The Pulse-Width Modulator with an emphasis on the post-modulator lowpass filter

James L. Tonne   W4ENE

Updated 28 July 2016

Summary

This note describes the generation of a pulse-width modulator (PWM) for an amplitude-modulated RF power amplifier for radio amateur service, with an emphasis on the output lowpass filter. The design procedure is applicable to both solid-state and tube-type equipment.

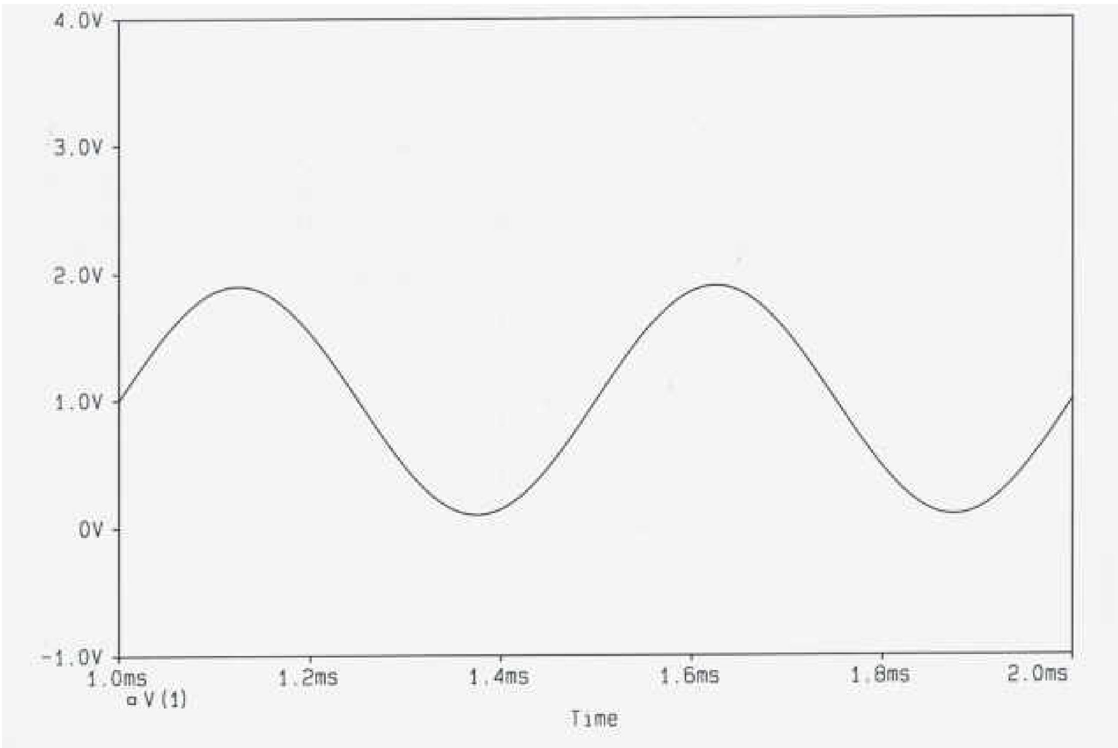
The switcher itself generates a complicated output spectrum which must be lowpass-filtered to remove the unwanted components. This filter must pass DC to an upper audio modulating frequency while rejecting other frequencies (the switching signal and its harmonics, with some sidebands for good measure). Other important items - often ignored - which will be taken into account here include the impedance seen by the switcher looking into the filter, transient response, envelope delay, and possible constraints on parts values.

This paper is based on observation of the waveforms encountered in the pulse-width modulation system and of their spectra. As will be seen, a series of tradeoffs are involved in the design of the filter. The design procedure includes control of the filter's transient response; this is especially important if any degree of clipping in the audio chain is used. Care taken in the design of the filter, along with the use of a nearly-direct-coupled modulator, can result in uncommonly-high modulation effectiveness for the AMer. The procedure outlined will not be particularly appropriate for the Kahn system wherein a phase-modulated signal is amplitude-modulated by the envelope component. This is because we are going to generally ignore the bandwidth and group delay requirements for such a lowpass and optimize the design for "ordinary" AM.

Generating the PWM waveform

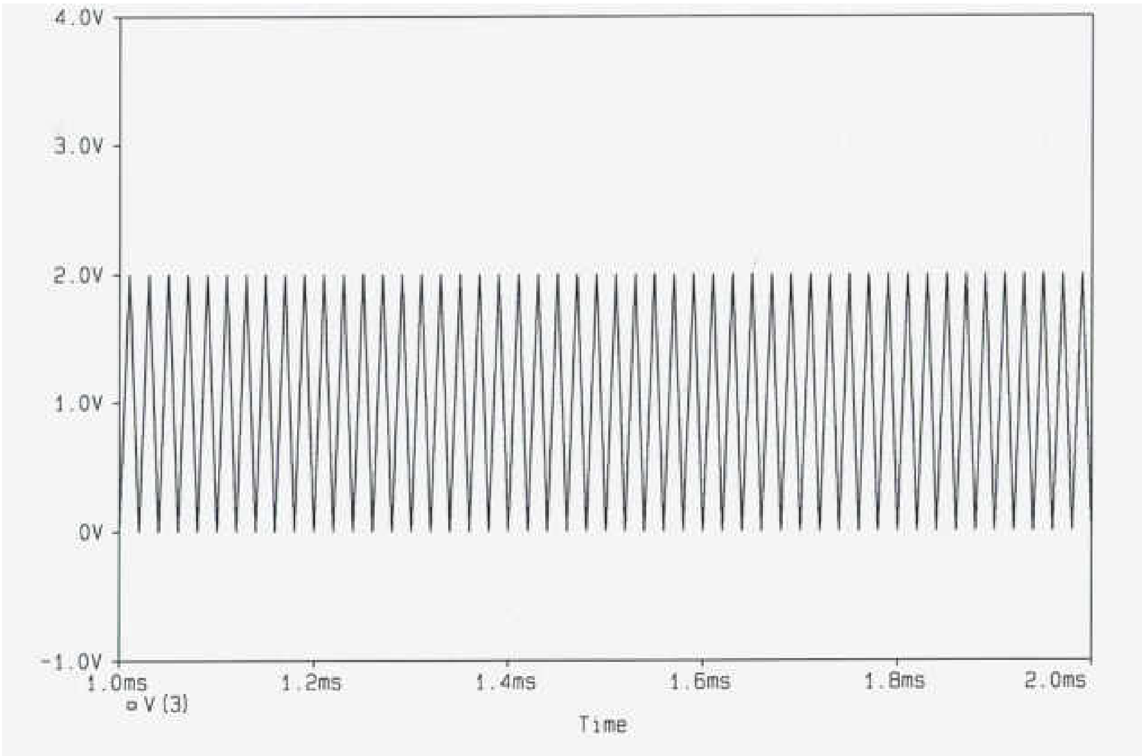
The PWM (pulse-width modulator, also known as pulse-duration modulator) waveform is generated in a straightforward manner. This is illustrated in the following graphics.

The modulating waveform is illustrated here:

