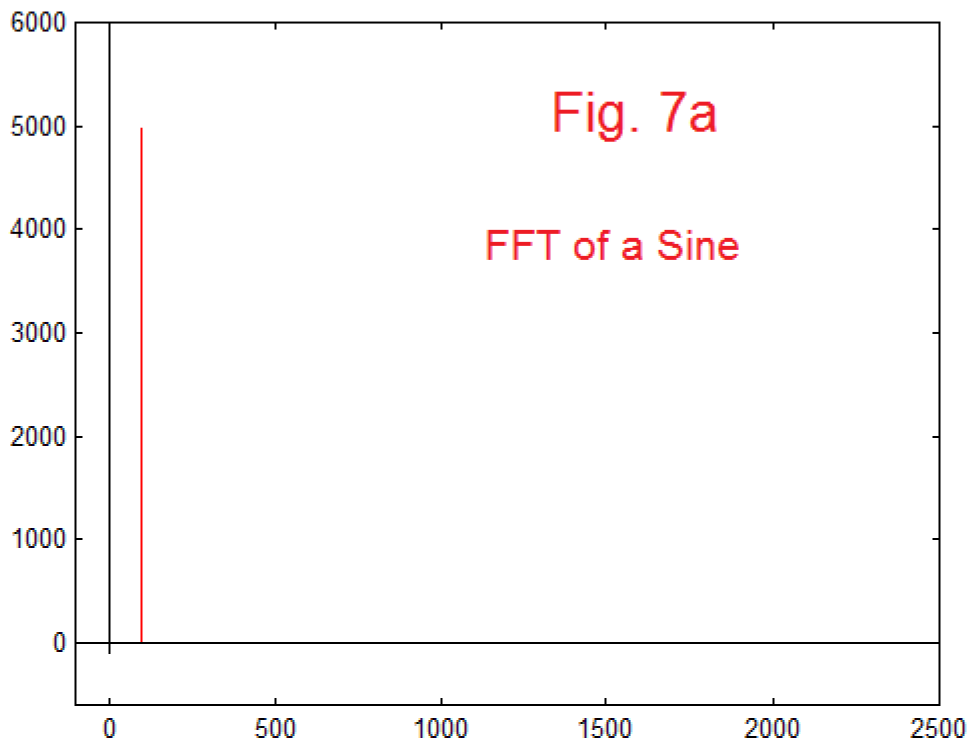
so we are in a position to evaluate equation (1) at values of f that are multiples of 1/10,000. This is the correct FT. How well does it compare to the FFT result. We know they are not supposed to be exactly the same, but here we have so many points, we expect good results. In fact, the two agree to about 1/2%, and the ratios are exact!

That is, we are essentially taking ratios from the data for Fig. 6b and dividing the results for k=1, 2, 3, 4, and 5 with the result for k=1, obtaining the ratios:

1.0000 0.5976 0.4152 0.3162 0.2548

We then use equation (1) with frequencies f = 0.0001, 0.0002, 0.0003, 0.0004, and 0.0005 and divide by the result for f = 0.0001. We get the exact same ratios to the four decimal place shown (they differ about the 7^{th} decimal place). So this part seems right. A few lines of Matlab code relating to these calculation, and as an example of the programs that generated the figures here, are shown as an appendix.

FFTs OF SINE AND PULSE TRAINS



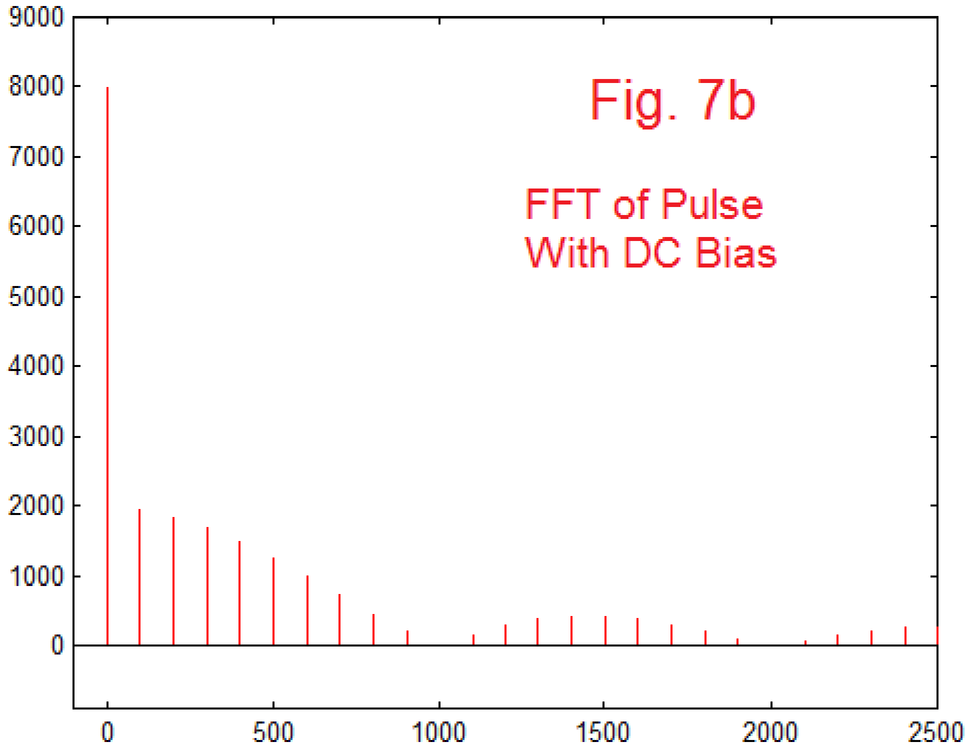


Fig. 7b
FFT of Pulse
With DC Bias
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
0
500
1000
1500
2000
2500

