ADVANCED POWER TECHNOLOGY®

APT0001

By: Richard Frey, P.E.

APPLICATION NOTE

High Voltage, High Efficiency MOSFET RF Amplifiers – Design Procedure and Examples

High Voltage, High Efficiency MOSFET RF Amplifiers – Design Procedure and Examples

Introduction

With the improvement in high power MOSFETs of late – lower gate charge, low loss gate structures, and much improved frequency capability – it has become more possible to employ these "switchmode" devices in rf generators at medium hf. The objective is to improve the 'power density' (W/m^3) and efficiency of the equipment. The most common problem associated with this application is getting the matching right. For the gate drive side it is matter of obtaining an efficient and repeatable match to a very small impedance. On the drain side it is a matter of choosing a proper operating class and power goal for the device and then designing a network to provide repeatable and cost effective realization. The discussion below outlines the procedure for both.

The Input

The gate of a power MOSFET can be modeled as a series RLC circuit. The C is the specified C_{iss} from the data sheet at the desired operating supply voltage. The R is the effective series resistance, ESR, of C_{iss} and includes dielectric losses, gate metalization (or polysilicon conductor), and bond wire conduction losses. The L is essentially bond wire inductance, generally package specific. The resonant frequency of this RLC for large switchmode devices is usually in the middle hf range, 10-30 MHz, and the Q is fairly high.

As an example, the APT 5010LLC data sheet C_{iss} is 5200pF at Vds of 25V. In practice, a higher drain voltage is employed and this value will be somewhat smaller. On data sheets where the C values are graphed, it is common for the graph to stop at 50 or 100 volts. When graphed on a log-log scale, the C can be extrapolated linearly to the proper operating voltage. The effective gate C in an amplifier is increased through the action of the Miller capacitance. Generally the 50 or 25 volt value is a good place to start.

The gate ESR is seldom specified on the data sheet. Measuring this on a RLC bridge using 5V of static bias

