Parallel Operation of Power MOSFETs

By Rudy Severns

INTRODUCTION

There are many reasons for operating power MOSFETs in parallel, either as multiple die on a common substrate or as individually packaged devices. Compared to paralleled bipolar junction transistors (BJTs), paralleled MOSFETs present fewer problems, require less derating, and provide higher performance. Parallel operation, however, is still not trivial. Problems can and do occur. To parallel devices successfully, the designer must understand both the causes and the cures for the problems. Fortunately, most of the cures are simple and are more a matter of attention to detail than of exotic or complex design.

This discussion attempts to identify all the problems that users potentially could experience and provides an explanation of each problem. Then it will tell how to recognize these problems and will give practical means for eliminating them.

Although the discussion is quite extensive, the final conclusions will show that all of the problems may be easily avoided and that paralleling MOSFETs is reasonable and, in many cases, highly desirable. However, because of the lack of prior, in-depth discussion of paralleling MOSFETs, it is necessary to substantiate this contention.

Throughout the discussion, the object will be to maintain a balance between simple, practical advice and reasonable, theoretical explanations.

Many designers experienced in BJT applications have a well merited aversion to paralleling devices unless it is absolutely necessary. While this is certainly appropriate for BJTs, such thinking is a distinct handicap when using MOSFETs because useful opportunities for improved performance or reduced cost may be missed. It is important to keep an open mind and to read carefully the following arguments.

MOTIVATION FOR PARALLELING

There are many possible reasons for paralleling multiple devices:

- Lower $R_{DS(on)}$
- Lower circuit inductance
- Improved thermal performance
- Compensation for radiation effects
- Lower cost
- Redundancy
- Higher I_D
- Derating

In switching applications, the majority of the power dissipation is due to $R_{DS(on)}$. If $R_{DS(on)}$ is made smaller, the power loss will go down. $R_{DS(on)}$ for a given device is determined by the active die area and the breakdown voltage (BV_{DSS}). $R_{DS(on)}$ decreases more or less linearly with increasing die area and increases exponentially (factor ≈ 2.5)

with BV_{DSS}. In theory, $R_{DS(on)}$ may be made as small as desired simply by increasing the die area, but the cost per unit area of the die increases exponentially when the die dimensions are greater than about 0.125″ × 0.125″. This increase in cost puts a practical upper limit on die size. Presently this is in the range of 0.25″ to 0.30″ square. In addition, as die are made larger, fewer and fewer package choices are available. When $R_{DS(on)}$ must be reduced even further, paralleling is the best alternative.

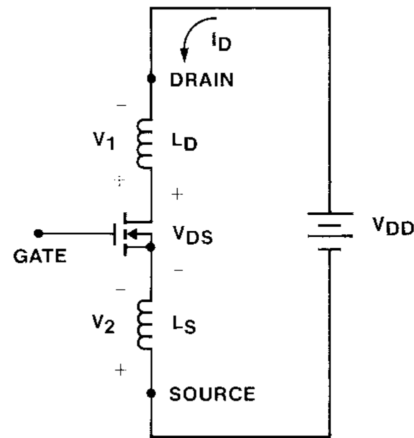
Because of the rapid increase in cost of the larger die and the restriction on suitable packages, there are often compelling reasons to use smaller die in cheaper packages even when a suitably low $R_{DS(on)}$ device is available in a larger die. For example, a size five die (0.25″ × 0.25″) will have one half the $R_{DS(on)}$ of a size four die (0.18″ × 0.18″). However, the size five die will not fit in a TO-220 package. It requires a TO-3. At current prices, the cost of two TO-220s would be twenty to twenty-five percent less than one TO-3. As prices drop further, production increases and competition grows, this price differential should grow wider.