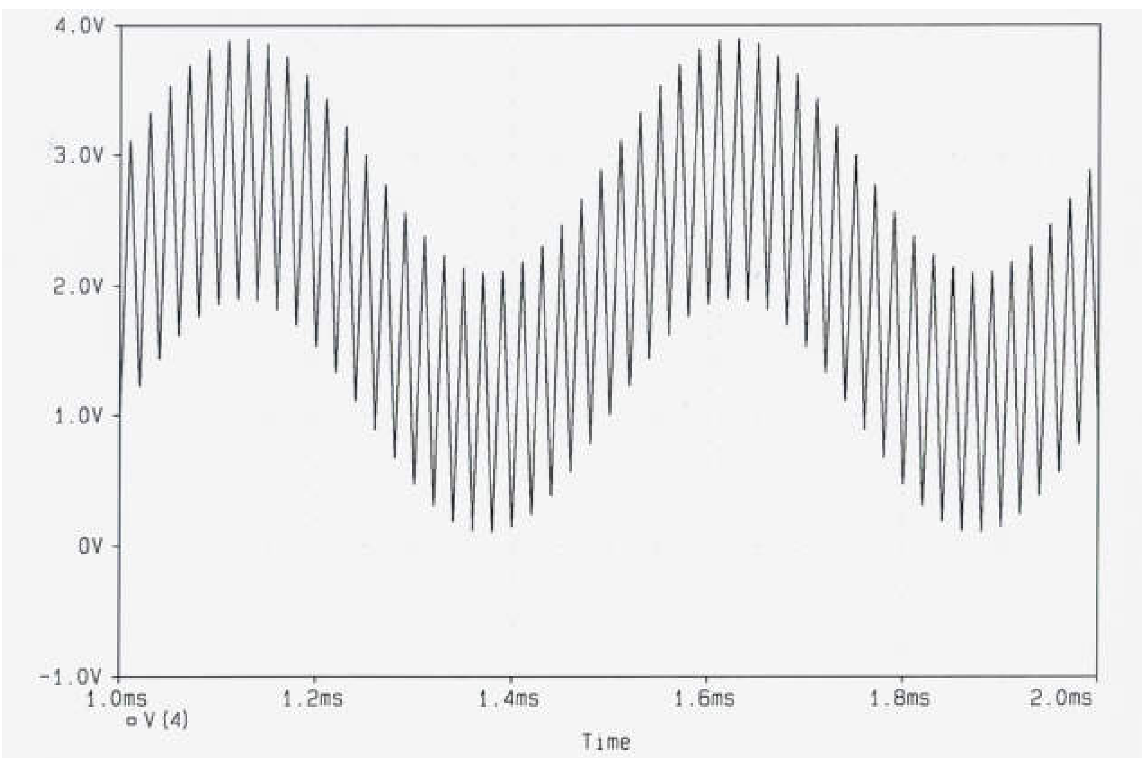


This waveform consists of two components. One is a DC signal corresponding to the carrier level and the other is the modulating sinusoid itself, set here to a frequency of 2000 Hz. The DC level has been set to 1.0 and the peak amplitude of the sinusoid has been set to 0.9, corresponding to 90% modulation.

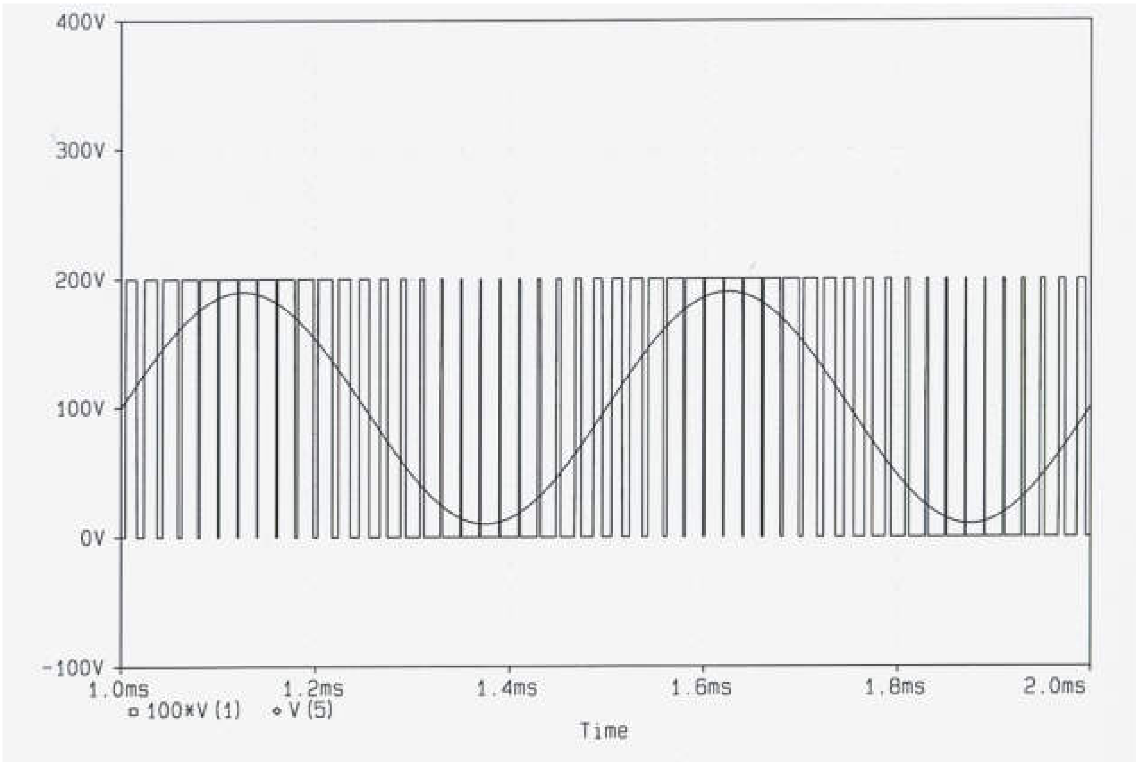
This signal is combined with a triangular-shaped waveform whose repetition rate is typically in the vicinity of 50 to 100 kHz. This waveform (at 50 kHz) is shown in the next graphic:



Those two signals (the modulating waveform and the high-frequency triangular waveform) are first summed (added together). This sum is shown in the next graphic:

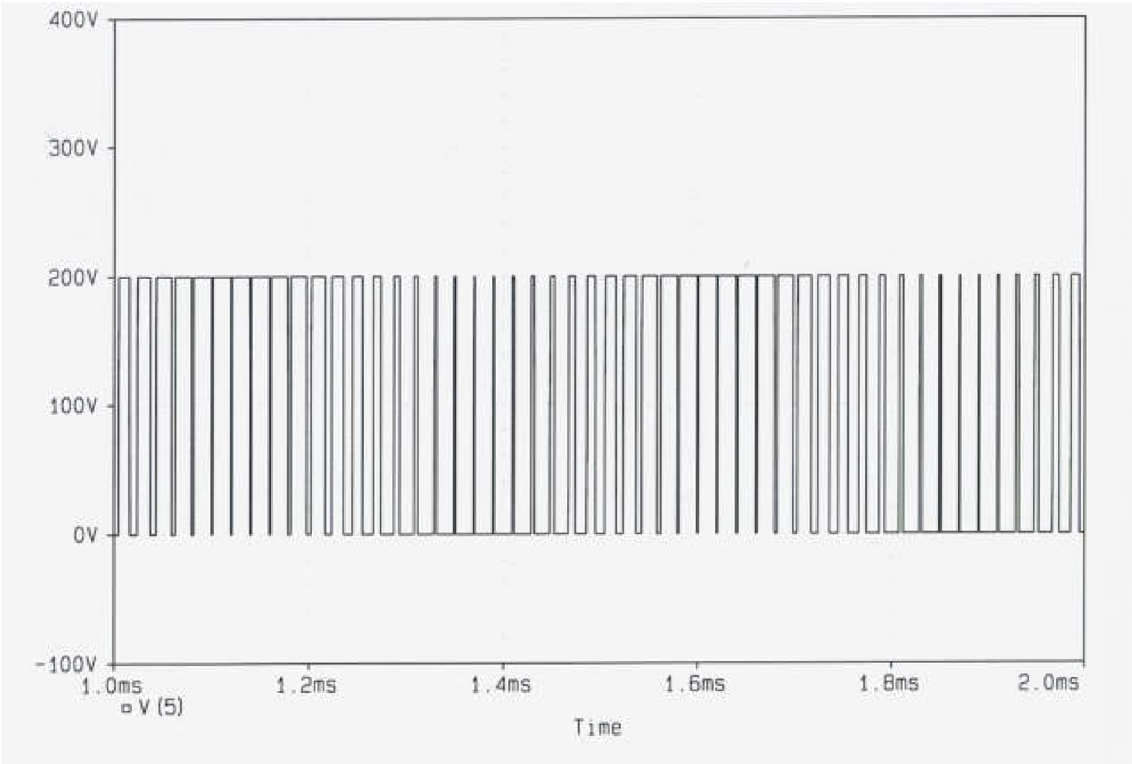


This signal is applied to a comparator which (in this case) generates an output when the sum exceeds 2 volts and generates no output when the sum is less than 2 volts. The output we are referring to here has been set to a value of 200 volts. This makes it easy to think in terms of "modulation percentage."



The modulating waveform has been superimposed on the comparator output to allow a simple comparison of the instantaneous value of the modulating waveform with the output of the comparator. As the modulating waveform value increases the duty cycle of the comparator output also increases.