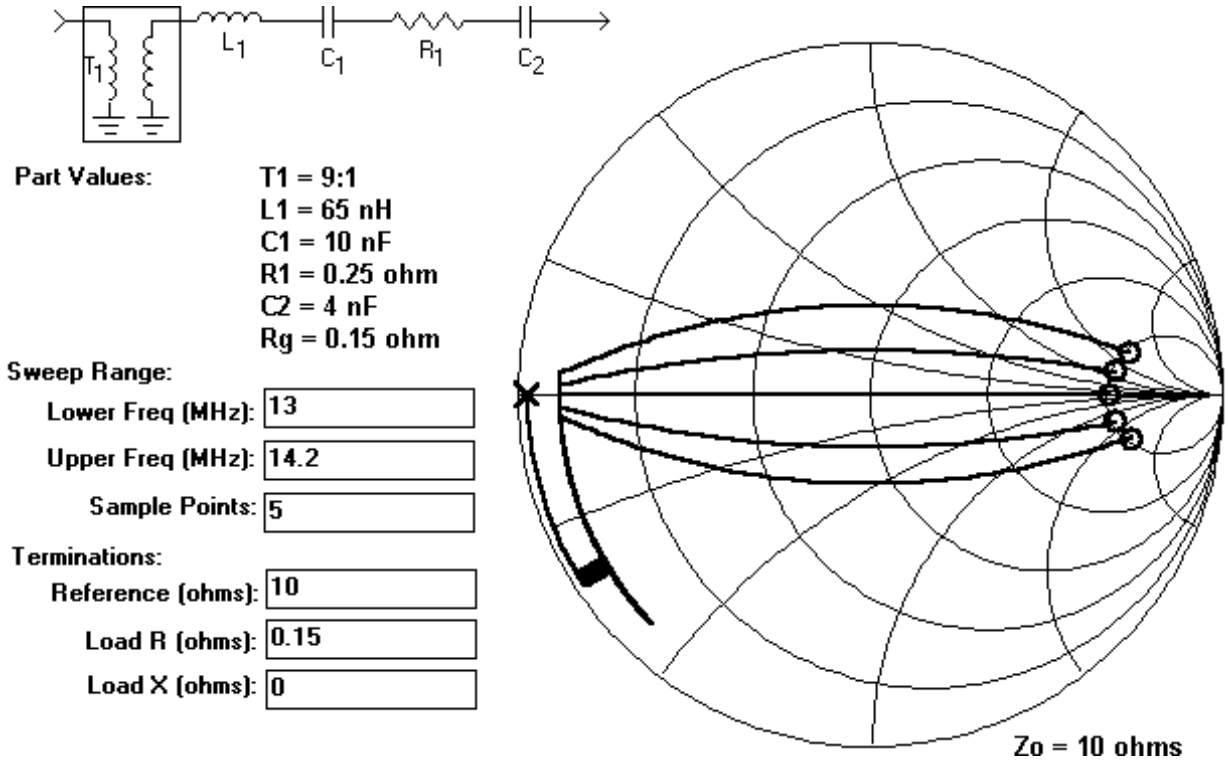

Figure 1 Input circuit design

with the drain-source shorted will yield a reasonable value. The 5010LLC has a gate ESR of approximately 0.25Ω. This value is highly dependent on the die geometry and fabrication process and varies widely between manufacturers of otherwise similarly specified parts. The higher this resistance is the lower the frequency limit for efficient operation and the more power will be dissipated in the gate structure. The gate ESR of APT devices is very low because of the all metal gate structure.

Gate parasitic inductance is generally between 5 and 15 nH depending on manufacturer, package and die size. 9 nH is a reasonable starting value for a TO-247 or TO-264 plastic package.

The circuit topology chosen for the drive is a ferrite loaded transformer with a single turn secondary. See figure 1. The graphics for this figure were produced by a Smith chart program called WinSmith[i], which was the tool used to design the networks described in this article. The "Load R" is the gate ESR. It is not represented in the series string in the figure but is the X where the trace starts on the chart. The transformer's leakage reactance plus the gate inductance is series resonated with the gate capacitance near the operating frequency. Adjustment of the resonant frequency is done using a series capacitor if the frequency is too low – it usually is if the device is large. The gate is referenced to ground or bias is supplied through a resistor shunting gate to source. See the discussion on biasing at the end of this article.

The procedure for designing the input circuit is fairly straightforward. The circuit is constructed using calculated and published values. Using an match indicator connected to the drive input connector, the circuit is energized with normal drain voltage, and the circuit is adjusted for best match at the operating frequency. The match indicator can be anything from a network analyzer to a noise bridge. The requirements are that it not inject enough signal to turn on the device and it must be sensitive enough to indicate a match to the driving impedance. The number of primary turns is adjusted to bring the input impedance real part close to 50Ω and then resonate the resulting total inductance by adjusting a series capacitor. In cases where a smaller device is used and the input capacitance is insufficient to resonate at the operating frequency, it may be necessary to add a small inductance. If bias is going to be used, this resonating should be done after the bypass capacitor is installed because it adds ESR and changes the total reactance.

The series capacitor is critical. The quality of the capacitor determines the total ESR. Any loss in the capacitor is drive lost and heating and resulting drift. However, if the capacitor is large enough the loss may be acceptable, or even desired if the gain of the amplifier is already too high. Further, lossy capacitors reduce the Q of the input and thus increase the bandwidth and lower the tuning sensitivity. The circulating current is high. Several capacitors in parallel will provide adequate power dissipation if high D caps are used and will offer a simple means of adjusting the value. The voltage rating can be 100V. For 13.5 MHz, the value used in the 5010LLC trial amplifier was 5nF composed of five 1nF Z5U 500V ceramic disc capacitors.