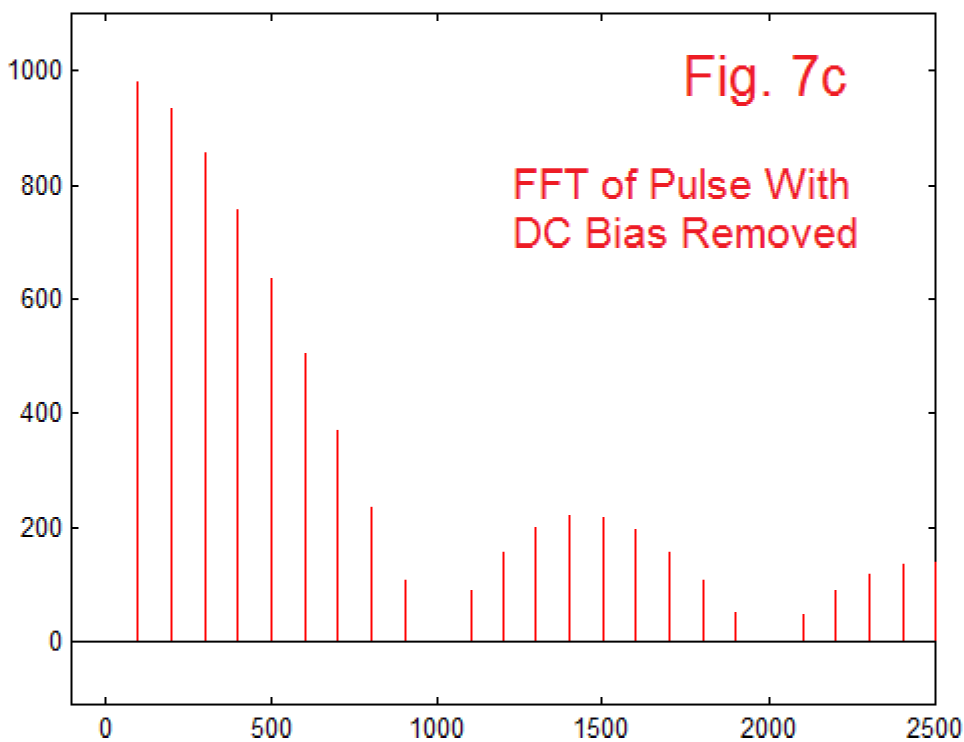
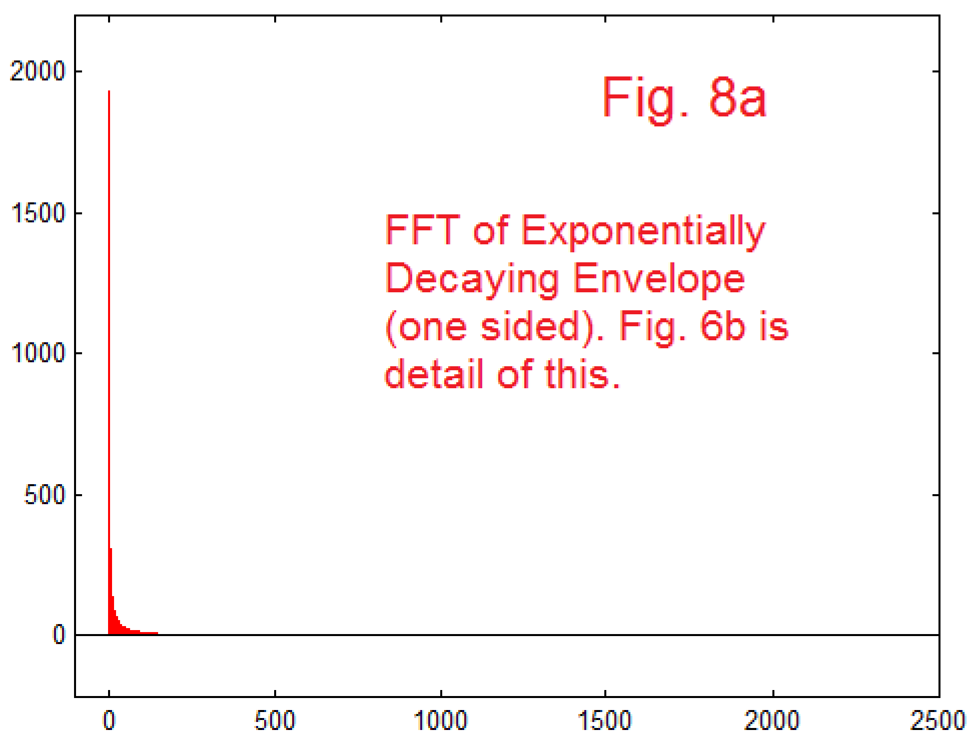


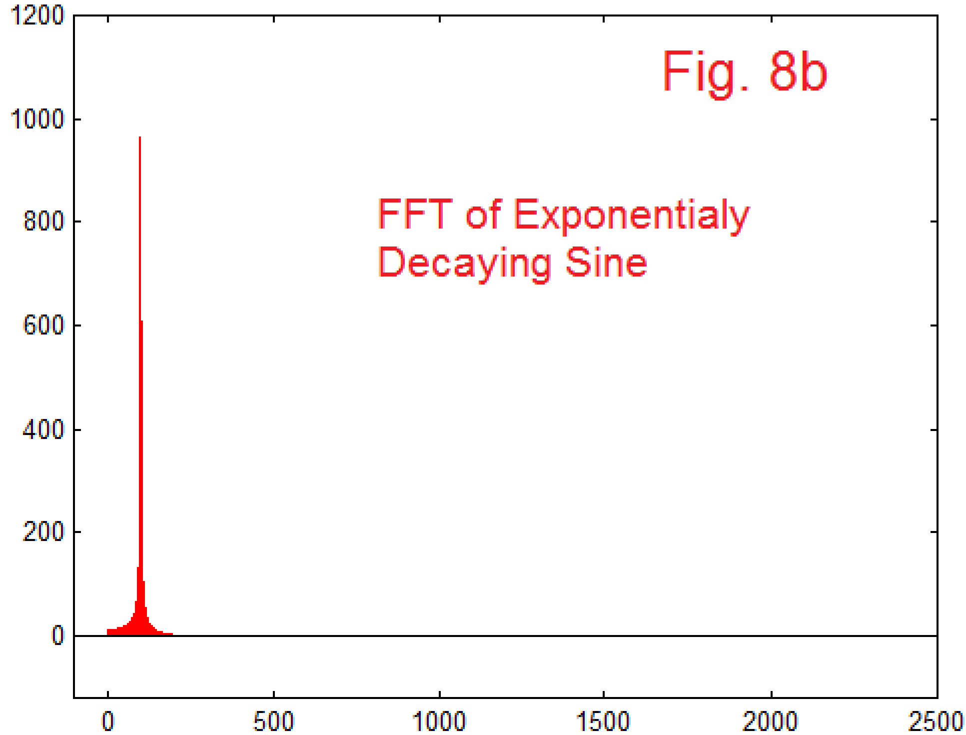
Fig. 7c
FFT of Pulse With
DC Bias Removed
1000
800
600
400
200
0
0
500
1000
1500
2000
2500

We are accustomed to using an FFT of generated data as a useful measure of the FT, and here we will look at the FFTs of a sine function (for reference) and of pulse trains with and without a DC bias. The FFT of the sine wave was for a sequence sin(2πt/100) so has a frequency of 1/100 and the FFT in Fig. 7a shows a single line at k=100. (The full sine signal has 100 cycle.) The pulse train we test with has a duly cycle of 1/10. Thus it starts out with a large DC bias of -0.8 times its amplitude. Again we choose a full period for the pulse train of 100 samples (thus a cycle is 10 +1's and 90 -1's). We expect a DC term at k=0 and harmonics at multiples of 100k. Further, we expect every 10^{th} harmonic to be zero. Fig. 7c shows exactly what we expect. In Fig. 7d we have removed the DC bias and the large spike at k=0 is gone.