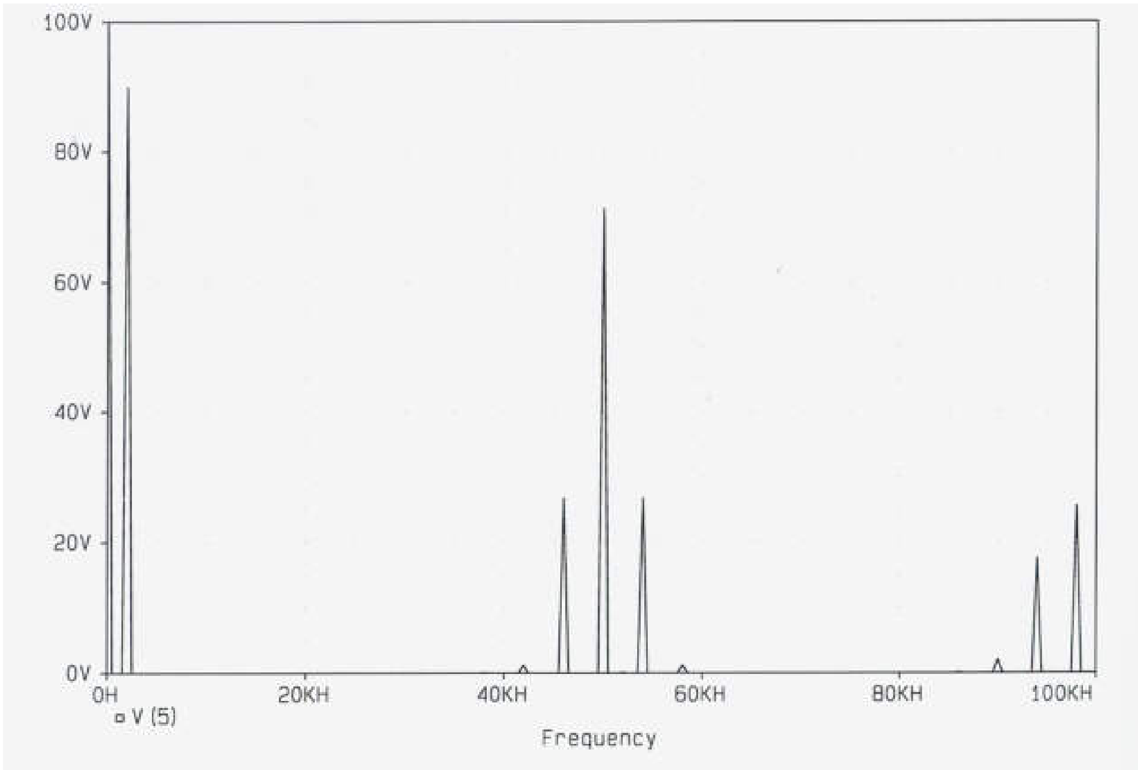
The comparator output alone then appears as in the following graphic:



This graphic shows the output of the comparator alone.

Examining the PWM waveform

If that signal is used directly to modulate the RF amplitude of an AM transmitter, it can be imagined that the transmitted signal bandwidth will be excessive. This can be confirmed by now looking at the spectrum of that waveform:



Here we see at 0 Hz 100 volts of DC corresponding to the idling carrier level. We can see 90 volts of a modulating signal at 2000 Hz.

Unfortunately we also see over 70 volts of a signal at 50 kHz, along with sidebands at plus and minus 4 kHz either side of that signal. Further, we see over 25 volts of a signal in the vicinity of 100 kHz. These harmonics of the switching signal extend spectrally without limit.

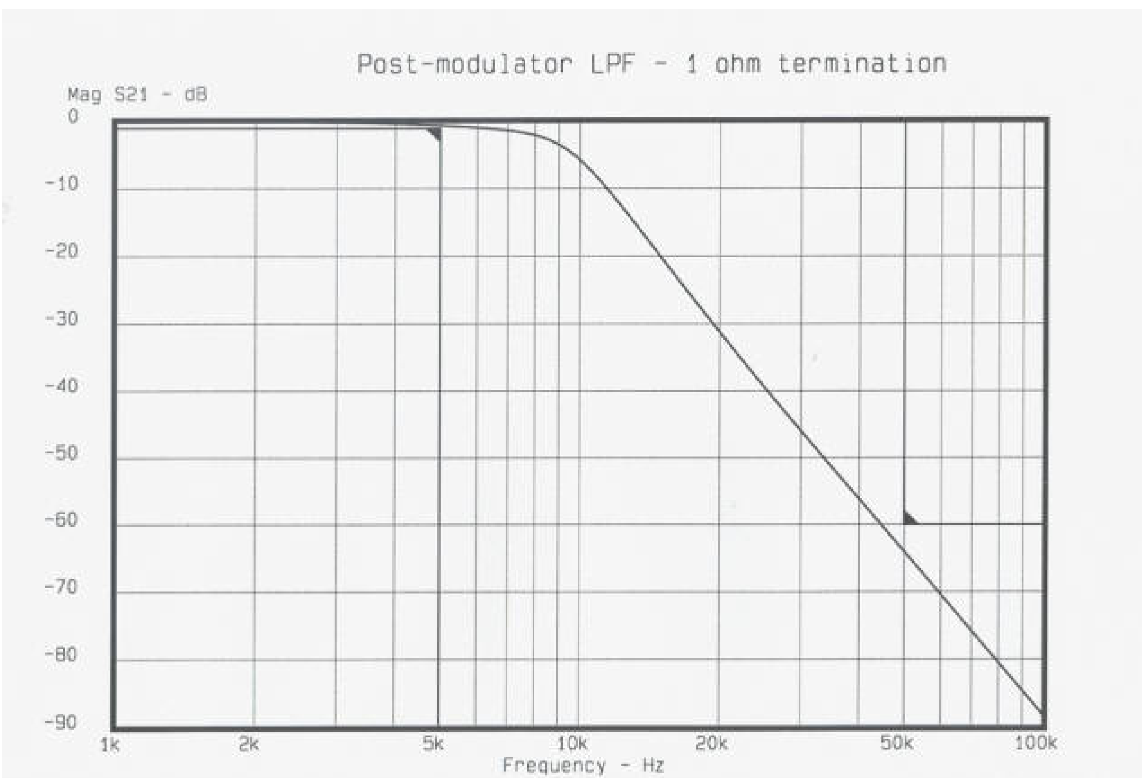
This is the spectrum of the signal from the comparator. If used to modulate the transmitter directly it would result in excessive occupied bandwidth.

The need for a lowpass filter

To restrict the occupied bandwidth of the transmitted signal we need to use a lowpass filter which will pass the DC and audio components while removing the switching signal and its related components (sidebands and harmonics). To be safe, these unwanted (switching signal) components as radiated in the form of sidebands about the transmitter's carrier should be suppressed 43 dB if the carrier power level is one watt. For higher transmitter power levels additional attenuation will be required. In an AM transmitting system the sidebands at full modulation are inherently down 6 dB from carrier and so the post-modulator lowpass filter can have its suppression of the switching-region components relaxed from the suppression target figure by 6 dB. This means that the *filter* response must be down 37 dB in the vicinity of the switching frequency for a transmitter with one watt output. Higher power levels would require greater attenuation to keep the spurious output level the same. For example, a transmitter with a 1000 watt output would need an attenuation in the filter of 67 dB at the switching frequency. These figures are slightly conservative but will keep the designer on the safe side.

This post-modulator lowpass filter should not be used as the primary transmission-bandwidth-limiting filter; that function is done more effectively by a sharper-cutoff lowpass filter employed in the lower-level audio processing stages.

The next graphic shows the magnitude response of a filter which would meet the requirements as just outlined. It passes DC up to about 5000 Hz with little attenuation while rejecting the switching frequency region by about 60 dB. Also shown on this plot are a pair of "limits" which define the allowable or at least recommended) responses.



This filter will pass the DC and audio components while satisfactorily rejecting the switching signal components.