The transformer used is a two hole "binocular bead" balun made from high permeability ferrite material. The primary has ten turns of #28 insulated wire tightly wound on the center of the core. Enamel coated wire is suitable if care is taken to insulate the core. The secondary is a single turn of 0.25" wide copper braid. In this case, solder wick braid. This transformer may run at considerable power. In the test amplifier, output was 750W and the gain was 15dB, corresponding to an input drive power of almost 25W. It runs warm. The primary wire should have high temperature insulation, enamel or teflon. The core material should have low loss.

Ferrite bead balun cores of this type are available from several manufacturers. Most are made for EMI suppression and the core material is intended to be lossy! In this case, the transformer core was from Steward, part number 28N0765-100. Their type 28 material has an initial permeability, μ_i , of 850. Steward's type 25 material with μ_i =125 would have less loss and almost equal leakage reactance. The μ_i=850 material is more common and proved satisfactory for the demonstration. However, in a severe commercial environment, the μ_i =125 material with its lower loss and higher Curie temperature would be a better choice.

It is interesting to note that the driver is being matched to the losses in the gate circuit. All the input power is dissipated in these losses. What drives the MOSFET is the voltage that appears across the input capacitance, but no power is dissipated there. Also, because the gate is a very low rf impedance, the signal that one can measure at the gate terminal of the device is not the signal across the gate capacitor. The gate inductance, just 9 or 10 nH, is a significant part of the gate terminal

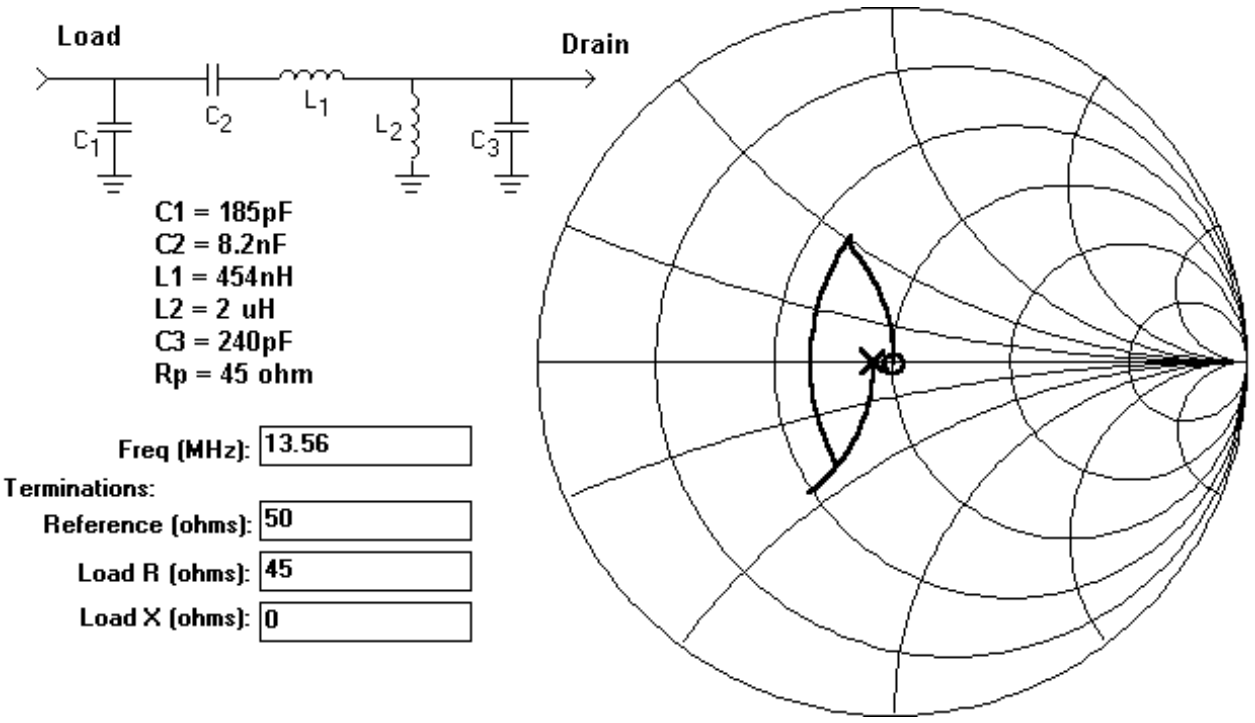

Figure 2 Output Network

input reactance. It can be calculated, but not measured directly.

