How to Make Linear Mode Work

Jonathan Dodge, P.E.
Applications Engineering Manager
Microsemi Power Products Group
405 S.W. Columbia Street
Bend, OR 97702

Introduction

Linear mode is when a power transistor is operated partially on rather than fully on or fully off. The control or limiting of current through a transistor in linear mode combined with the ability to turn the transistor fully on are very useful for applications like electronic loads, circuit breakers, solid state relays, inrush current limiters, and even transient voltage suppression. The versatility of a power transistor enables combining these functions into a single unit, like a "smart" solid state relay that protects itself and the load.

Theoretically, linear mode operation is very easy. Simply bias the gate to deliver some desired amount of current or power, and stay within the manufacturer's Forward Safe Operating Area (FSOA) according to the datasheet. The reality however is that linear mode operation is one of the trickiest power applications of all, turning many "simple" designs into a reliability nightmare. This article points out the pitfalls of linear mode operation and provides guidance and examples for highly reliable linear mode designs. The discussion focuses on power MOSFETs but applies equally to IGBTs.

The Linear Mode Challenge

There are three fundamental challenges with linear mode design:
1. The information in manufacturers' datasheets is often inadequate or even incorrect concerning linear mode operation.
2. Thermal instability makes linear mode operation much more delicate than indicated by maximum power dissipation or die junction temperature ratings.
3. Insulated gate devices (MOSFETs and IGBTs) vary significantly in threshold voltage and transconductance from part to part.

Thermal Instability and Gain

The drain current is easily controlled by adjusting the gate-source voltage. However, there is inevitably some variation in temperature across the die, causing thermally-induced variation in current across the die. If the temperature-dependent change in current density is thermally unstable (which is normally the case in linear mode), the result can be hot spotting and current hogging within the die. This hot spotting phenomenon leads to a failure mode similar to second breakdown in bipolar transistors. This failure mode limits the true FSOA to a much smaller area than an FSOA based only on thermal resistance, which is often published in datasheets. Avoiding this failure mode is the biggest challenge with linear mode design. Therefore, it is worthwhile to understand the cause of this failure mode.

Linear mode operation is in the "saturation" region of the transfer characteristic (as opposed to the Ohmic region), as shown in Figure 1.
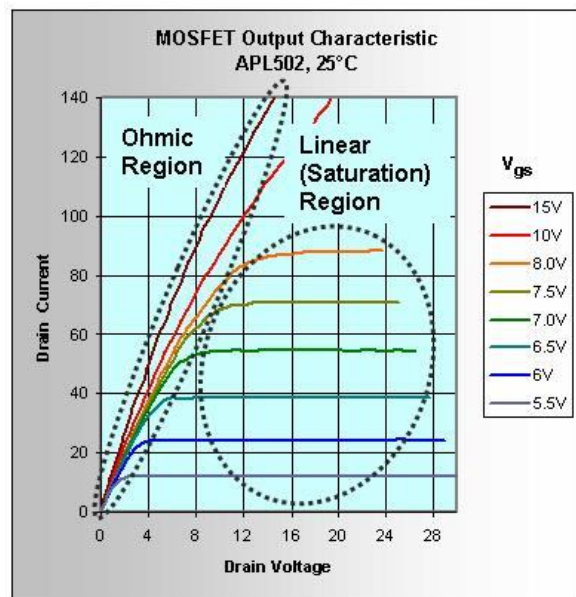

Figure 1 MOSFET Output Characteristic

In this area of operation, the drain current is related to the gate-source voltage v_{gs} and the threshold voltage V_{th} by the equation:

$$i_D = \kappa(v_{GS} - V_{th})^2 \qquad (1)$$

where:

$$\kappa = \frac{\mu_e \cdot C_{OZ} \cdot W}{2L}$$

and μ_e is the electron mobility, C_{OZ} is the gate oxide capacitance, W is the channel width, and L is the channel length. Gain and κ are related in that the wider the channel width W is and the shorter the channel length L is, the higher the gain is.

Since μ_e decreases with temperature, κ also decreases with temperature. (The capacitance does not change value with temperature but does with drain-source voltage). Likewise, V_{th} decreases as the temperature increases. As a device operating in linear mode heats up, the reduction in electron mobility tends to reduce the drain current, thus leading to thermal stability. This is countered by the reduction in threshold voltage, which tends to increase the drain current. The negative temperature coefficient of the threshold voltage leads to thermal instability.