Passband input impedance

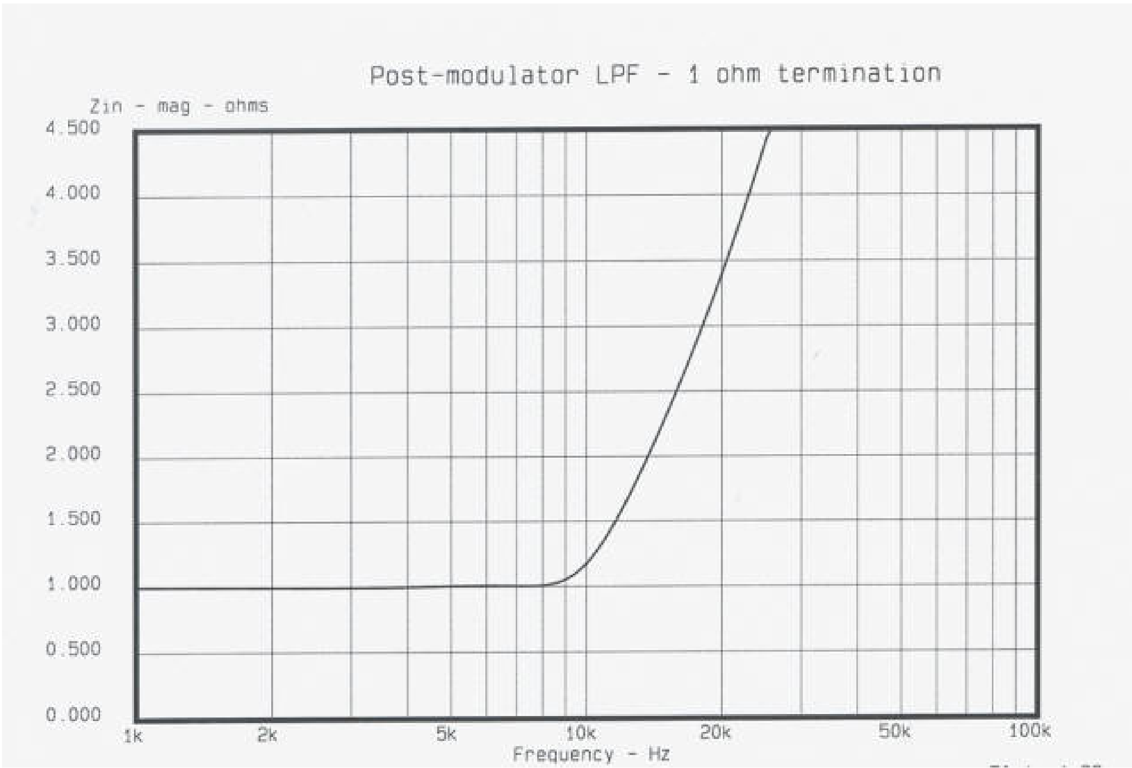
Seldom discussed in the design of the post-modulator lowpass filter is the input impedance as seen looking into the filter across the audio passband. The impedance seen by the switcher looking into the filter should be constant across the audio passband. The proposed filter has a remarkably constant input impedance up to 10 kHz.

If the input impedance of the filter falls to a low value at some modulating frequency, then the modulator will have to supply increased current at that frequency to maintain full modulation. If the impedance remains within reason then this is not a problem. If the modulator is current-limited (and it should be) then full modulation may not be achievable at that audio frequency because of insufficient output current capability on the part of the modulator.

If the input impedance rises to a high value at some modulating frequency, full modulation may not be achievable at that frequency because of insufficient voltage swing capability at the output of the filter.

Especially in the case of high power designs there exists a limit to the current capability of the modulator and a voltage-swing capability. It is folly and indeed expensive to design a modulator which has excess voltage swing capability and excess current-supplying capability. It is much better to design a system (this includes the post-modulator lowpass filter) which "just barely does a very good job." For full modulating *capability* at all audio frequencies, the impedance seen looking into the filter should be constant across the audio passband.

The input impedance of the recommended filter design is shown in the following graphic:



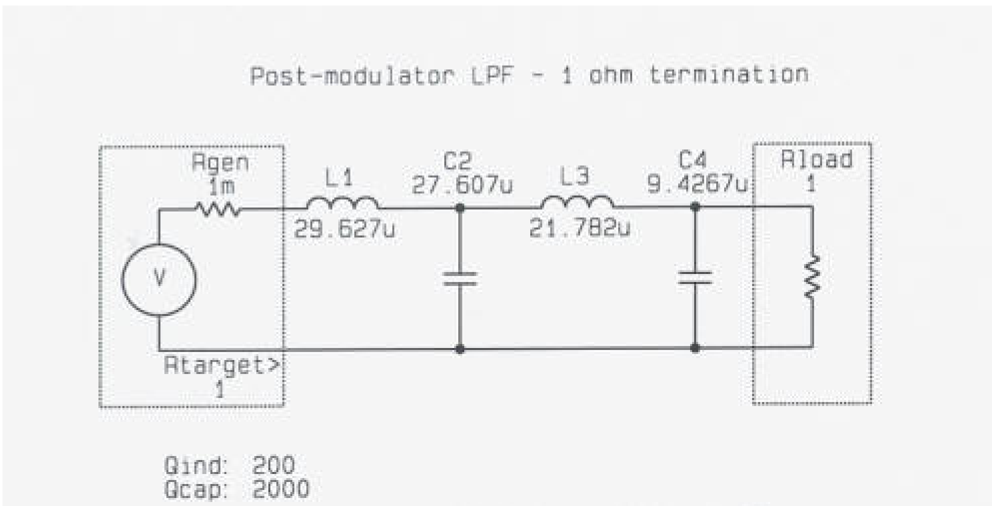
This subject of modulation capability is an item distinctly separate from that of modulation frequency response (or "magnitude" response). It is possible to have a flat magnitude response while at the same time having a limited modulation capability at one or more audio frequencies. This commonly manifests itself as an inability to modulate fully at, for example, the upper audio frequencies.

Stopband input impedance

The usual modulator (switcher) in this application acts as a voltage source whose output impedance is typically a fraction of an ohm for solid-state versions or perhaps 100 ohms for a vacuum tube version. In either case the source impedance (filter's input termination) is perhaps 1% of the load impedance (the filter's output termination. In any event, the impedance seen by the switcher looking into the filter should be constant across the audio passband. It will normally rise and become inductive in the stopband. This is harmless; in fact it unloads the modulator (switcher) at extremely high audio frequencies, where it may be inefficient. The load on the output of the filter is typically 5 to 25 ohms for a solid-state transmitter or perhaps 1000 to 5000 ohms for a vacuum-tube transmitter.

Because this lowpass filter is being driven by a voltage source, the first element must be a series inductor. For this reason the input impedance will become inductive at frequencies far above the audio range. The modulator becomes unloaded in this range, which includes the switching frequency region.

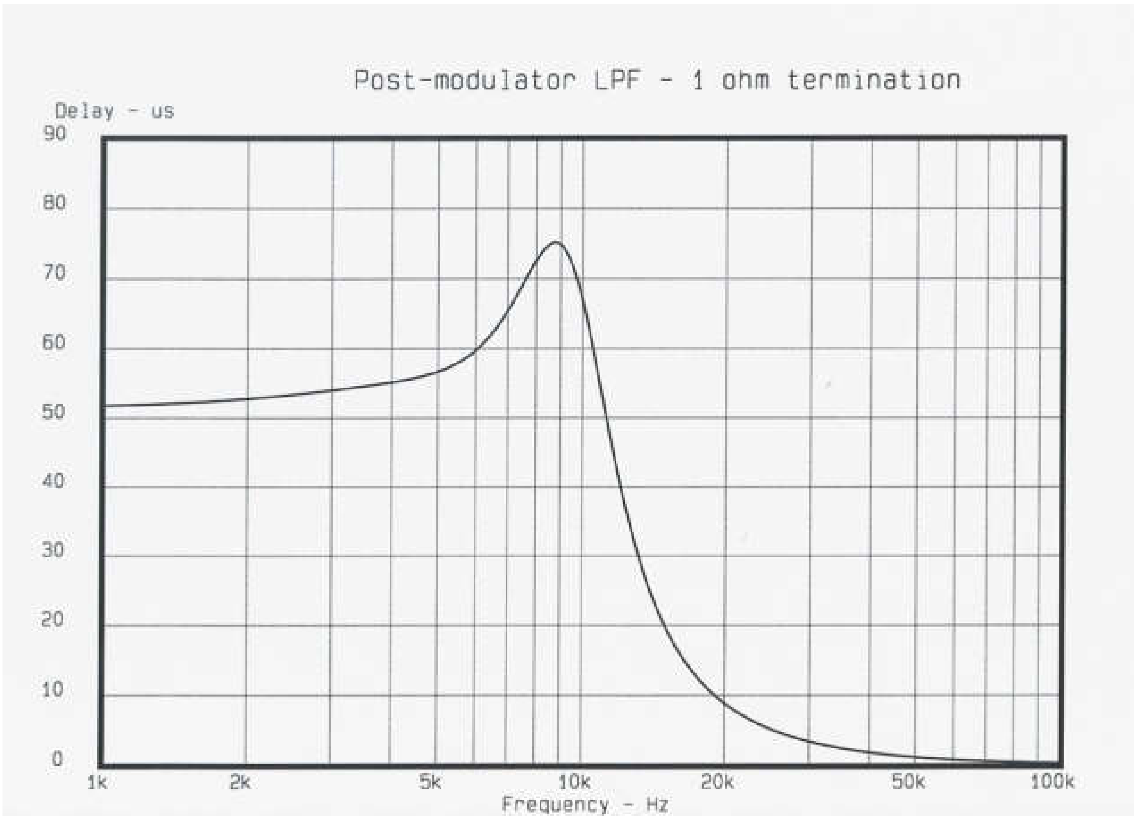
The schematic of the recommended filter for this application is shown in the next graphic. **The component values shown are normalized to a one ohm output termination.** They must be scaled properly for an actual application.