The Output

The load impedance required by an amplifier is determined by the class of operation, the output power and the supply voltage. The classic formula provided in every solid state text book defines load resistance:

formula 1 $R_L = V_{dd}^2 / 2P_o$

The factor 2 in the denominator works for linear amplifiers running class AB where the drain swing is reasonably sinusoidal and the conduction angle greater than 180°. The factor diminishes with decreasing conduction angle as the class of operation passes through Class C to the high efficiency classes where zero voltage switching is intended, like Class E. To get the proper phase relationship between drain current and voltage for these higher efficiency modes of operation, the phase of the load impedance will typically be positive.

The drain model of the MOSFET is the output capacitance shunted by a "source resistance". This is necessary to perform the matching calculations but has no physical realization. If there were such a source resistance within the MOSFET drain, the efficiency could never exceed 50%! The value of C_{oss} can be read from the data sheet. Its value should be that for half the drain supply voltage to account for the nonlinear value change over the total drain swing. The Q of the output capacitance is a major contributor to efficiency. "Super Junction" high voltage MOSFETs are notoriously poor in this regard and should be avoided for rf amplifier service.

The class of operation is defined by the drain conduction angle; the portion of the operating frequency period when the drain is conducting. Class A operation is fully conducting, 360°. It never turns off. It requires a positive dc bias voltage applied to the gate. It is linear but not efficient. Class B is a special case where the gate bias is adjusted for drain current cutoff. The conduction angle is 180°. It is linear and more efficient than class A but not more than 75%. Class C is less well defined but is characterized by conduction angles of less than 180°. There is usually no bias voltage provided except by the drive signal, and the efficiency is as much as 90%.

There are more classes. Saturated class C, sometimes called C-E is a higher drive version of the same thing. It can be 95% efficient. Class D is a particular configuration using two transistors as switches. Its use

is limited by the parasitics of passive and active parts to about 10MHz. It can be nearly 100% efficient. Class E is a special zero voltage switching mode. It relies on the proper phase relationship between drain voltage and current to provide no overlap. The efficiency can be 100% with perfect parts. There are other classes, somewhere between C and E, see US patent 5817580 for example.

For the purpose of this paper, the class will be sinewave driven Class E[ii]. This does not require a large amount of drive. At 13.56MHz, the small signal gain of these MOSFETs is well more than 25dB. For saturated class C, enough drive is used where it can be considered almost a square wave equivalent. APT MOSFETs have a gate breakdown voltage well past 40V. The proper drain load has been determined empirically. In this case, the load was measured after the amplifier had been tuned for best gain and efficiency. The factor in the load impedance formula was adjusted to fit the result. The inductive component was derived from the measured results and has proven effective in subsequent amplifiers. In any case, the calculations will yield the values needed for a functioning amplifier, close enough to permit minor tweaking for best operation.