These relationships can be expressed mathematically by differentiating (1) with respect to temperature, and substituting in the relationship for power dissipation and temperature $\partial T = \partial i_D \cdot R_\theta \cdot V_{DS}$, such that we have a stability factor S:

$$S = R_\theta \cdot V_{DS} \cdot \left[-2\sqrt{\kappa \cdot i_D}\, \frac{\partial V_{th}}{\partial T} + \frac{\partial \kappa}{\partial T} \cdot \frac{i_D}{\kappa} \right] \qquad (2)$$

The larger the value of S is, the more thermally *unstable* the device is, meaning that a localized temperature increase is regenerative. If S is negative, the device is thermally stable in linear mode. Note that the values of $\dfrac{\partial \kappa}{\partial T}$ and $\dfrac{\partial V_{th}}{\partial T}$ are always negative. From (2) we can see that a device becomes more thermally stable (S becomes smaller) as:

1. The thermal resistance is reduced
2. The drain-source voltage is reduced
3. The drain current increases
4. The gain (and hence κ) is reduced
5. The magnitude of the threshold voltage temperature coefficient $\dfrac{\partial V_{th}}{\partial T}$ is reduced

The fourth and fifth items depend entirely on the device design. Thus a device can be designed to be more thermally stable, resulting in a wider safe operating area in linear mode. This is what has been done for the Linear MOSFET series and most ARF series RF MOSFETs from Microsemi Power Products Group (formerly Advanced Power Technology).
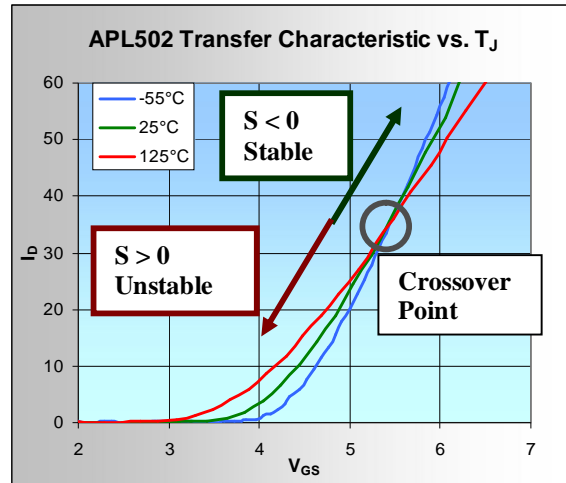
Figure 2 MOSFET Transfer Characteristic

Figure 2 shows the transfer characteristic of a MOSFET at three temperatures. This graphically illustrates the thermal stability factor expressed in equation (2). There is a crossover point through which each temperature curve passes. Below this point, the threshold voltage effect dominates, and localized changes in current are thermally unstable. Above the crossover point, the change in gain dominates, and the device is thermally stable.

The Failure Mechanism

Since the transfer characteristic crossover point is at a relatively high current, linear mode operation is almost always in the thermally unstable area below the crossover point. The problem is that the hotter areas on the die have higher current density, thus reinforcing hot spotting.

Every MOSFET and IGBT has an intrinsic bipolar transistor. The gain of this transistor increases as the device heats up, and as the drain-source voltage increases. The bipolar transistor base resistance also increases with temperature, and the base-emitter voltage drops. All these factors combine to increase the likelihood of generating enough voltage across the base resistance to turn on the bipolar transistor as the die heats up. Thus if a hot spot on the die gets hot enough, it can turn on the bipolar transistor in the area of the hot spot. This is what causes linear mode failure: the device latches up in a hot spot, and the resulting thermal runaway and extreme heating cause a burnout spot, shorting the drain to the source and possibly the gate to the source as well. Some damaged devices can actually be turned on, but when turned off

they can only support voltage with a huge leakage current through the damaged area.

Linear Mode Design Guidelines

The first step to create a reliable linear mode application is to contact a device manufacturer's applications engineer. You may find a treasure trove of information and tips that are not published in datasheets. The second step is to determine the true FSOA for devices you are seriously considering to use. Unfortunately this cannot be done by simulation because simulation models do not tell you when a device would actually fail. A number of devices must be tested to failure to determine the operating FSOA. This is a benefit of step one: this work is probably already done. Once you have a graph of failures at various voltages and the corresponding currents, you can create a curve fit or mathematical model for the data. By adding some safety margin, you get a true FSOA.