It has been shown (1) that paralleling is an effective means for reducing the junction-to-heatsink thermal impedance ($R_{\Theta js}$). To use the previous example, $R_{\Theta js}$ for a size five die in a TO-3 package is about 1.2°C/W; for a size four die in a TO-220 package, $R_{\Theta js}$ = 1.3°C/W; but two TO-220's in parallel would reduce the effective $R_{\Theta js}$ to 0.65°C/W. For the same $I_{D(rms)}$ and heatsink, $R_{DS(on)}$ will be lower because T_j will be lower. This may allow a further reduction in die size or as an alternate, the heatsink may be made smaller thus reducing its cost. The thermal design issues are treated in detail in Reference (1). For high power pulse applications, the limitation may be on $I_{D(peak)}$ rather than $R_{DS(on)}$. Again paralleling provides a means for increasing the allowable peak current. Very often in pulse applications, rapidly rising (high di/dt) current waveforms are required. Sometimes the limiting factor on speed is the package inductance. One means of reducing this inductance is to parallel several smaller devices so that the total effective inductance is decreased in proportion to the number of devices paralleled. Also smaller packages will tend to have lower inductances.

Another problem which can arise in fast pulse applications— if low voltage high current devices are used — is a degradation in the external BV_{DSS} capability caused by voltage spikes generated by the internal connection inductance (2). Figure 1 shows a MOSFET with the internal parasitic inductances. At turn-off, I_D is flowing. As the device is turned off, the voltage polarities will be as indicated. The actual voltage across the junction will be:

$$V_{DS} = V_{DD} + (L_D + L_S)\frac{dI_D}{dt}$$

For example, if a 50 ampere, 100 V device with 20nH of internal parasitic inductance is switched in 20nsec, the internal voltage spike between the drain and source terminals (which is not



Mechanism Responsible for Internal Voltage Spiking at Turn-off
FIGURE 1

visible externally) will be 50 V! The usable value for BV_{DSS} has been reduced to 1/2 its data sheet value. Paralleling several smaller devices, possibly in smaller, lower inductance packages, can reduce the effective value of parasitic inductance and increase the usable operating voltage.

When MOSFETs are subjected to large neutron fluxes (> 10^{13} neutrons/cm^2), $R_{DS(on)}$ may rise dramatically. One means to ensure that the post irradiation $R_{DS(on)}$ is not unacceptably high is to start with a low value for $R_{DS(on)}$ (much less than required) in the unirradiated circuit. To achieve this, it is usually necessary to parallel devices.

CONCERNS ABOUT PARALLELING DEVICES

When considering a design using parallel FETs, a number of concerns and possible problems arise. Generally these concerns fall into four categories:

1. Steady-state current sharing
2. Thermal stability and maximum T_j
3. Dynamic current sharing during switching
4. Parasitic oscillations

As is so often the case, many problems that exist in theory are not significant in the actual hardware. Nevertheless, in the following discussion, many possible problems will be explained.

STEADY-STATE CURRENT SHARING

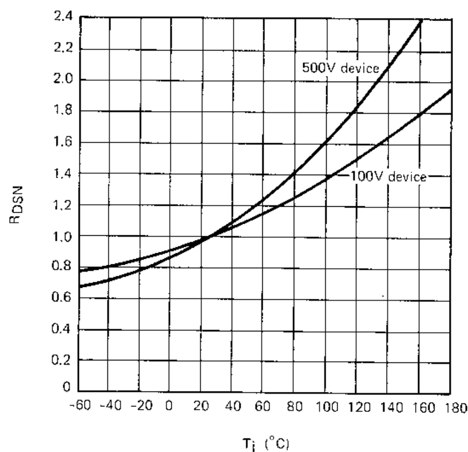
The distribution of current among parallel devices is a concern to the designer. With regard to the effects of asymmetrical current sharing, three questions need to be answered:

1. What is the maximum junction temperature among the devices?

2. Does the asymmetry cause a significant increase in total dissipation?
3. Is any device operating outside of its safe operating area (SOA)?

When the current is not distributed equally, some devices may run hotter than others. Because the operating reliability is directly related to T_j, it is important to identify the maximum T_j which can occur. Obviously if any device is operating outside of its rated SOA, the reliability will also be greatly reduced. Furthermore, it is important to know if the conduction asymmetry is creating a power loss penalty.