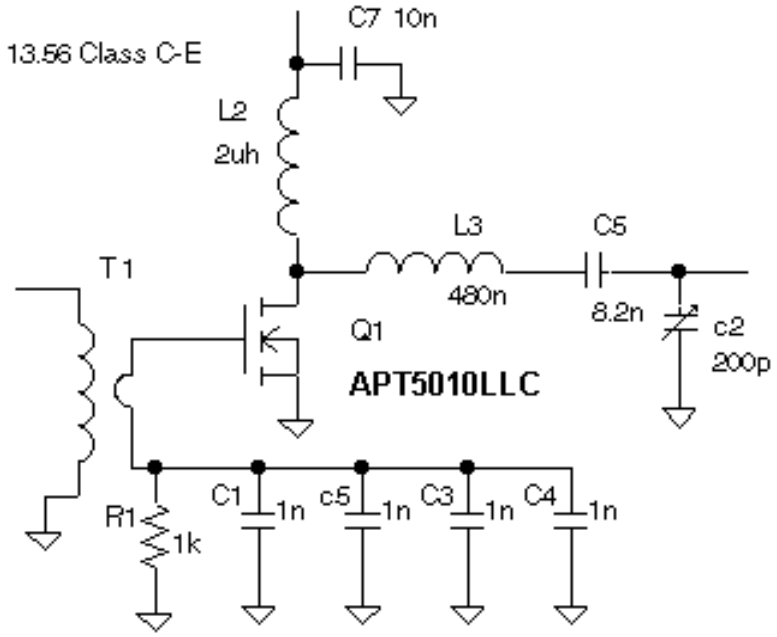

Figure 3 Schematic as Built – 1kW @ 13.56 η=89%

The output circuit is shown in figure 2. This is deceptively simple and, with minor changes in part values, drive level and bias, can be tuned to run any class – A through Z. The calculation for load impedance is per formula 1 except we will use a value of 1.2 for the denominator factor. This compensates for the large harmonic content in the drain waveform and is partly the result of experience. The phase is lagged somewhat to 'unlap' the voltage and current waveforms and improve efficiency.

The APT5010LLC has a rated power dissipation of 520W (only when the junction temperature is at its maximum of 150°C and the case is at 25°C). A practical useable value is half that. This takes into account the thermal resistance of a good drain insulator and a lower maximum T_j. In any good design, power dissipation is determined by total thermal resistance. Assuming a 70% efficiency for the amplifier, the maximum output power can be 600W. If the actual efficiency is better than this, the design margin improves. 600W corresponds to a drain load of 37.5Ω for a drain supply of 150V.

The output C_{oss} is 1030pF at 25V – not a particularly helpful figure. With a 150V supply, the C_{oss} is probably closer to 700pF. This capacitor in parallel with the 37.5Ω "source" impedance will be used to determine the output network. From this point the procedure is very mechanical. The parallel combination is transformed to an equivalent series impedance, 6.25 -j14. An L network is used to transform this impedance to 50Ω. How one chooses to design the L network is a matter of preference. Handbook formulas will work as will a Smith chart. The real part is transformed to 50Ω. The required series L is increased by +j14 to cancel the series output C.

In a practical circuit there are some slight adjustments to this procedure. A method to feed the drain voltage is

needed. A shunt choke is the typical method. If the value is not too large, it will cancel some of the C_{oss}. A 1 μH value was chosen which reduces the effective C_{oss} to 562pF. The calculation for the series equivalent impedance and L network values is repeated. Some form of blocking capacitor is needed to keep the supply voltage out of the load. If placed right at the output connector, the tuning capacitor will have both dc and rf on it. It is thus better to put the capacitor in series with the inductor unless the drain load is very low - as it would be in a 50V rf device.

Choice of parts is a major factor determining the overall success of the circuit. The passive parts are as critical as the active device. As the stage power increases, the stress on individual parts can be severe. A case in point is the blocking capacitor. Placed in series with the L network coil, the rf current through it at 700W is considerable, 4.3A. Unless this is a low loss capacitor, it will heat, drift, and even fail. Surface mount parts are nice but not for this application. There are "leaded surface mount" parts, stacked MLC types made for SMPS service. The one used here is an AVX part number SM067A822KAJ120. They are available in 1kV COG characteristic from several sources in values up to 10nF. A high Q coil is necessary. Placing this coil near a ground plane or compartment wall will lower the Q and increase the losses. It can get hot enough to unsolder itself!