CONVOLUTION

So we pretty much understand the FFT results which are essentially the same as a Fourier Series result. However, these applied to infinite duration periodic instances of the signals, and we need to deal with finite duration sequences. So we need to multiply by a windowing envelope. For practical purposes, the exponentially decaying envelope of Fig. 6a decays fast enough to be considered finite duration. We looked at a detail of the FFT of this envelope in Fig. 6b. Fig. 8a shows a larger view of the FFT consistent with further examples to come, although the lines are far too close together to be resolved in the figure. We see significant spectral content from 0 to perhaps 150 or so,





Thus it is a simple matter to multiply the sinusoidal sequence by the envelope. This will be a convolution in the frequency domain. Thus the FFT in Fig. 8a, and its negative side not shown, are convolved with Fig. 7a. This is the same as picking up the FFT or the envelope and moving it to the position of the sinewave frequency (k=100). Note the effective folding back of the negative side. Overall, the convolution is a smearing. What this means of course is that a sinewave enveloped by an exponential (or otherwise time-limited for that matter) is no longer a pure frequency. This is well-understood.

We move on to applying the exponential envelope to the two cases of the pulse train, again invoking convolution, with the result that the two figures (Fig. 8c and Fig. 8d) are understood in terms of superposition – each of the original components in Fig. 7b and Fig. 7c is present but smeared. Other than some changes of vertical scale, that's all that happens because of the enveloping.

Now, comparing Fig. 8c with Fig. 8d (ignoring the vertical scale) we see that there is only one significant difference, the smeared DC term. All this was totally predictable – we knew in general how this had to turn out. Now we must concentrate in the lower left corners of these two figures. The details of these corners are shown in Fig. 9a and Fig. 9b, where we see a fairly dramatic difference that would seem to be important. So this is probably an appropriate point to say that what we see in the mathematics is not necessarily what we hear.

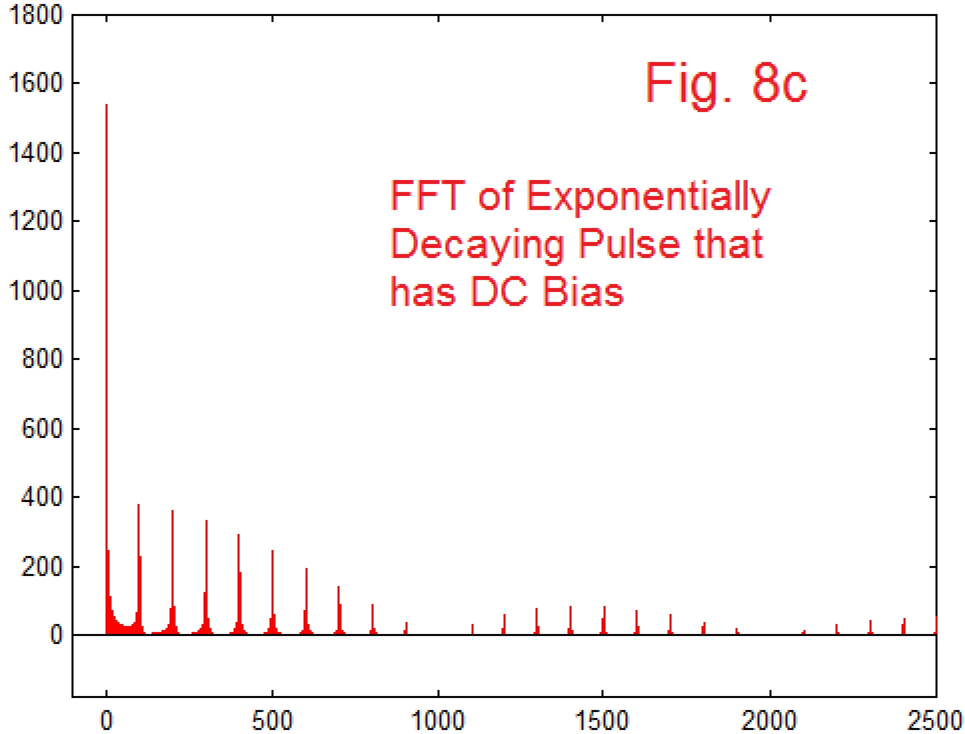


Fig. 8c
FFT of Exponentially
Decaying Pulse that
has DC Bias
1800
1600
1400
1200
1000
800
600
400
200
0
0
500
1000
1500
2000
2500

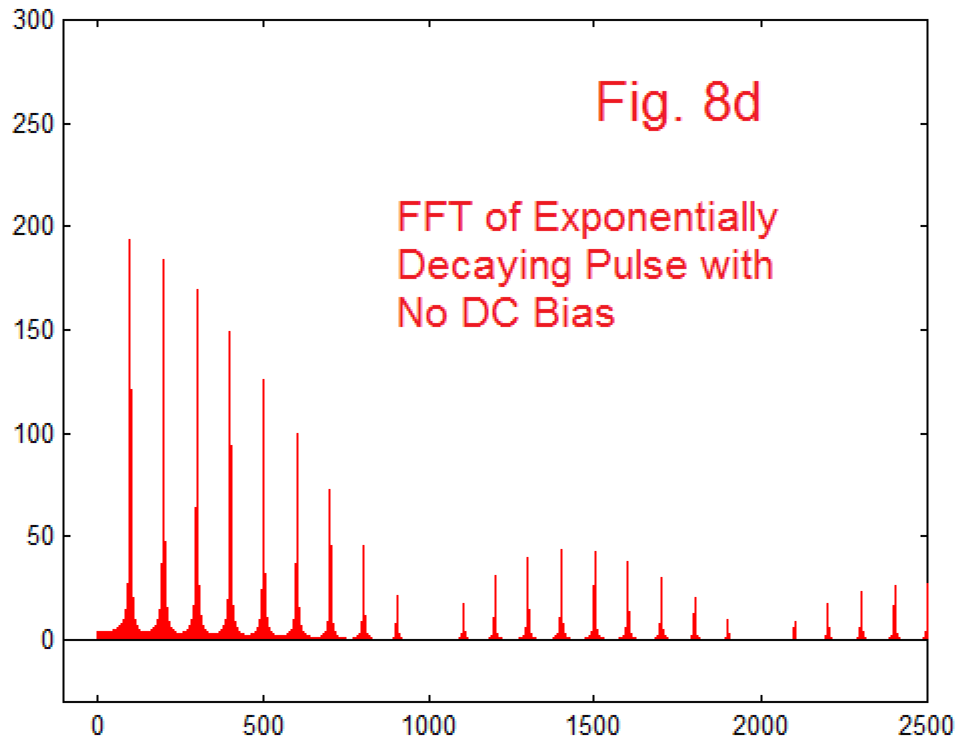


Fig. 8d
FFT of Exponentially
Decaying Pulse with
No DC Bias
300
250
200
150
100
50
0
0
500
1000
1500
2000
2500