Table 1 shows linear mode power dissipation data for an APT200GN60J IGBT, where the collector-emitter voltage was fixed and the linear mode current was increased until the device failed. Results at several collector-emitter voltages were recorded.

| V_{CE} (V)	I_C (A)	Power (W)
500	0.227	114
450	0.25	113
400	0.338	135
350	0.413	145
300	0.473	142
250	0.565	141
200	0.68	136
150	1	150
100	1.84	184

Table 1 Linear Mode Failure Data: APT200GN60J

The parts were mounted on a liquid-cooled heat sink. The measured case temperature T_C was about 75°C at failure. Based on curve fitting, the average junction temperature at failure is estimated to be about 175°C, which happens to be the same as the rated maximum junction temperature. It is important to note that some devices fail in linear mode with an average junction temperature well below the rated maximum.
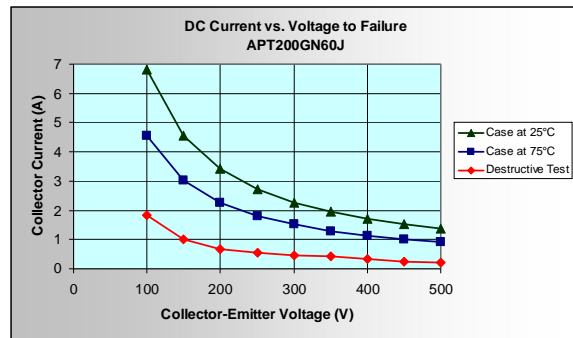
Figure 3 Measured & Theoretical FSOA

Figure 3 shows the data in Table 1 as well as the theoretical FSOA curves based on constant power dissipation with $T_J = 175$°C and $T_C = 75$°C and 25°C. Notice how much less the true FSOA is than what would be predicted based on constant power dissipation, limited only by thermal resistance (represented by the curves with $T_C = 25$°C and 75°C). Most datasheets publish an FSOA curve with the case at 25°C. Designing to that curve would result in six times more current at high voltage than the device can actually handle! Even derating to the measured case temperature of 75°C would result in too much current unless the hot spotting failure mode is taken into account. The only way to do that is to test several devices to destruction.
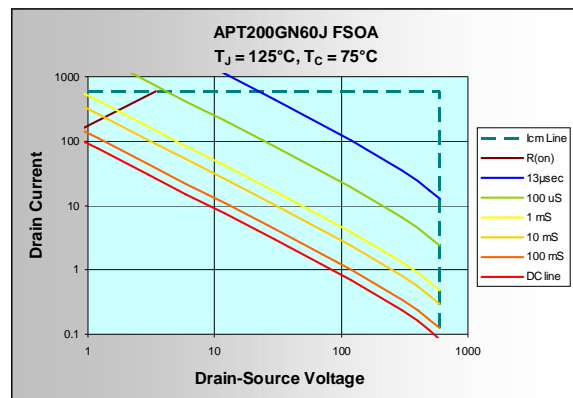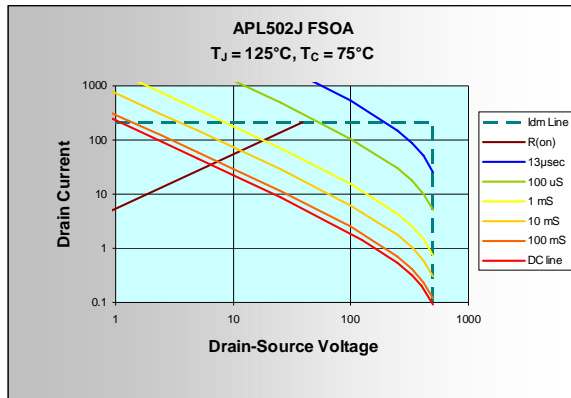

Figure 4 Usable FSOA: APT200GN60J

A curve fit was applied to the DC current destructive test data in Figure 3. Then pulse FSOA curves were generated based on transient thermal impedance data. The results are shown in Figure 4. This is a usable FSOA for the APT200GN60J part. By setting the junction temperature to 125°C (below where the part fails), some safety margin is created. Notice the DC curve is shifted down in current compared to the destructive test curve in Figure 3. It is always recommended to stay well away from the maximum rated junction temperature in linear mode, and there

should be at least 20°C margin from the average junction temperature at the point of failure. For Figure 4, 125°C is the maximum recommended junction temperature, providing a healthy 50°C margin from the failure temperature.