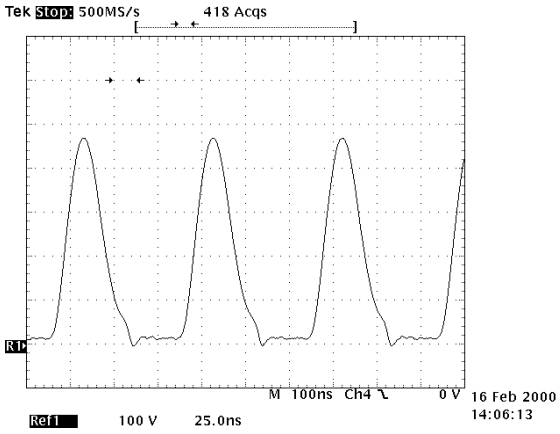

Figure 4 5010LLC Drain voltage waveform

Tuning the output circuit is a matter of adjusting the series coil L_2 and shunt capacitor C_4 for best output power and efficiency. Initial tuning is best done at a low duty cycle, 20 to 50 %. For a trial value of L_2 , adjust the value of C_4 for best performance. Adjust L_2 slightly and repeat the sweep of C_4. Make sure the most efficient power does not result in a peak voltage swing too close to the BV_{dss} rating for the device. In it not unusual for the most efficient power to be well in excess of the voltage capability when the conduction angle and drain current-voltage overlap are minimized. If the peak voltage is too much, add some external C_{oss}, reduce L2 and tune for a lower power, or reduce the supply voltage. The peak voltage can be three to four times the supply voltage.

The circuit of Figure 3 was also used with an APT10050LVR, a 0.5 Ω 1kV part similar to the 5010LVR. With only slight adjustment of L3 and C2, similar results were obtained.

Note in both Figure 4 and Figure 6 that the peak drain voltage is more than the rated breakdown voltage of the part used. Clearly, there is some avalanche occurring but the total avalanche energy is very small as indicated by the good efficiency. This is not a desirable situation and can be remedied in several ways. Reducing the supply voltage is the most obvious but is not always an option. The next best way is to reduce the impedance of the drain circuit by adding additional shunt capacitance across the drain and retuning the output network. In so doing, the peak drain voltage is reduced without materially effecting the output power because the impedance is also lowered.

There are programs available which permit designing analyzing, and optimizing class E performance before doing anything in the lab. This can be very useful for evaluating and comparing different devices, and for adjusting a circuit for a particular maximum peak drain voltage.[iii]

It will be necessary to go back and readjust the input match because Miller capacitance, Crss, does indeed change the input impedance. Watch for signs of instability – the tendency for oscillation. If the layout is poor, or the gain is too high, it will oscillate. There is another phenomenon called snap-on which occurs when the change in input impedance between full drive and no drive interacts with the driver. When the input has been adjusted for best SWR at full output the driver provides maximum power to the gate. When drive is reduced, the output goes to zero. When the drive is increased from zero, the device input impedance is different and the driver will not be able to provide enough drive at the normal setting. As it is increased further, still with a fairly poor match, the input will eventually have enough drive to make output. As the C_{rss} feedback changes the input match, it improves, the power transfer into the gate improves and suddenly we have full output.