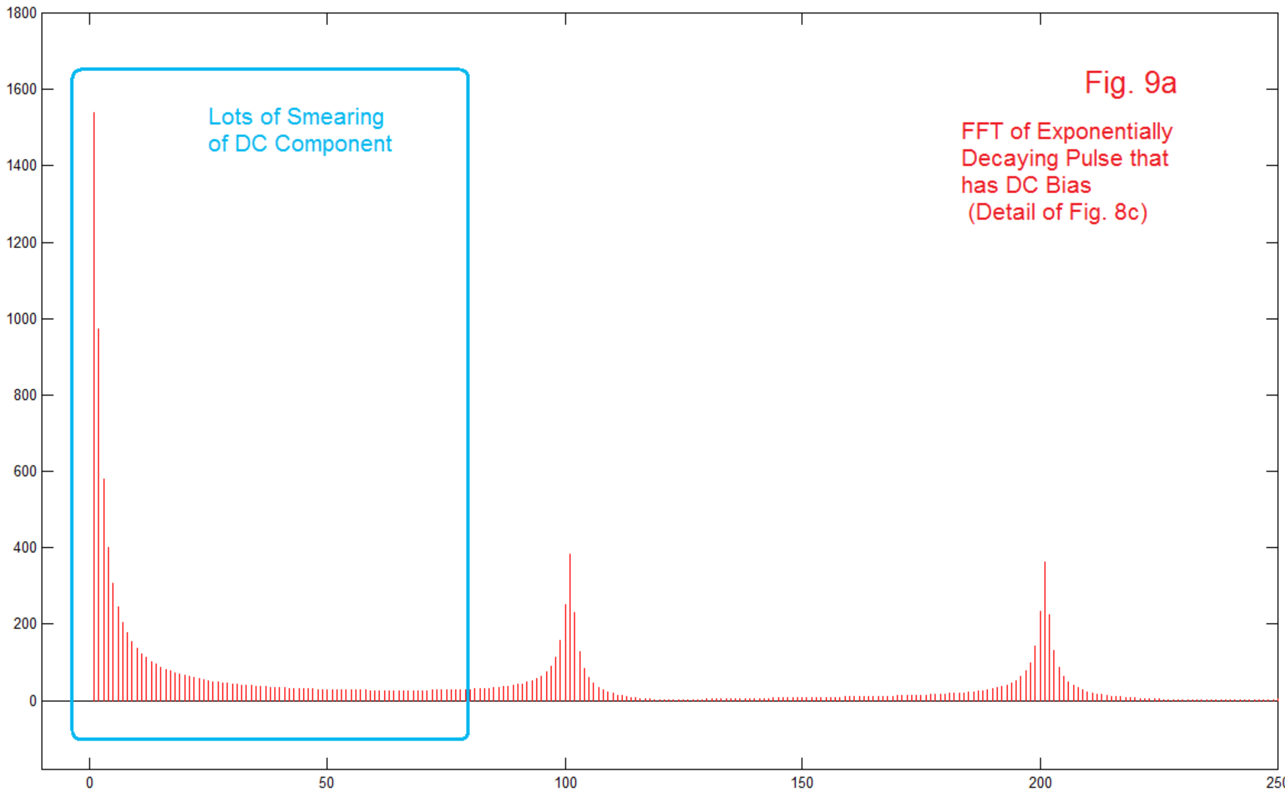


Fig. 9a
FFT of Exponentially
Decaying Pulse that
has DC Bias
(Detail of Fig. 8c)
Lots of Smearing
of DC Component
1800
1600
1400
1200
1000
800
600
400
200
0
0
50
100
150
200
250