Envelope delay

For this paper we are going to largely ignore envelope (group) delay values *per se*. Delay values must be known and taken into account if the Envelope Elimination And Restoration system of SSB generation is used. This is because the phase-modulation path (at RF) must be delayed by the amount of time taken by the amplitude-modulation path in those systems. Further, it may be advantageous to have a time delay that is uniform across the audio spectrum. The need to know the envelope delay is an item by itself, above and beyond how it affects the system's transient response, which item is indeed of importance. For this paper we are going to place envelope delay in the background. We will, however, pay attention to transient response.

The envelope delay of the recommended filter is shown here:



Transient response

The transient response of the filter to the modulating waveform **is** of concern. The modulator and filter taken as a system should have little (preferably *no*) overshoot on an applied audio-frequency squarewave. This all comes about because it is very common to apply what is euphemistically called "processed" modulation to a transmitter. The object of this processing is to make the transmitted signal louder. The only way this can be done with a given transmitter power is by some kind of dynamic range reduction of the modulation. This usually includes both an automatic gain control system and a clipper. The clipper might operate only on transients escaping the AGC. Clipped audio is sometimes called "trapezoidal" modulation. The post-modulator lowpass filter we are designing will probably be required to handle some degree of modulating-waveform clipping. A more commonly available and standardized waveform to be used for testing is the squarewave; such a signal exercises the system to the maximum.

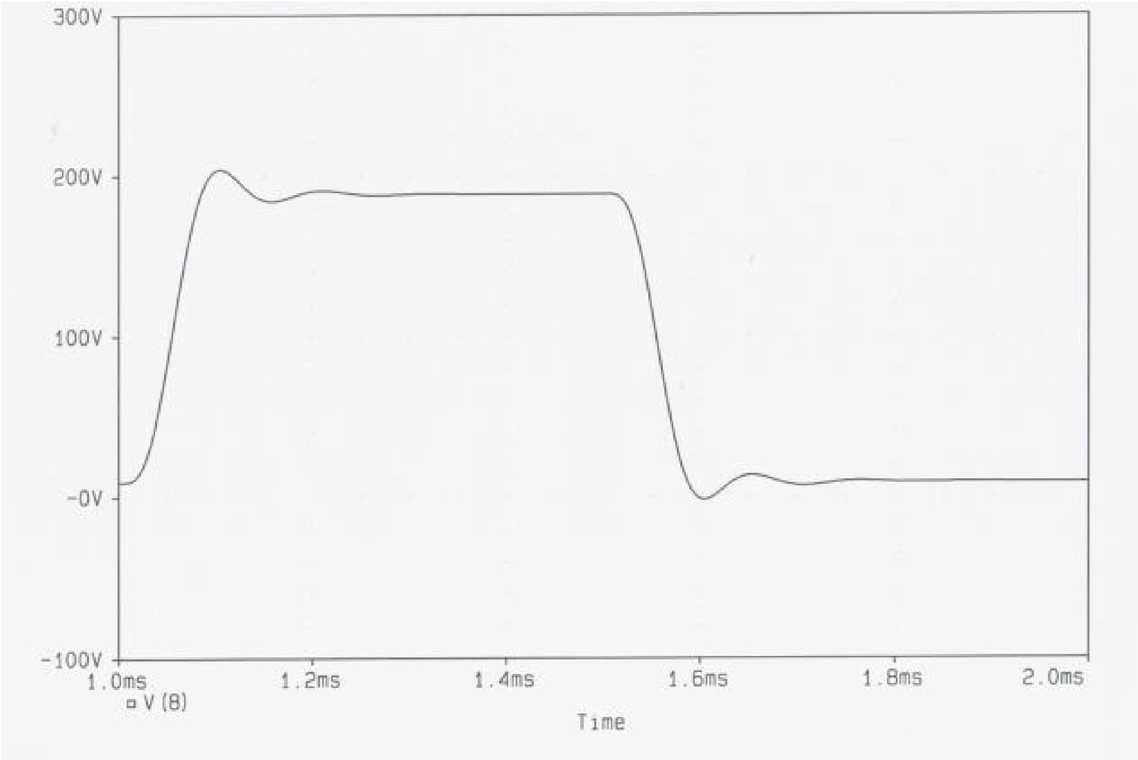
Due to the requirement of passing audio frequencies up to some point while offering significant rejection (perhaps 60 dB) of frequencies perhaps a decade above that point, the transition region between passband and stopband is relatively narrow. A direct result of this is a relatively sudden truncation of terms of a complicated waveform with a resulting overshoot added to such a wave. As the switcher frequency is moved upwards the filter can be simplified (the number of poles lowered). But, especially in a solid-state design, the efficiency may suffer. It appears possible to move the switching frequency up to about 120 kHz or even 150 kHz in vacuum-tube equipment without too much trouble, provided that the interconnection between the comparator and the switching tube can accommodate the required bandwidth.

The design of this filter involves a compromise among several variables, including passband bandwidth, switching frequency and the attenuation at that frequency, input impedance in the passband, passband magnitude flatness and overshoot on an applied audio-frequency squarewave.

If the bandedge of this lowpass filter rolls off gently, it will typically also have a relatively constant time delay; the two traits are normally interrelated. If a filter cuts off sharply, it will typically have a rise in the group delay characteristic at bandedge. The filter shown has a compromise between passband flatness and overshoot on a squarewave.

If heavily-processed (clipped, for example) program material were to be passed through a sharp-cutoff filter, overmodulation could result at high levels of modulation. At low levels of modulation the overshoot is harmless. A filter with a slightly rounded "nose" would have less of a tendency to allow overmodulation.

The response of the recommended design to a 1000 Hz squarewave is shown here:



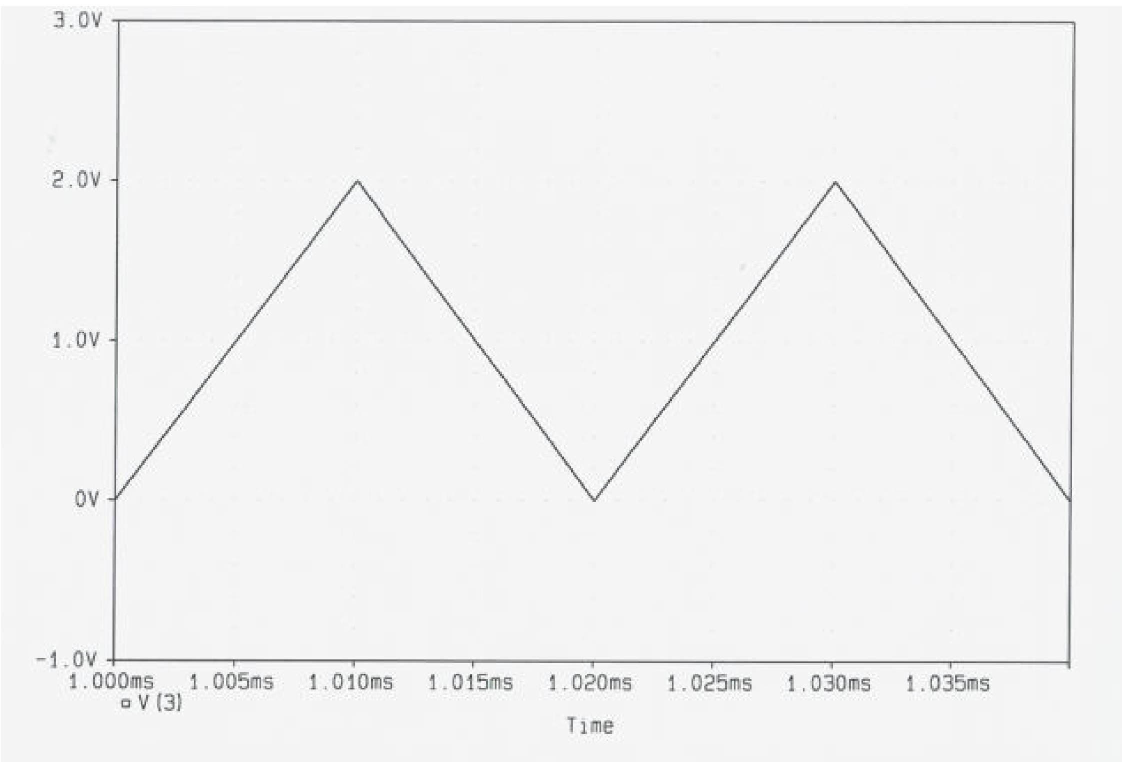
The post-modulator lowpass filter is the only item contributing to alteration of the modulating waveform. The overshoot seen here, which is quite modest and acceptable, is due entirely to the filter. Rephrased, the filter alone has an identical effect on the modulating waveform.