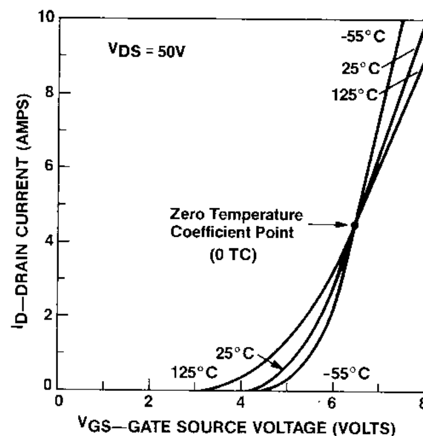
The answers to these questions depend on the operating state of the devices. There are two possibilities: linear or switching operation. When the devices are used as switches, V_{GS} will be large (10-15V), and the devices will be fully enhanced. In this mode the device acts like a positive temperature coefficient resistor. Typical $R_{DS(on)}$ versus T_j characteristics are shown in Figure 2. Note that $R_{DS(on)}$ has been normalized (R_{DSN}) to 25° C for comparison purposes.



The Relationship Between $R_{DS(on)}$, T_j and BV_{DSS}
Figure 2

When the devices are operating in the linear mode, the behavior is quite different. Figure 3 shows a typical transfer characteristics graph (I_D as a function of V_{GS}) with T_j as a parameter. The interesting feature of this graph is that above 4.5A, the temperature coefficient (TC) is positive, but below 4.5A, the TC is negative. For this device $R_{\Theta js}$= 1° C/W, and if a perfect heatsink is assumed such that t_c = 25° C then the maximum value for V_{DS} at 4.5A is:

$$V_{DS} = \frac{T_j - T_c}{R_{\Theta jc}\, I_D} = 27.8\ \text{V} \qquad (2)$$

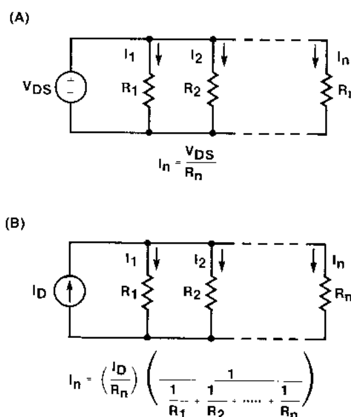


Typical Transfer Characteristics for an IRF330
FIGURE 3

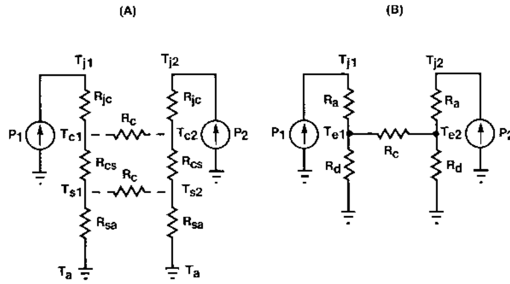
Since this is a 400V device, it is unlikely that the device would be used in the linear mode with such a low value of V_{DS}. Most linear applications would use the device at currents well below 4.5A to exploit the BV_{DSS} capability (400V), and therefore, they would be operating in the negative TC region. Since this is exactly the opposite of the switch mode, the two types of applications will be treated separately.

Current Sharing While Fully Enhanced

When V_{GS} is large (6-8 V above V_{th}), a MOSFET is essentially a positive TC resistor, and the current will divide among the paralleled devices in proportion to their individual $R_{DS(on)}$ as illustrated in Figure 4. As was shown in Figure 2, the TC will be positive, so there is a tendency for a device which has a greater than average current to heat up more than the other devices, increasing $R_{DS(on)}$, which in turn reduces its I_D. The degree of thermally forced current sharing present has been analyzed (3), and an outline of the analysis is presented here.



Current Sharing in Parallel MOSFETs
FIGURE 4



Thermal Models for Two Parallel MOSFETs
FIGURE 5

A thermal model for two parallel devices is given in Figure 5A. The normal thermal impedances for junction to case ($R_{\Theta jc}$), case to heatsink ($R_{\Theta cs}$) and heatsink to ambient ($R_{\Theta sa}$), are present as well as the thermal coupling between the devices (R_c). Three possibilities are indicated:

1. R_c between nodes T_{c1} and T_{c2} represents die mounted on a common case or header.
2. R_c between nodes T_{s1} and T_{s2} represents the situation in which separate packaged devices are paralleled on a common heatsink.
3. The situation where there is no common coupling between devices, i.e. separate heatsinks, is represented by $R_c = \infty$.