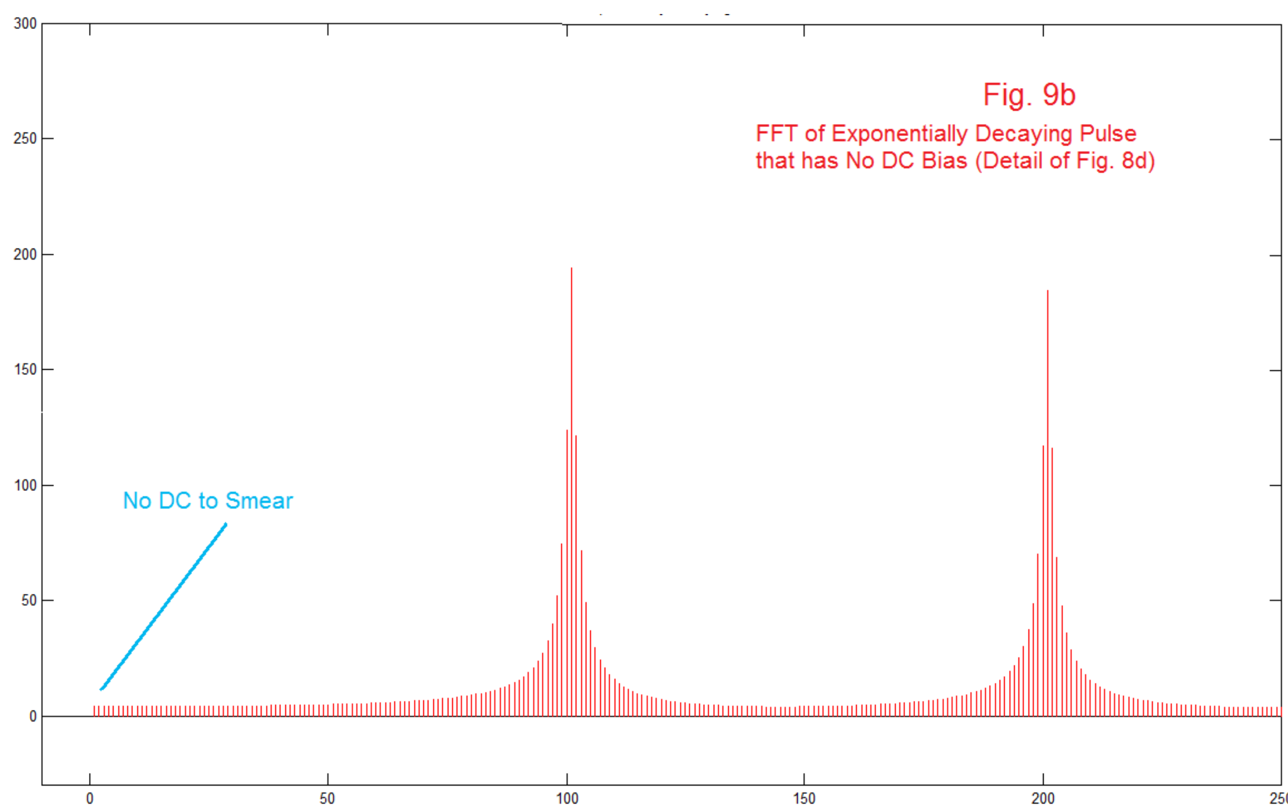


Fig. 9b
FFT of Exponentially Decaying Pulse
that has No DC Bias (Detail of Fig. 8d)
No DC to Smear
300
250
200
150
100
50
0
0
50
100
150
200
250

THE DC COMPONENT SMEARED TO THE AUDIBLE RANGE

In Fig. 9a and Fig. 9b we get to look at the clutter in the lower left corners (again do not be concerned with vertical scales – perhaps use the peak around k=100 as the common reference. We know that frequencies around k=100 are audible – indeed k=100 is the fundamental frequency of the tone. So while k=0 is still inaudible DC, there are certainly frequencies below k=100 that are audible, and these are clearly broadbanded (lots of adjacent peaks of similar height). In Figure 9a, there is a lot of content there (the blue box). From Fig. 9b we see that some of this content is the smearing downward of the k=100 peak. But in Fig. 9a it is clear that for low values of k, there is the strong contribution from the DC (k=0) term being smeared up.

It would seem that the case with the DC bias has more low-frequency broadbanded content, and one might suppose that this would be audible (and objectionable). Another point might be that if we had higher pitches, the smeared DC term might be less audible. So if it turns out that we don't really hear a difference at a given level of pitch, perhaps we will hear it at lower pitches. As stated earlier, the two sound much the same if not identical. At lower levels of pitch (by adjusting the playback rate) there is a slight difference. But if anything, I like the DC case better as having a bit more "punch" which is what we are likely after with percussive sounds. So much for good ideas. But experiments don't lie.

WHY DO WE HEAR NO SIGNIFICANT DIFFERENCE?

This is a question I cannot answer for sure. Knowing the truth helps support speculation. If we suppose that the FT is what we actually hear, then equation (1) introduces a ridiculous notion that everything from $-\infty$ to $+\infty$ is included. Even with a finite length envelope (or a sufficiently fast decaying exponential) we can't say the ear includes everything that is non-zero. It is known that the ear has a "time-constant" of something like 50 msec. This is fortunate. If not for the fact that the ear "forgets" older excitations, we would only hear accumulating noise. It is reasonable to suppose that there is some sort of forgetting envelope that fades out older stuff. By older we are means something older than perhaps 200 msec.

So consider a signal with a sudden onset. This arrives at the ear suddenly at high level and then does something else, like perhaps decays. However, for a time period similar to the ear's time constant, there is <u>no notion of what is going to happen</u>. Everything is more or less just coming in through the newly opened door.

Fig. 10

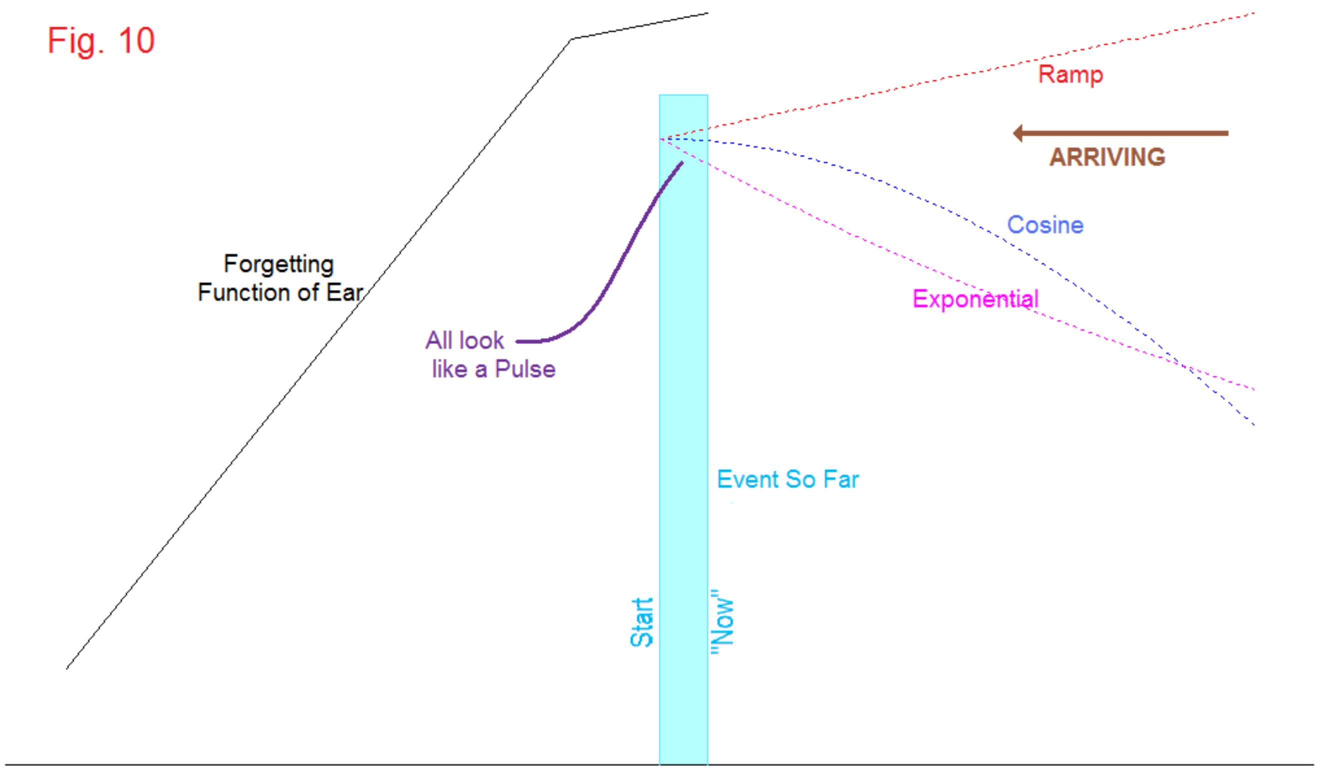


Fig. 10 suggests the sort of thing that is apparently going on. Here we suppose that we have three possible different waveforms arriving from the right side at high level**:** a cosine, a decaying exponential, and an upward ramp. The blue box in the diagram indicates an existence window labeled as "Event So Far" with the left side being the start of the event and the right side is "Now". As pictured, the blue event box is considered to be somewhat more narrow than the forgetting function so that very little of the waveforms is <u>not</u> still in play. Eventually, as the "Start" side of the box moves to the left, the start of the waveforms will begin to be forgotten in a tapered manner. Clearly there is no information at time "Now" as to what is going to happen immediately to the right of "Now". For all we know, the three waveforms may plunge instantaneously to the zero baseline (black at bottom). Accordingly, <u>all three waveforms are roughly equivalent, and equivalent to a pulse</u> the same as the blue window itself.