Since the filter's transient response is important, there may be a temptation to design for a gentle (Gaussian-like) rolloff. The problem with making this filter having a gentle rolloff is that the response in the audio passband will droop at the upper modulating frequencies if the required attenuation is maintained at the switching frequency and its sidebands. Increasing the order (complexity) of the lowpass filter doesn't help much in this regard. Further, the linear-phase filter families do not normally have a constant input impedance in the passband.

It should be mentioned at this point that it is possible (by misdesign of the filter) to have a rolloff followed by an actual peak in the response. This is undesirable for two reasons: one is that the overshoot resulting from application of a squarewave modulating signal can result in overmodulation (i.e., the transient response is substandard). The other is that noise in the low-level portions of the audio circuitry can be magnified by that peak and will be exaggerated. This noise peak may degrade the signal to noise ratio in the demodulated RF signal, and it may show up as a pair of peaks of noise on each side of the carrier as seen on a spectrum analyzer.

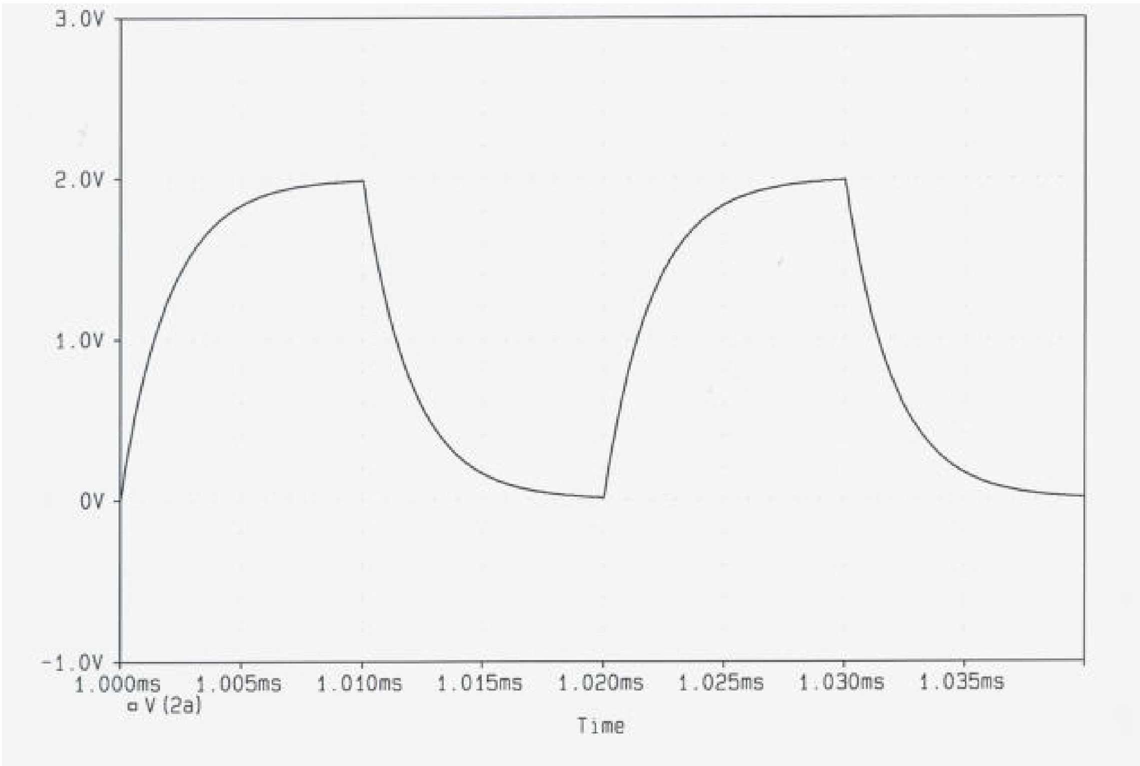
Triangle Waveform Linearity

Let us now look at the linearity of the trianglular- shaped comparison waveform. We have looked at that waveform previously; here we look at it again in detail, with the timebase of the plot being reset to show just two cycles of that signal.

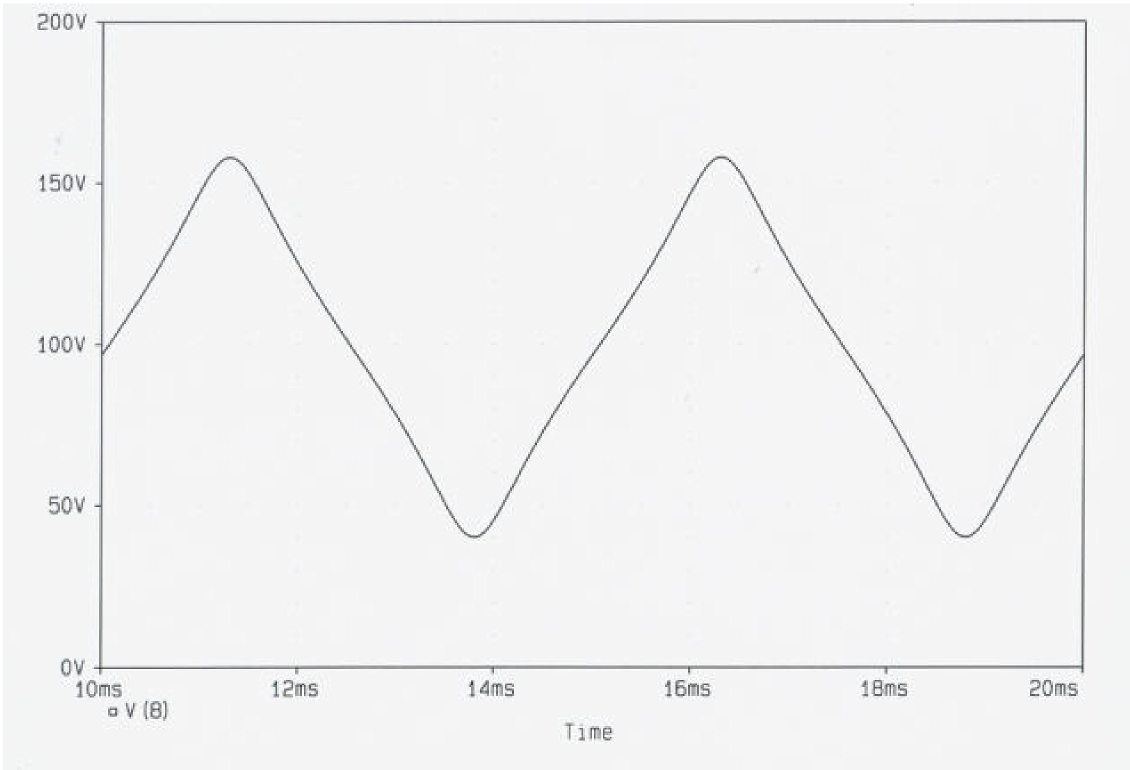


This should be considered a well-behaved triangular wave. It is in fact difficult to generate and for this reason an integrated circuit designed specifically to produce such a signal is commonly used. There may be a temptation to use a squarewave followed by a simple waveshaping network in an attempt to generate the triangular wave.