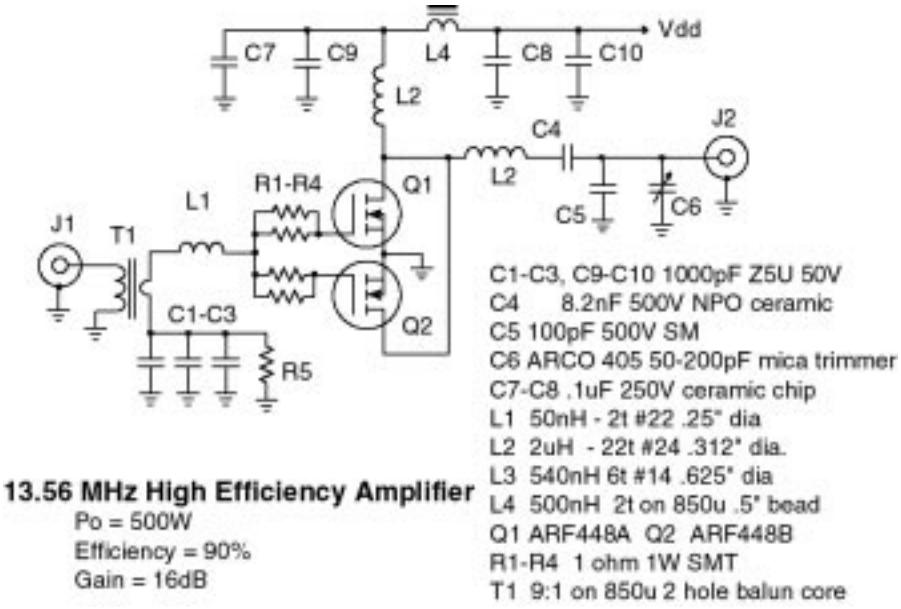

Figure 6

This can be seen as drive hysteresis. A stiffer driver, more series resistance in the gate, or additional fixed shunt capacitance across the gate will improve this situation. In the test amplifier of Figure 5, 1nF was added across each gate to ground. The input match was readjusted and the hysteresis was almost completely eliminated.

If several devices are being paralleled, it is necessary to insure they will share the load. It a common misconception that paralleled MOSFETs inherently share the load. No, they do not! Even though the temperature coefficient of $R_{DS(ON)}$ is positive, this only effects the fully saturated condition. The temperature coefficient of the gate threshold voltage is negative. With constant voltage bias, the drain current goes up as the device heats up. The hotter device is turned on harder.

In an ideal device that is conducting current and dissipating power, it can be assumed that current distribution across the die is uniform and that the distributed junction temperature is also uniform. In reality, however, if there is any trace of non uniformity in the current distribution, the junction temperature will be higher in those regions where the current is highest. Non uniformity can be precipitated by small voids in the die attachment solder, non uniform junction doping, and by a host of other variables. Hot spotting is then triggered by the negative temperature coefficient of gate threshold voltage acting in a localized manner. As one part of the die heats up more than another, the local junction Vgs(th) falls and Gm then causes that area to turn on harder drawing even more current. Since Gm increases with increasing current, the runaway process is accelerated even more. The fact that $R_{DS(ON)}$ has a positive temperature coefficient is irrelevant in that the MOSFET is operating in the linear region, and not in "saturation".

The bigger the MOSFET die, the higher the power dissipation and the greater probability of experiencing this problem. The alternative is to use multiple smaller devices and buffer them with source degeneration resistors.

Gate Bias

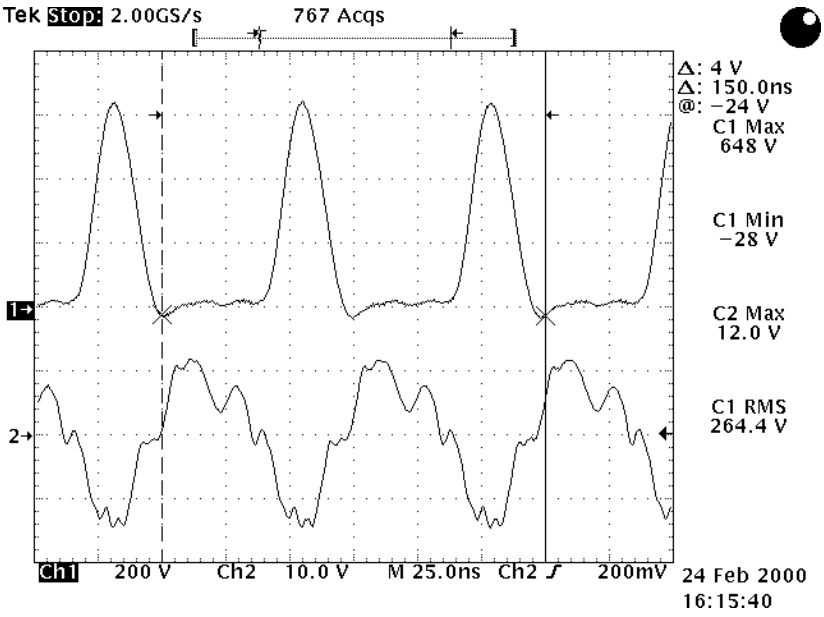

Figure 5 2x ARF448 waveforms. P_o = 600W, η = 90%.