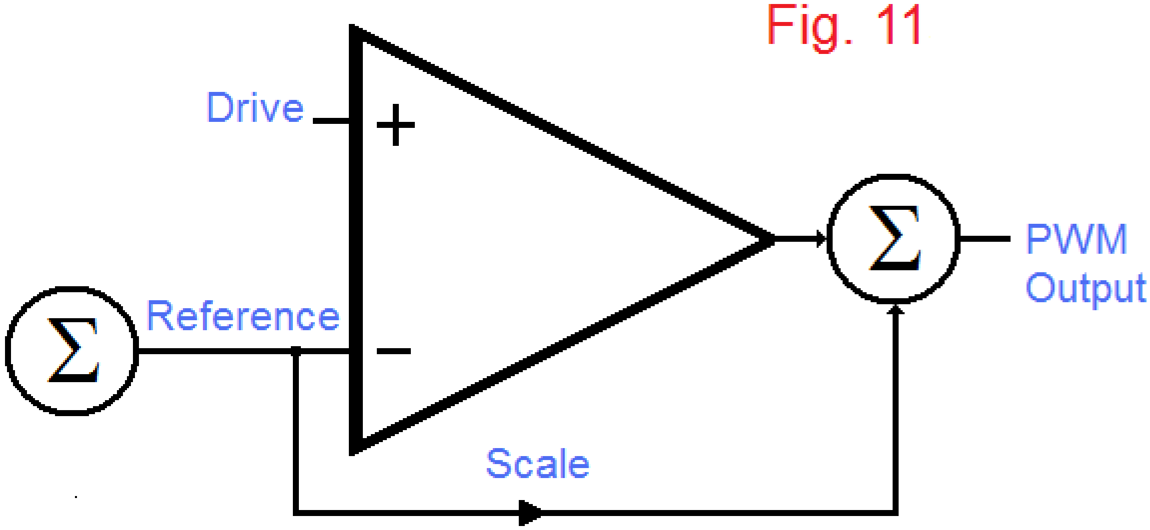
So all three waveforms are just a narrow pulse, the sort of onset click we are familiar with, such as when we plug a set of earphones into a jack or otherwise make an instantaneous switching. The click is annoying. Yet we apparently have a tolerance for this click if it is in the context of being followed by an exponential (or similar) decay. The characterization as a click rather than as a sudden onset followed by a decay seems to be delayed. Possibly this is because we deal with so many percussive-like event in normal hearing as well as in music.

This is not much more than speculation.

CIRCUITRY FOR REMOVING DC BIAS

We have NOT made the case for removing the DC bias because there seems to be no objectionable audible effect of the bias of the PWM signal itself. [If however, this were used in a double modulation, like PWM driving FM, we could well be concerned with envelope feed-through.] We do note that it seems to be straightforward to remove the bias, as indicated in Fig. 11 by scaling and adding the reference level (details follow). Note that in Fig. 11 (and see below) we are adding the reference level, because as we have set the inputs to the comparator, a higher reference level means less positive extent, and thus a lower DC bias. (Reversing the (+) and (-) inputs would call for subtraction.) As shown, the DC bias level turns out to be the negative of the reference.

Fig. 11



An objection to implementing this would seem to be the need for an output summer, an additional op-amp. Note that the summer at the input setting the reference is commonly there already. However, we recall that it is also common to scale, and simultaneously set a PWM output impedance to a standard 1k value with a 3k and 1.5k voltage divider (e.g., Fig. 13b). What if instead of setting the lower leg of the divider to ground we return it to the reference. This double voltage divider [9] allows us to get a summation without an op-amp. We do need to position this divider between available op-amp outputs or else we defeat the purpose of not having to add an op-amp. And we did want to establish that output impedance of 1k. There is a price to pay – a reduction in output level. But there is no extra op-amp.