Now let us consider the FSOA of a MOSFET that was designed specifically for linear mode operation: the APL502J.

Figure 5 Usable FSOA: APL502J

Figure 5 shows the usable FSOA for the APL502J. Compared with the FSOA of APT200GN60J in Figure 4, the APL502J has a wider FSOA. The tradeoff is conduction loss versus FSOA ruggedness. Turned fully on and with a 200A load, the typical collector-emitter voltage of the APT200GN60J is only 1.7V when hot (1.5V at room temperature). By comparison, at 26A, 125°C, the APT200GN60J has about 6 times less conduction loss than the more rugged APL502J.

Notice in both Figures 4 and 5 that the FSOA curves drop at higher voltage (plotted on log-log scales). The FSOA curves based on constant power dissipation are straight lines. If you see a straight-line DC FSOA curve in a datasheet, beware! The graph is probably not usable for linear mode.

Design Examples

DC Solid State Relay

The APT200GN60J works well in a DC SSR application where it can operate in linear mode to limit current while charging a large capacitor bank, then operate fully on with minimal conduction loss. The capacitor charging current would need to be very limited to stay within the FSOA of the IGBT. This may not be an issue if there is no stringent charge time requirement.

Requirements

Suppose that we need to charge a 1500µF capacitor bank from 0 to 400V. We do not care how long it takes. The heatsink can keep the case temperature of the SSR at 75°C or less.

Solution

According to the FSOA graph in Figure 4, the current is most constrained at the maximum applied voltage, which is 400V in this case. From data used to create Figure 4, at 400V we can safely charge the capacitors with 0.16A (which according to Table 1 happens to be about half the current at the point of failure, so there is good safety margin). At 0.16A charge current, the capacitor bank will be charged from 0V to 400V in 3.75 seconds. Certainly it would be faster to charge the capacitors by following the DC SOA curve, thus increasing the charge current as the capacitor voltage rises (collector-emitter voltage falls). However, we don't care about charge time, and a constant charge current simplifies the control circuitry.

Staying within the DC FSOA is only half of the problem. The other thing to consider is the peak power dissipation and the resulting peak junction temperature. Since charge current is kept at a fixed (DC) value, the collector-emitter voltage will decrease linearly from 400V to almost 0V as the capacitor bank charges. So the power dissipation will peak at 64W (0.16A · 400V) as soon as voltage is applied and will decrease in a linear fashion, resembling a triangular waveform if plotted versus time.
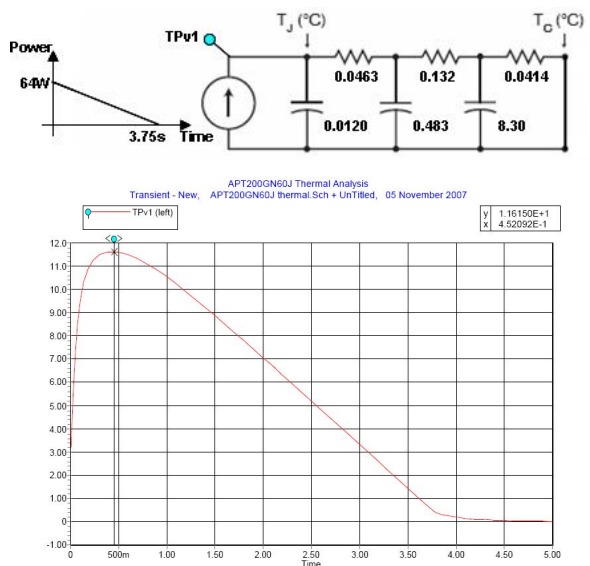
Figure 6 Transient Thermal Simulation

Figure 6 shows the result of a simulation using the APT200GN60J transient thermal impedance RC circuit model, applying a linearly decaying 64W peak power pulse. The peak junction-case temperature is about 12°C. If the case temperature reaches 75°C, the average junction temperature would reach 75°C + 12°C = 85°C, well below our maximum allowed 125°C.

Electronic Load

The Linear MOSFET APL502J works well where a wider FSOA is required, such as an electronic load. In this application, many parts would be connected in parallel to meet the power dissipation requirement as well as the maximum on-state voltage requirement.