For analytical purposes, the model in Figure 5A can be simplified as shown in Figure 5B where the values for R_a and R_d depend on the position of R_c. For example, for devices paralleled on a common heatsink:

$$R_a = R_{\Theta jc} + R_{\Theta cs} \quad (3)$$

$$R_d = R_{\Theta sa} \quad (4)$$

From this network, it is possible to calculate the values of $R_{DS(on)}$ for each device (R_1 and R_2). Given the values for R_1 and R_2, one can then calculate the values of I_D (I_1 and I_2), power dissipation (P_1 and P_2), and the junction temperatures (T_{j1} and T_{j2}). Kassakian (3) has derived the following expressions for this network.

$$R_1 = R_{10}\left\{1 + \frac{I^2 R_1 R_2 A}{(R_1 + R_2)^2}\left[R_2\left(\frac{(R_d + R_a)(\frac{R_c}{R_d} + 1) + R_a}{\frac{R_c}{R_d} + 2}\right) + R_1 \frac{R_d}{\frac{R_c}{R_d} + 2}\right]\right\} \quad (5)$$

$$R_2 = R_{20}\left\{1 + \frac{I^2 R_1 R_2 A}{(R_1 + R_2)^2}\left[R_1\left(\frac{(R_d + R_a)(\frac{R_c}{R_d} + 1) + R_a}{\frac{R_c}{R_d} + 2}\right) + R_2 \frac{R_d}{\frac{R_c}{R_d} + 2}\right]\right\} \quad (6)$$

Where:

R_{10} = 25°C value for $R_{DS(on)}$ in device #1
R_{20} = 25°C value for $R_{DS(on)}$ in device #2
A = Temperature coefficient of resistance. This can vary from 0.5 to 3%/°C depending on the device.
I = $I_1 + I_2$

By itself, this set of equations is not generally useful since the equations are non-linear: i.e. they contain cross products and powers of R_1 and R_2 (the dependent variables). Using numerical methods, this type of equation is usually solved on a computer. However, in most cases this is not necessary as the following example solution demonstrates.

Assume for example, that 10A, 100V devices are being used which have the following characteristics:

$R_{\Theta jc}$ = 1.67°C/W R_{10} = 0.12 Ω
$R_{\Theta cs}$ = 1.00°C/W R_{20} = 0.16 Ω
$R_{\Theta sa}$ = 1.47°C/W   I = 20 A
T_c = 0.67 %/°C

Assume further that the die are mounted on the same header:

R_a = 1.67°C/W
R_d = 2.47°C/W

If Equations (5) and (6) are solved for values of R_c from 0 to 100°C/W, the graph shown in Figure 6 results.

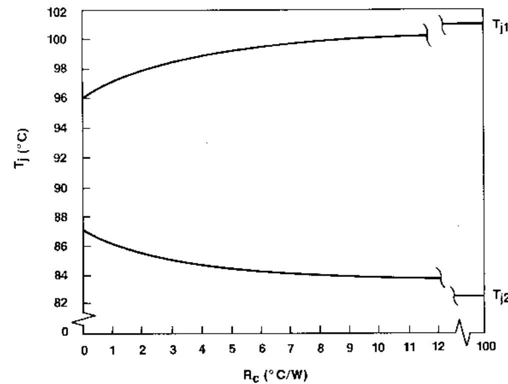


FIGURE 6

Notice that when R_c is small the difference between T_{j1} and T_{j2} is also small (about 9°C) even though R_{20} is one-third larger than R_{10}! The effect on current sharing is almost non-existent (about 0.3A).

As long as R_c is comparable to or smaller than R_a, the difference in junction temperatures should remain small. In practice this should not be difficult to achieve. For multiple die on a common header, this requirement holds very well. For

multiple individual packages mounted on a common heatsink, this requirement can be met by having a reasonably heavy common web between devices and by mounting these devices relatively close to each other.

In the past, many people (this author included) have touted thermally forced current sharing as a major feature of parallel device operation in MOSFETs. Yet thermally forced current sharing is not prominent except in the case of very high voltage devices with large TCs and separate heatsinks.