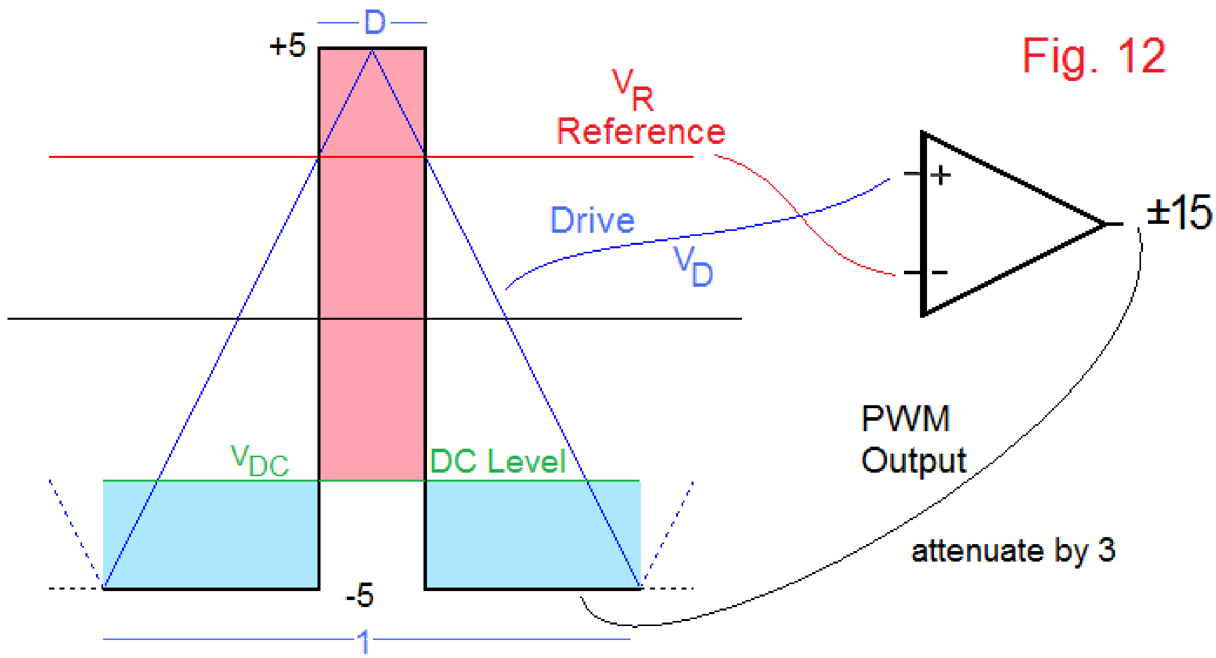


Fig. 12 shows the relationships between the drive (blue), the reference (red), and the output of the comparator/op-amp (black), along with the DC level (Green). Note that the black waveform is high only when the blue waveform exceeds the red line, and is otherwise low. With reference to the green DC level line, we note portions more positive than the green line (pink) exactly balance portions more negative (light blue), which is the definition of a zero DC level. Here we are assuming that the levels of the actual op-amp (for convenience taken to be ±15) are attenuated to the amplitude of the drive (typically ±5).

We need a few simple equations here. First we want to establish the relationship between the reference level (V_R) and the duty cycle D. This applies to a drive with straight-line segments such as the triangle shown (or it could have been a saw). When the reference V_R is at -5, the drive is always above and the duty cycle, the ratio of high to total cycle length (1 here) is D=1. When the reference is 0, D = 1/2, a square wave. When V_R reaches +5, the drive is always below V_R, and D=0. As drawn in the example, V_R = 3 and D=1/5. In general, the relationship is linear for straight-line drive segments as:

$$D = \frac{5 - V_R}{10} \tag{4a}$$

which is the fractional time the PWM waveform is high. The fractional time it is low is thus:

$$1 - D = \frac{5+V_R}{10} \tag{4b}$$

The DC level of the PM output is the high level multiplied by D plus the low level multiplied by (1-D) all divided by the length which is 1 in this case:

$$V_{DC} = 5D + (-5)(1-D) = 10D - 5 \tag{5}$$

and when we substitute (4a) and (4b) in here, we get simply:

$$V_{DC} \;=\; -V_R \tag{6}$$

Thus we have the simple result that the DC level of the PWN output is the negative of the reference. This seems believable based on careful examination of Fig. 12.

We don't always expect such a simple answer. <u>We just need to add the reference to correct the DC bias</u>. If this were addressing a serious problem, we would not hesitate to try something like Fig. 13a, which would cost us an op-amp. (Note that the resistor marked 300k here would likely be adjusted down to allow for the fact that the op-amp output levels may be about a volt lower than the ±15 supplies.) Is there, however, an easier way to remove the DC bias? Yes, but we get some attenuation too.

Fig. 13b reminds us of the usual way the op-amp comparator output is reduced from ±15 to about ±5, while at the same time establishing a standard 1k output impedance. We use a voltage-divider. There are two resistors, and it is the ratio that maters for the divider – hence the 2:1 ratio of resistors. But there are plenty of 2:1 choices. We select exactly the 3k**:**1.5k choice because the parallel combination of these two is exactly 1k. This starts us thinking – can be connect the lower leg of the divider to the reference and add it back in using a voltage divider?

Fig. 13c reminds us of the case where we are interested in the voltage that is at the junction of two resistors when both the ends of this series chain are driven. Indeed, this yields to superposition [9], and we get the equation shown in the figure. What we get is not a summer. It is a weighted attenuator. Thus we can relatively weight the voltages, but must also make them smaller. So we have one signal that is ±15 and another that is ±5. This calls for a resistor ratio that is 3:1. The parallel combination of resistors in a 3:1 ratio that had a resistance of 1k is 4k and 1.33k, as seen in Fig. 13d.

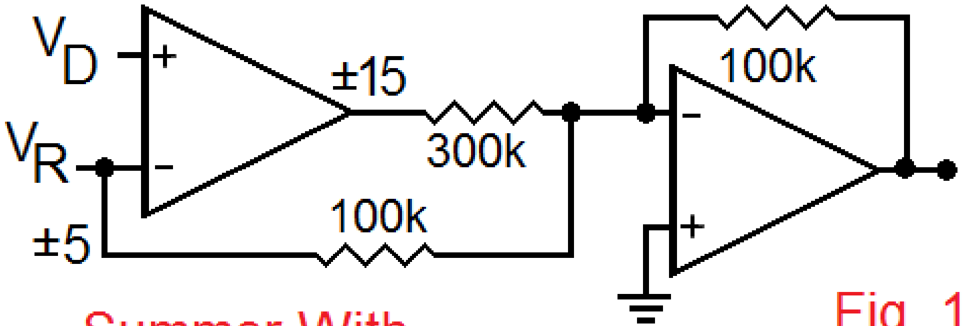


Fig. 13a

Summer With
Extra Op-Amp

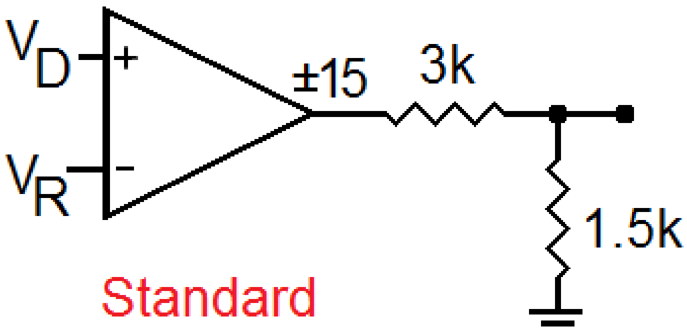


Fig. 13b

Standard

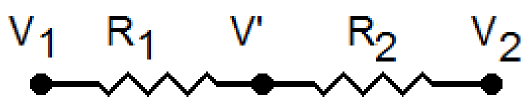


Fig. 13c

$$V' = (V_1R_2 + V_2R_1) / (R_1+R_2)$$

Generalized Voltage-Divider

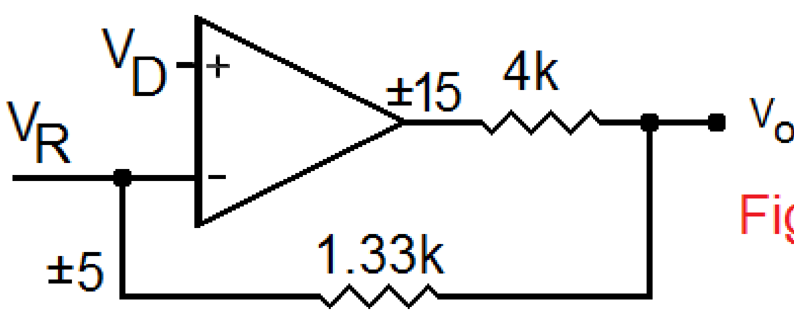


Fig. 13d

Modified for 0 DC Bias

Let's just check this. The output V_o of Fig. 13d , when the op-amp is at +15 is (with R = 4k/3):