In an N-channel MOSFET, zero volts between the gate and source results in no drain current; the device is cut off. These devices can run class C without any additional voltage on the gate. To run class B, a positive voltage is applied to the gate to a point where the device just starts to turn on. Class AB bias would be set for some amount of static no signal drain current. The gate threshold voltage of all power MOSFETs has a negative temperature coefficient. It is usually shown in a figure on the data sheet as Id vs Vth at several temperatures. It means that a fixed amount of class B bias applied to a cold device can be enough to put it well into class A after the application of rf drive due to the change in T_j. This will kill the part by over dissipation if the drain supply voltage is more than 75V unless a system is in place to actively adjust bias for changes in T_j.

High voltage MOSFETs do not work well with forward bias but bias adjustment can be useful. It has been mentioned in the literature for years that one of the benefits of the MOSFET is the ability to control the gain of the device by adjusting the gate bias. The typical gate threshold voltage is 4V. That means an 8V p-p signal is needed on the gate to start turning it on. If 1V of fixed dc bias were put on the gate, the effective threshold would be reduced by that amount and a 6V p-p signal would have the same effect on the drain. The gain is higher. The reverse is true. A negative bias will increase the threshold and effectively reduce the gain. This has great implications for gain/power control.

Systems

The output power required for a typical generator system is always more than can be obtained from a single device. In order to minimize the amount and cost of the "glue" required to drive and combine the outputs of several amplifiers needed to produce the required power, the opportunity exists to use custom devices containing multiple die. The motivation is improved reliability, efficiency, power density, and lower cost. Very large die are subject to hot spotting. Several smaller die with suitable buffering can be incorporated within a thermally efficient package to improve the performance and increase ruggedness and reliability. The cost of the package is more than offset by the savings in combining parts and circuit simplification.

The more efficient the amplifier, the more dependent it becomes on having a particular load for proper operation. Class A amplifiers are most often used in wideband general use lab amplifiers which are required to guarantee stability under all load conditions. Class C-E or saturated C and class E amplifiers have less tolerance to load disturbance. In order to provide a working system in 13.56 MHz plasma generators, a matching device is needed between the load and the amplifier. It will 'follow' load changes and provide a more constant load to the amplifier.

Harmonics are not a real problem for the plasma generator. While some specifications put limits on harmonic energy delivered to the load, the effectiveness of the plasma itself is not effected by harmonic energy. The most significant problem is the effect of the harmonic energy on the VSWR detectors used to control the matching network. It is sometimes useful to place a lowpass filter on the rf samples before detection to insure a good null at the operating frequency. If an output low pass filter is used, its effect on the stability of the amplifier must be considered. Harmonic energy reflected from the lowpass filter circulating in the output network can cause serious component heating, loss of efficiency through modification of the drain waveform, and instability when the match is disturbed. Use of a highpass/lowpass diplexer[iv] with the highpass section terminated in a load can ease the problems associated with harmonic energy.

Conclusion

If proper drain load is chosen, the output power and efficiency goals for the amplifier stage can be reached. Proper functioning of the circuit in a commercial environment is very much dependent on the quality and selection of the passive parts in the circuit as well as the cooling and load protection systems necessary to support it.

---

[i] winSMITH 2.0, 1998, available through Noble Publishing Corp.,4772 Stone Drive, Tucker, GA 30084.

[ii] J. F. Davis, D. B. Rutledge, "Industrial Class-E power amplifiers with low-cost power MOSFETs and sine-wave drive" Conference Papers for RF Design '97, Santa Clara, CA, Sept 1997, pp. 283-297.

[iii] N. Sokal, HEPA-PLUS/WB v3.29, Design Automation, Inc., Lexington, MA

[iv] Sabin, Schoenike, et al, *Single Sideband Systems and Circuits*, McGraw-Hill 1993, p. 492.