The most common applications for MOSFET switches involve a load impedance that is large compared to $R_{DS(on)}$. In these applications, the total current (I) through all of the devices is determined by the load. In this case, the total power loss is:

$$P_T = I^2 R_T \tag{7}$$

where: R_T = the total on-resistance of the parallel devices.

In this case, the distribution of current between the devices is not considered; all that matters is the final value for R_T. The fact that some devices will be slightly warmer than others will increase R_T slightly, but this is usually a second order effect.

From these analyses and other work (5), the following general observations can be made:

1. The current sharing between parallel devices is in proportion to $R_{DS(on)}$.
2. When the devices are well coupled thermally, the differences in T_j are small even when the difference in $R_{DS(on)}$ is substantial.
3. For thermally coupled devices, forced current sharing is insignificant.
4. For high voltage thermally uncoupled devices, some forced sharing can occur, but since this is achieved at the expense of higher T_j and R_T, in most cases thermally uncoupled parallel operation is undesirable.
5. Parallel devices should be mounted on a common heatsink or substrate with a minimum common thermal impedance (R_c).
6. Matching of devices for $R_{DS(on)}$ is usually not necessary unless the range of variation (±20%) would allow too large a value for R_T. Rather than matching or screening devices for $R_{DS(on)}$, it is frequently cheaper and easier to add another device in parallel.
7. If a limit on the maximum value for R_T is desired and matching is acceptable, the matching for $R_{DS(on)}$ should be done at the anticipated average current for each device and the planned value for V_{GS}.
8. To minimize differences in $R_{DS(on)}$ and the final value for R_T, it is important to enhance the devices fully. A V_{GS} of 10 to 15 V is sufficient.
9. The relevant issue in paralleling is not current sharing, per se, but rather the junction temperature differences and any additional power losses. If the ΔT_j and ΔP_T are small, the asymmetry in the current is irrelevant.

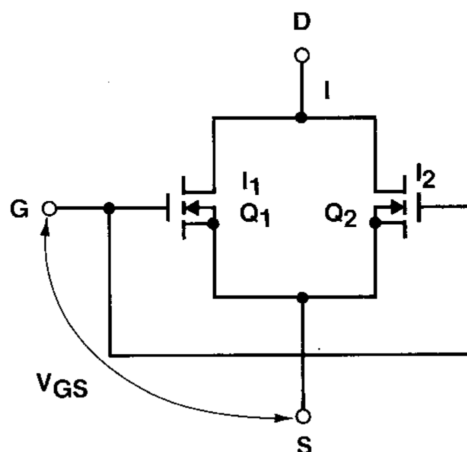
Current Sharing During Linear Operation

In the linear mode of operation, the temperature coefficient of ON-resistance is negative. This means that if one device, in a group of parallel devices, is conducting more than its share of current, its temperature will rise. This will further increase this device's share of the total current. This process can lead thermal runaway and is very similar in nature to the thermal instability present in BJTs. In the case of the MOSFET, the transconductance ($g_m = \Delta I_D/\Delta V_{GS}$) is much lower, and the tendency towards instability is correspondingly less.

Thermal regeneration of this type combined with normal device characterization variations can cause three problems:

1. Large differences in current sharing can occur. The current distribution among parallel devices will vary with temperature and, in some cases, with total current (I).
2. The quiescent operating or "Q" point is ill-defined and varies with temperature.
3. Thermal runaway and subsequent device failure are possible.

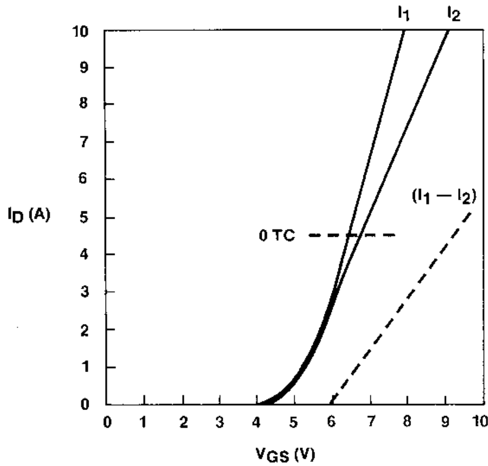
The following discussion will examine these problems and demonstrate a simple cure. For the purposes of this discussion, two parallel devices will be used, but the principles exposed apply to multiple parallel devices. Figure 7 shows the reference model for the discussion.