We will now see the result of a poor triangle wave by dramatically reducing the time-constant of the shaping network until we see the following signal, here deliberately over-done for ease of viewing:



If that poor triangular waveform is used as the comparison wave to generate our PWM signal we will see the following final signal from the output lowpass filter, a modulator output which is obviously quite substandard:

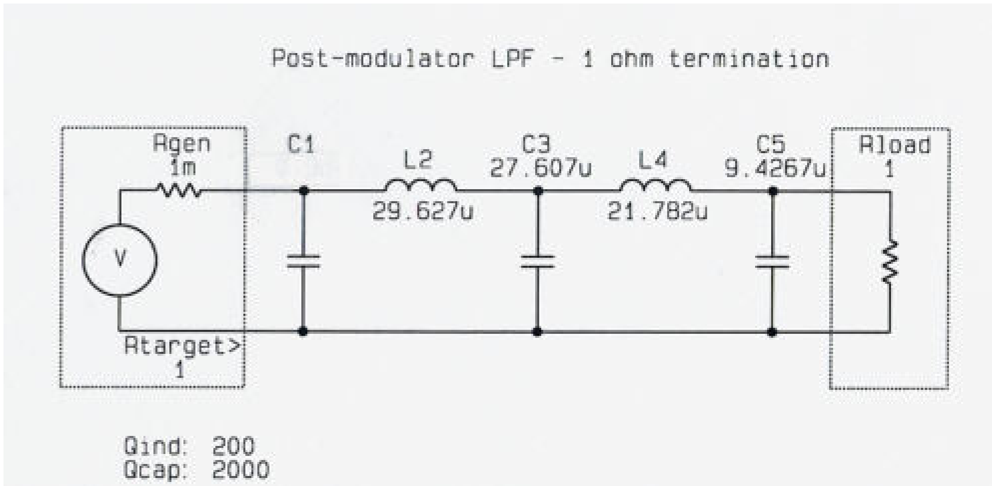


Analysis of this obviously defective sinusoid reveals a significant harmonic content. The waveform shows that its peaks are stretched slightly. A more palatable distortion would be to have the peaks compressed a small amount. Further, the overall amplitude has been reduced with this defective triangle comparison wave..

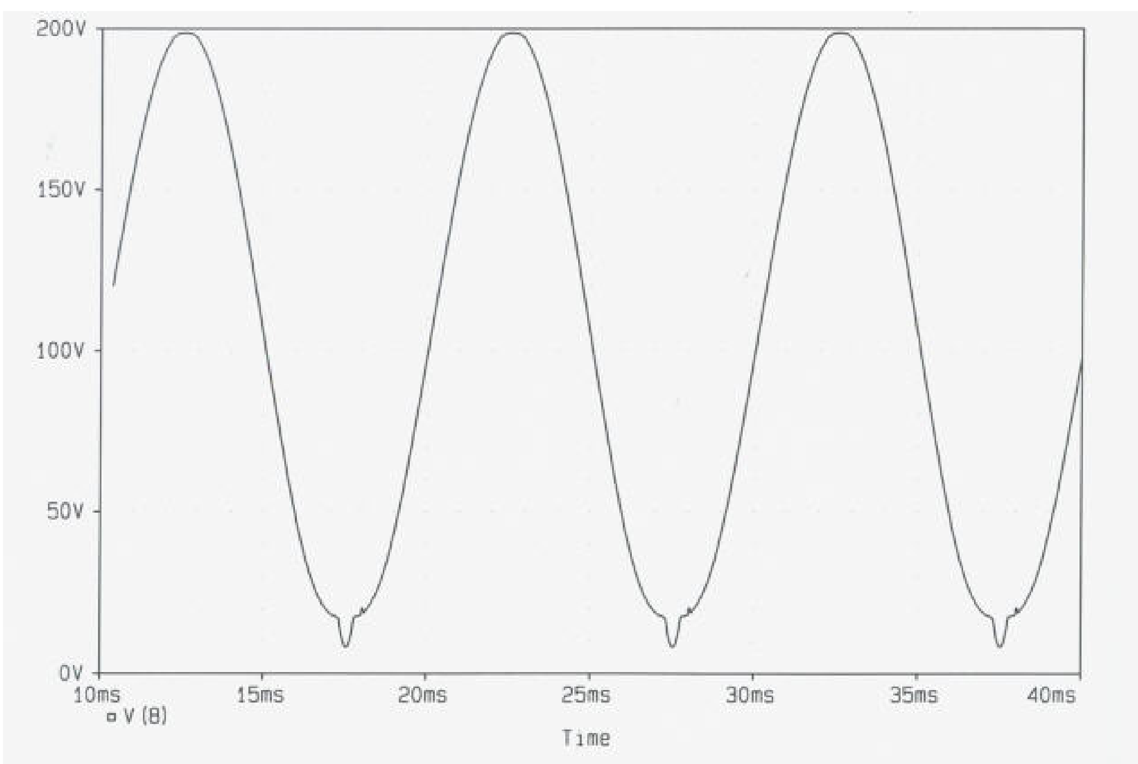
The linearity of the triangular comparison waveform has a significant impact on the modulator's output.

That Evil Stray Capacitance

The junction of the modulator output and the input of the post-modulator lowpass filter will have a stray capacitance to ground. This might be a byproduct of the modulator device (tube or transistor) or it might be due to stray capacitance across the first filter inductor or, more commonly, it might be due to a combination of the two. This capacitance will result in a load on the modulator, which must supply the current required to charge the capacitance. But a far worse effect is that such capacitance can seriously alter the modulation waveform on the negative peaks when full modulation in the negative or downward direction is approached. The capacitance we are talking about here is shown in the following graphic as C1:



The effect of this capacitance on a sinusoidal modulating waveform is shown in the following graphic. Note the point on the bottom of the waveform and also the fact that an appreciable part of the negative part of the sinusoid is shifted in the positive direction.



If the capacitance to ground from this node is appreciable, the switcher may charge that small capacitance in a single cycle but the capacitance can discharge only through the remainder of the filter into the output load. The result is serious distortion of the modulation envelope on the negative peaks ("troughs").

The waveform illustrated was obtained by adding to the analysis a value representing about one third of a percent of that first actual shunt capacitor. In a one-kilowatt transmitter using vacuum tubes, that first actual shunt capacitor in the filter might be a few thousand picofarads. The waveform degradation as illustrated would then be the result of an unwanted stray capacitance of ten or twenty picofarads. It is difficult to reduce the value of stray capacitance to such a low value. In the vacuum-tube transmitter the capacitance of the filament transformer secondary to ground would be greater than this unless a very special winding technique is used. The distributed capacitance in the input inductor in the lowpass filter must be added to that value for the total.

Several schemes to counter this effect have been used over the years. Patents covering this have all expired; some were of dubious utility. Unpatented design efforts involving trade secret are outside the scope of this paper.

Summary: keep the capacitance to ground at the input node of the lowpass filter to a minimum. This capacitance will include the output capacitance to ground of the switching device (tube or solid-state device), wiring, and stray or distributed capacitance across the first (series) inductor in the lowpass filter.

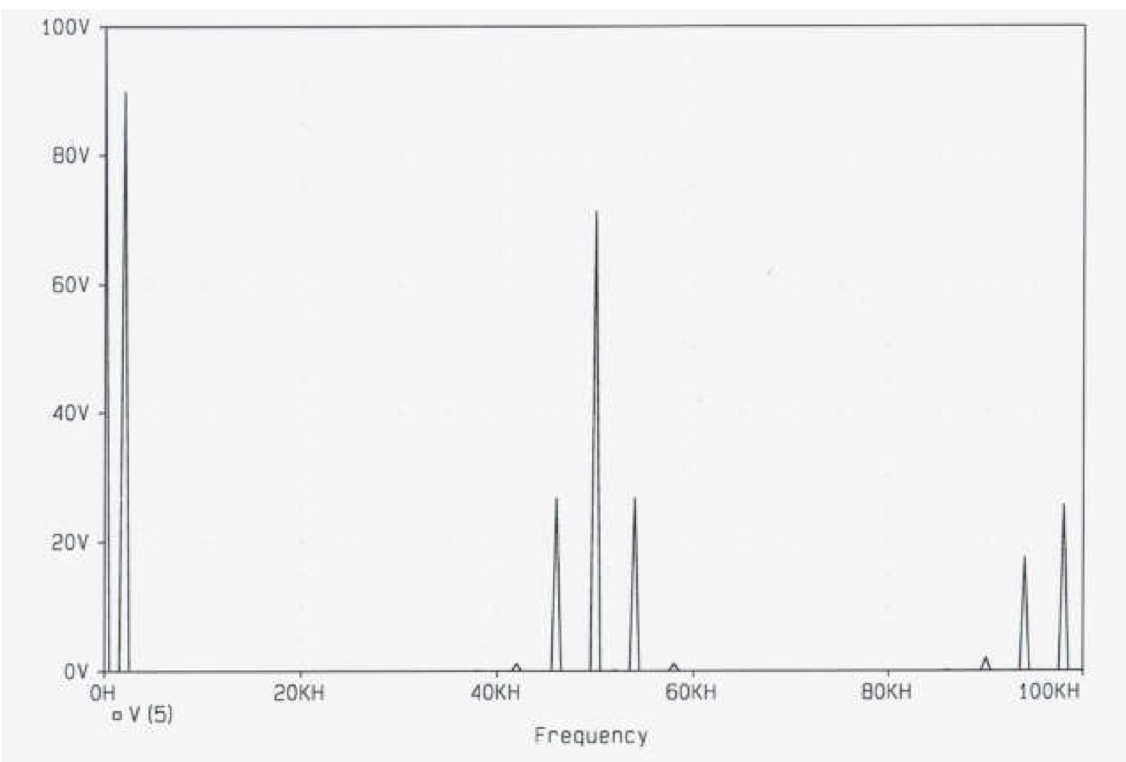
It may be of interest to consider using a very small inductor located physically right at the switching devices (transistors or tubes), then following that electrically (and so physically) small inductor with the remainder of the filter. It would seem that this approach might minimize the effect of stray capacitance to ground as seen at the input of the main filter. It may minimize the effect of, for example, stray capacitance in the filter's first inductor.