Requirements

For this example, our homemade load has a working range up to 400W, 400V, 20A, and a fully-on voltage drop of 1V or less at 20A. The heatsink can keep the case temperature of the devices at 75°C or less.

Solution

We will use the FSOA curves in Figure 5, so we want to keep the junction temperature below 125°C. First check the on-state requirement. The APL502J has an $R_{DS(on)}$ of 0.090Ω maximum at room temperature (and 26A). At 125°C the $R_{DS(on)}$ is double, so 0.180Ω per part. The maximum total resistance allowed is 1V / 20A = 0.050Ω. The minimum number of parts required to meet the on-state voltage requirement is found as: $0.180\Omega / 0.050\Omega = 3.6$, so 4 parts minimum. Note that we need some voltage drop headroom for current sense resistors, which will be discussed.

Considering the FSOA limitation, the least amount of power can be dissipated at the highest applied voltage; in this case 400V. For the APL502J with the case and junction temperatures at 75°C and 125°C respectively, the maximum current per part at 400V is 0.2A at a power dissipation of 80W. We find the minimum number of parts to handle the 400W total load as: 400W / 80W = 5 parts minimum. All our worst-case conditions are met with a minimum of 5 APL502J parts in parallel.

It would be tempting to simply parallel the parts directly, installing a separate gate resistor to each part to prevent oscillation; and monitor the current from a single point. If we simply did that however, the result would certainly be failed devices.

Now we address the last remaining challenge of linear mode design: part-to-part variation of the threshold voltage. In linear mode, parts cannot be directly connected in parallel; each part must be forced to carry its share of current. This can be done in a number of ways.

If the maximum on-state voltage requirement allows it, a fairly high resistance can be installed in series with each MOSFET, thus carrying a significant portion of the thermal load (the resistors get hot). The resistors can also be used to somewhat balance the current between the MOSFETs by connecting a resistor between the source of each MOSFET and the gate drive return, providing negative feedback to each gate. Perfect balancing would not be possible. Also, sorting parts based on threshold voltage is not feasible since a tiny difference in threshold voltage between MOSFETs results in a substantial current mismatch.

With the low on-state voltage requirement of this design example, a cost-effective technique is to use a current sensor and an amplifier circuit to individually control the current by adjusting the gate-source voltage of each MOSFET. Figure 7 shows a conceptual schematic with three parallel MOSFETs. Low-value resistors or Hall Effect sensors must be used to keep the total voltage drop within specification.
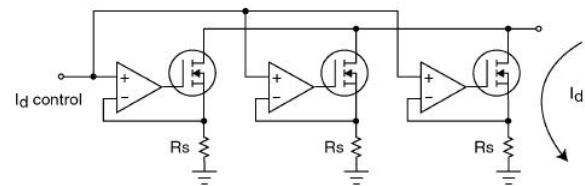
Figure 7 Linear Mode Paralleling Concept

To simplify assembly and minimize system size and cost, Microsemi Power Products Group is introducing a series of parts intended mainly for linear mode operation, although suitable for switch-mode applications as well. These parts, in a compact SP1 package, include a suitable power transistor (Linear MOSFET or Field Stop IGBT), a low-inductance current sense resistor in series with the source, and a temperature sensor.

Figure 8 Transistor, Current Sense Resistor, and Temperature Sensor in SP1 Package

The integrated current sense resistor is mounted on the same ceramic insulator as the power transistor, minimizing inductance and providing cooling for the resistor, which dissipates only a few Watts at maximum load. This makes it simple to monitor the drain-source voltage, the drain current, and the case temperature simultaneously. This information can be processed digitally such that the FSOA curve can be followed, allowing full utilization of the device and minimal system cost.

References

[1] David Blackburn; "On the Thermal Instability of Power MOSFETs", National Bureau of Standards

[2] Alfio Consoli, Francesco Gennaro, Antonio Testa, Giuseppe Consentino, Ferruccio Frisina, Romeo Letor, Angelo Magri; "Thermal Instability of Low Voltage Power MOSFETs", IEEE Transactions on Power Electronics, Vol. 15, No. 3, May 2000

[3] H. Yilmaz, T. Tsui, I. Bencuya, T. Fortier, K. Owyang; "Safe Operating Area of Power DMOS FETs", IEEE Power Electronics Specialists Conference, 1989

[4] Harold Ronan Jr., "MOSFET SOA – Thermal Current Focusing Affects Power MOSFET Operation", PCIM Magazine, August 1998