Two MOSFETs in Parallel, Reference Circuit
FIGURE 7

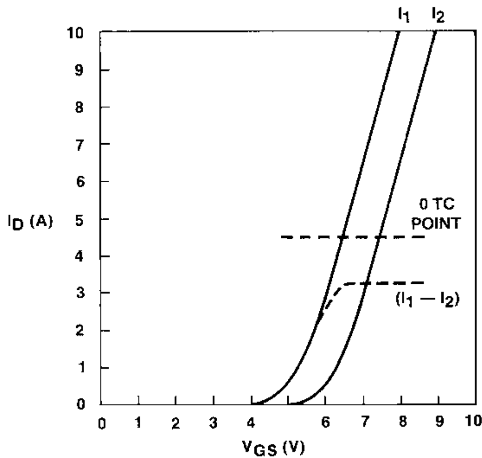
When two non-identical devices are paralleled, a variety of situations can arise depending on the differences between the devices. Figures 8, 9, and 10 show several possibilities as well as the resulting current imbalance. For the moment, the effect of T_j changes will be ignored.

Figure 8 illustrates the effect of differences in g_m. In this case, the current differential increases as I_D is increased.



Current Difference Between Two MOSFETs where gm is the same but Vth is different
FIGURE 8

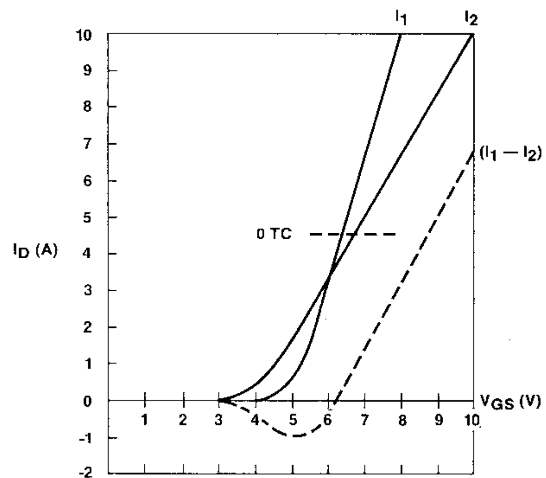
Figure 9 illustrates the effect of a one volt difference in V_{th}. In this example, all of the current flows through one device until V_{GS} reaches 5 V. At that point, the second device begins picking up some of the current. The current asymmetry remains essentially constant for $V_{th} > 6.5V$. The problem here is that if the "Q" point is below 3A, one device will hog nearly all of the current!



Current Difference Between Two MOSFETs where gm is the same but Vth different
FIGURE 9

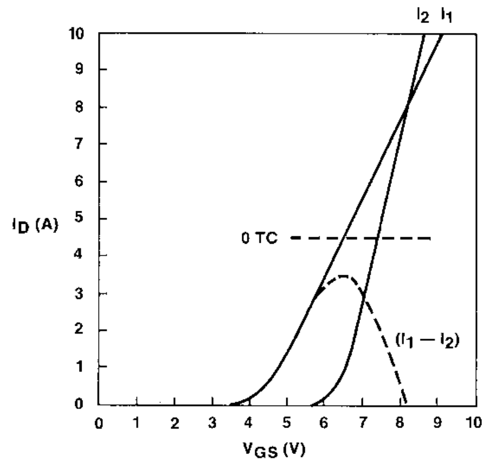
Most practical applications will represent some combination of the two previous examples. A typical combination is shown in Figure 10. In this particular example, the device carrying the greater portion of the current will depend on the "Q" point.

Now it is necessary to include the effect of the junction temperature on the current imbalance. Looking again at Figure 9, notice that $I_1 > I_2$. It is reasonable to assume that Q_1 will heat



Current Difference Between Two Mismatched Devices
FIGURE 10

up and Q_2 will cool down. An approximation of what will occur is shown in Figure 11. From this figure, it can be seen that the peak value of the difference between I_1 and I_2 is increased, and the range of V_{GS} over which Q_1 takes all or most of the current is expanded. In most applications, the effect of differential heating in linear operation is to make the problem worse. One way to minimize the effect of asymmetrical currents is to maximize the thermal coupling between the devices (make R_c as small as possible). This is the same conclusion reached for parallel devices operating as switches!