Ye scribe is amused by a tale which has been perpetuated for decades revolving around "collapse of the field" or some other magic. We have a potential problem revolving about stray capacity. The system as a whole including the filter components is linear. It does not suddenly display nonlinear behavior at a high modulation level. The switcher proper, with its diode-like output circuitry and with the commutating diode, will exhibit the effect shown when there is stray capacity added to the system. The inductors are not involved. The Spice file for the system, at the end of this paper, will allow proof of these statements.

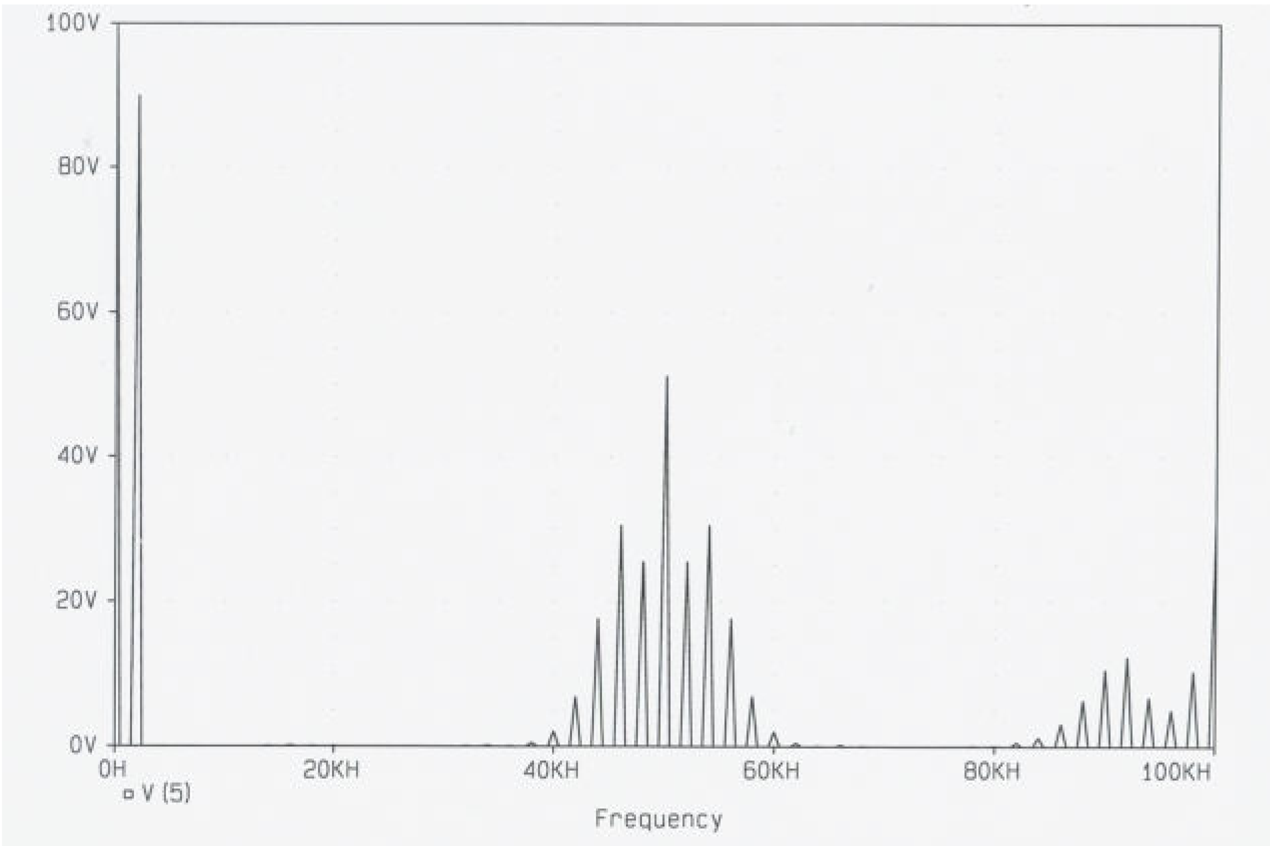
Spectrum note

The spectrum of the signal applied to the post-modulator lowpass filter in the PWM system is somewhat dependent on the shape of the comparison waveform. Sidebands appear about the switching signal and its harmonics. If the comparison signal in the switcher is of the *sawtooth* type then sidebands will be spaced about that frequency at all multiples of the modulating rate. If the comparison signal in the switcher is of the *triangular* type then sidebands will appear about the switching frequency only at even multiples of the modulating rate and their amplitudes fall off more rapidly with the higher orders than is the case with the sawtooth switching signal. In either case the comparator output has an essentially identical appearance. The *spectra* are noticeably different for the two.

Here we see (again) the switcher output spectrum in a system using a *triangular* comparison signal:



Here we see the switcher output spectrum in a system using a *sawtooth* comparison signal:



Either commparison waveform (triangle or sawtooth) should work well.

Design of the Post-modulator Lowpass Filter

The output impedance of the switcher itself may be considered as a low (essentially zero) value. This requires that the filter's design be that of a singly-terminated lowpass with the first element being a series inductor. (Even if the modulator could drive a shunt capacitor to ground, which it can't, that capacitor would be swamped by the source impedance of the modulator. See also the previous paragraphs about capacitance to ground at the input to the filter.) Set the topology to an inductor-input lowpass.

When such a design is attempted with a switcher operating in the 50 to 70 kHz region, it will be found rather promptly that at least four elements will be required for the filter. If the switcher is operating at 120 kHz or above it may be found that a two-pole design (a series inductor followed by a shunt capacitor to ground) may be satisfactory. For our high-efficiency design using currently-available transistors, set the order to four.

Start the design by specifying the bandwidth slightly greater than the desired audio passband. For amateur radio applications the filter bandwidth should be set to 5 kHz.

Select a Butterworth family, which has a smooth passband response. A computer-aided tuning process will be used to modify the design to trim the responses to have a compromise between passband flatness, modest envelope delay peaking at cutoff and a near-constant input impedance.

We can use Elsie (go to the software page on this site) as the final design tool. Elsie's tuning mode will enable us to home in on a good design very quickly. Use the tuning mode while monitoring the magnitude response and input impedance. The final design will have a trivial amount of droop in the magnitude response and an input impedance across the audio passband which is very constant. Envelope (group) delay can and doubtless will have a modest peak near cutoff.

The design shown is for a filter with an output termination of one ohm. Notice we have shown the input termination to be negligible - one milliohm - and the filter's output termination is one ohm. This is a "normalized" design, normalized to one ohm. To convert it to any other impedance just multiply the inductor values by the ratio of the new termination to one, and divide the capacitor values by the ratio of the new termination to one.

Design Constraints

It may be found during exploration of the design phase of the post-modulator lowpass filter that the design may be improved in one or more areas by using as a final element a shunt capacitor which is smaller than the design of the RF system dictates. If this is found to be the case - and it may be at the lower ham bands (for example the 160-meter band or at an LF or VLF frequency) then the RF system will be found to be controlling the overall system performance to some extent. The only way to circumvent this problem is to lower the audio passbandwidth of the lowpass filter. For LF and VLF applications (think Future!) this should not be a problem. The final value of shunt capacitance at the output of the filter should be that which the RF system requires. An optimizer, if used, must then use that value of capacitance as a minimum.

Normalized values

The filter as presented is a good design for stringent applications in radio amateur service. It may be used for other impedances (supply voltages, degrees of loading, etc.) by a very simple scaling process. **Again, please note this filter design is for a one-ohm system.** To go to another impedance, scale the inductor and capacitor values appropriately. For example, to go to a 2000 ohm load, divide all the capacitor values by 2000 and multiply all the inductor values by 2000. By simply scaling that filter to a new impedance value, all of the other traits of the filter (its magnitude and transient responses, constant input impedance, etc.) will then be maintained.