$$V_o = \frac{15R + V_R(3R)}{4R} = \frac{15}{4} + \frac{3}{4}\,V_R \tag{7a}$$

and when the output is -15 it is:

$$V_o = \frac{-15R + V_R(3R)}{4R} = -\frac{15}{4} + \frac{3}{4}\,V_R \tag{7a}$$

Then using equations (4a) and (4b) for D and (1-D) we can check the DC level here:

$$V_{DC} = (\tfrac{1}{2} - \tfrac{V_R}{10})\,(\tfrac{15}{4} + \tfrac{3}{4}\,V_R) + (\tfrac{1}{2} + \tfrac{V_R}{10})\,(\tfrac{-15}{4} + \tfrac{3}{4}\,V_R) \tag{8}$$

Unlikely as it might seem, this comes out to 0. We did cancel the DC. Note however that the difference between the high level and the low level is 15/4-(-15/4) = 15/2 = 7.5 so the amplitude of the pulse is not 10 volts peak-to-peak but 7.5 volts peak-to-peak. Yes, we do need to settle for an attenuation of 3/4 in this case.

If I were still building, I would add this as a feature so that I could use the oscillator and compare the results in more than just the Matlab test. We can still keep the original method (Fig. 13b) for comparison. Perhaps an extra output jack for the bias-free PWM output.

REFERENCES:

[1] B. Hutchins, "Clipping: Friend or Foe," Electronotes Application Note No. 397, June 15, 2013
http://electronotes.netfirms.com/AN397.pdf

[2] B. Hutchins, Musical Engineer's Handbook, Portion of Chapter 2c on PWM.
http://electronotes.netfirms.com/MEHPWM.PDF

[3] I. Fritz, "Reader's Circuits" (Double-Pulse Waveform Shaper and two others), ***Electronotes***, Vol. 8, No. 72, Dec. 1976, pp 21-22

[4] W. Hartmann, "Comments on the Triangle-to-Sawtooth Conversion Problem," ("Symmetrized Ramp"), ***Electronotes***, Vol. 8, No. 67, July 1976

[5] B. Hutchins, "A New Look at Pulse Width Modulation – Part 1", **Electronotes,** Vol. 11, No. 108, Dec. 1979, pp 9-16

[6] B. Hutchins and L. Ludwig, "A New Look at Pulse Width Modulation – Part 2", **Electronotes,** Vol. 12, No. 110, Feb. 1980, pp 3-17

[7] B. Hutchins and L. Ludwig, "A New Look at Pulse Width Modulation – Part 3", **Electronotes,** Vol. 12, No. 118, Oct. 1980, pp 3-24

[8] B. Hutchins, "Efficient Calculation of the Fourier Series of Rectangular Functions" Electronotes App Note 160 Jan 11,1980
http://electronotes.netfirms.com/AN160.PDF

[9] B. Hutchins, "Node Voltages by Superposition", Electronotes App Note 394, Jan. 20, 2013
http://electronotes.netfirms.com/AN394.pdf

Appendix – Program

Below is a runable program that is used (1) to show the general coding that produced many of the graphs here, and (2) to show the code that did the verification of the FT and FFT results.

```
% Example Matlab Code for EN#216

t=0:9999;
g=0.95
e=g.^(t/100);
e=[ones(1,1000),zeros(1,9000)];
```

```
s=e;
S=abs(fft(s));
figure(11)
plot([-540 540],[0 0],'k')
hold on
plot([-1000 100000],[0 0],'k')
length(S)
S= S(1:251);
S=[ S(251:-1:2), S ];
length(S)
for k=1:500
   plot([k-251,k-251],[0,S(k)],'r')
end
hold off
axis([-30  30  -220  2200])
title('Detail - FFT of Exponential Decay')
figure(11)

% Compute ratios from FFT
S(245:255)
format long
r1=[S(252)/S(252), S(253)/S(252),S(254)/S(252),S(255)/S(252), S(256)/S(252)]
format short

% Compute FT and ratios
f=0:.0001:.01;
beta=0.01*log(0.95)
FTD=beta+j*2*pi*f;
FT=abs(1./FTD);
FT(1:20)
figure(13)
plot(f,FT,'*r')
axis([-.001 .011 -100 2500])
figure(13)
format long
r2=[FT(2)/FT(2), FT(3)/FT(2), FT(4)/FT(2), FT(5)/FT(2) FT(6)/FT(2)]
format short
```

NOTE ON LISTENING TEST

It is very often the case that the interesting listening cases are those which are in part ambiguous. In particular here, I noted the differences between the PWM output of Fig. 2 and Fig. 3, as compared to the unmodulated pulse. There is absolutely no doubt that when we compare the unmodulated pulse to either modulated case that the apparent pitch of the modulated cases is much lower. Indeed, it is something like 8f (unmodulated) dropping to f (modulated). This we can understand in terms of the periodicity of the pattern going to a frequency of 1/8 the original.

Further, I have stated that both the modulated cases have the same pitch. I need to qualify this a bit. It is strange for a couple of reasons. For one, there is a different impression if you play the two rapidly back to back, as compared to waiting a couple of seconds between examples. If you play them in rapid succession, your first impression is that the second (Fig. 3) has a lower pitch. If however you play the first one, hesitate perhaps 4 seconds getting the pitch fixed in your mind, and then play the second, you will likely say the second pitch is exactly the same (the timbre is clearly different). The second strange thing is the by looking at the spectra, you would have guessed the pitch of the second might, if anything, have been apparently higher.

Here we are talking about the spectrum of modulated signals, and trying to guess the pitch that one "should" hear. Only special cases (setting to exact ratios) are even harmonic. Hearing a clear pitch is very difficult with balanced modulation, FM, and indeed here with PWM. Normally we have thought of PWM in terms of a dynamic <u>envelope controlled harmonic spectrum</u>. This case with <u>two audio range periodic signals is different</u>. We use other modulations this way all the time, but not usually PWM. Why have we ignored this – or is it just myself? We need to recognize that the result here is not untypical of modulation effects in general.

So some strange things are going on. At the same time, it is certainly possible that I have an error in my procedure. Thus I invite others to try these experiments in analog (or by simulation) and report any results.