Current Imbalance from Figure 8 when Heating Effects are taken into account
FIGURE 11

Good thermal coupling by itself, however, is not enough to assure good current sharing; it merely reduces the degree of mismatch. For linear applications, some further means to force sharing must be taken. One obvious solution is to match devices. Unfortunately to get really good current sharing, it is necessary to match the entire transfer characteristic. This level of matching would rarely be practical. A useful compromise would be to match the devices at the "Q" point. This is fine in a stable thermal environment, but if the ambient temperature (T_a) varies over a wide range then it is unlikely the devices will remain matched. In any event, matching of devices can be costly and a considerable nuisance in production.

A better solution would be to take randomly selected devices and force them to share. This can be done most easily by using small source resistors to provide negative feedback, as shown in Figure 12A. As an added advantage, this will also stabilize the "Q" point.

The effect of the source resistor can be quantified by examining its effect on the "Q" point of a single device as indicated in Figure 12B. It has been shown (6) that the effective g_m (g_m') will be:

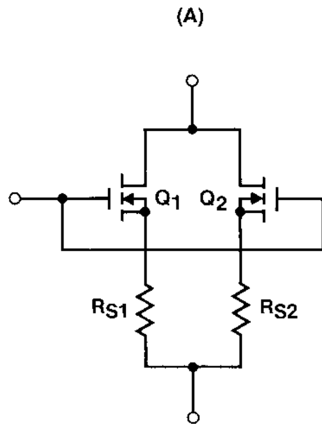
$$g_m' = \frac{g_m}{1 + R_s g_m} = \frac{1}{R_s + \frac{1}{g_m}} \quad (8)$$

To minimize the effect of differences in g_m' it is necessary that:

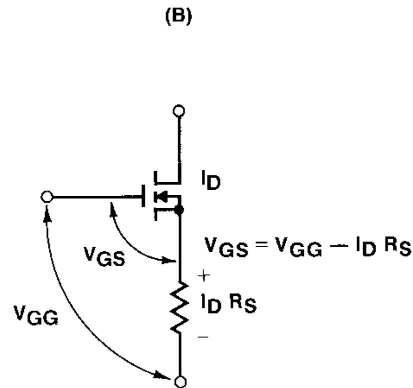
$$R_s >> 1/g_m \quad (9)$$

Typical devices will have values for $1/g_m$ in the range of 0.1 to 1.0 ohms.

(A)

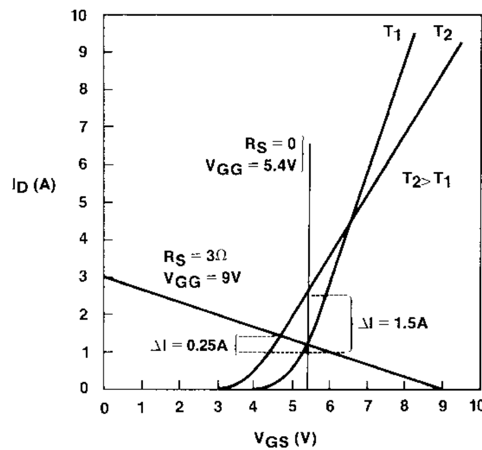


(B)



Forced Current Sharing using Source Resistors
FIGURE 12

The effect of R_s on the "Q" point stability (with fixed V_{GG}) is shown in Figure 13. If the operating point is selected as $V_{GG} = V_{GS} = 5.4V$ and $R_s = 0$ then $I_D = 1.2A$ at T_1, but it will rise to $I_D = 2.7A$ at T_2 — an increase of more than 2:1! By making $R_s = 3$ ohms (as indicated by the sloping line), the change in $I_D(\Delta I)$ when the temperature changes from T_1 to T_2 is reduced to 0.25A! This is a very great improvement. In this example, g_m is approximately 2.5S so that $R_s \approx 8/g_m$. Even if R_s is reduced to 1Ω, Δ I is still only 0.6A.



Operating Point Stabilization
FIGURE 13

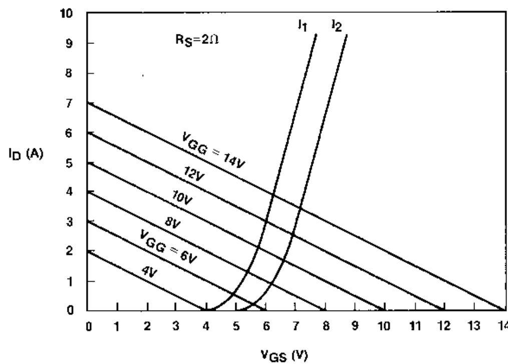
For those readers not familiar with the graphic technique just used, a few words of explanation may prove helpful. The straight line representing R_s = 3 and V_{GG} = 9V is shown in a graph representing the values of V_{GS} for given values of I_D. For example, if I_D = 1.667A, the drop across R_s is 5V, viz. 3 × 1.667A. From Figure 12 we know that

$$V_{GS} = V_{GG} - I_D R_s \qquad (10)$$

so that V_{GS} in this case is 4V.

What we are seeking is a simultaneous solution to Equation (10) and the equation represented by the graph of the transfer characteristic. This is done by graphing Equation (10) (the straight line) on the transfer characteristic graph and by noting the intersection of the graphs. In Figure 13, the intersection representing the "Q" point is at V_{GS} = 5.4V, and I_D = 1.2A.

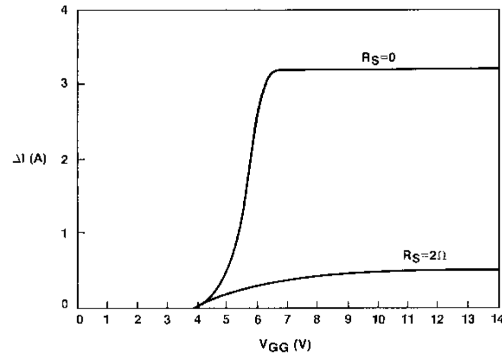
This same technique can be applied to parallel devices since stabilization of the "Q" point also stabilizes the current sharing. Using the example of Figure 9, the effect on sharing by adding a 2 ohm resistor in series with each device can be determined. A series of parallel lines are drawn for several different values of V_{GG}, each having a slope corresponding to 2 ohms. Each line establishes a "Q" point with a given value of ΔI. The values for ΔI corresponding to the different values for V_{GG} can now be plotted and compared to the values for the original example (Figure 9). This is shown in Figure 15 where it can be seen that ΔI is reduced from 3.2 to 0.5A! Clearly this is an effective means for equalizing the current distribution as well as stabilizing the operating point.



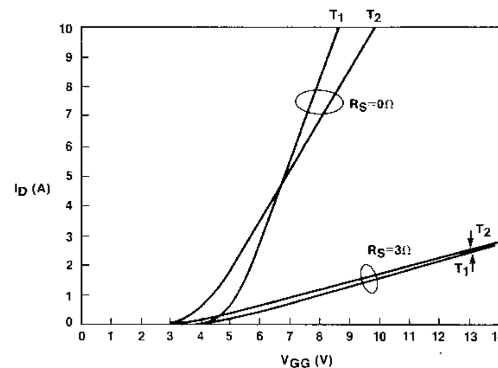
Rs Load Lines Plotted on the Transfer Characteristic Graph
FIGURE 14

Source resistors can also improve thermal stability by reducing differences between the transfer characteristics as T_j is varied. This effect can be illustrated by regraphing the example in Figure 13 as a function of V_{GG} rather than V_{GS}. This is done in Figure 16. In effect, the thermal regenerative gain is much lower and the stability greatly enhanced.

Note that Figure 16 is also an alternative method for determining ΔI between non-identical devices; however, it is usually simpler to draw the resistive load lines for R_s, as was done in previous examples (Figures 13 and 14), than to redraw the graph in the form used in Figure 16.



The Effect of Source Resistors on Current Sharing
FIGURE 15



The Effect of Rs on the Transfer Characteristics over Temperature
FIGURE 16

The use of source resistors has several advantages:

1. The operating point is stabilized.
2. Excellent current sharing and equalization of power dissipation is achieved.
3. Transconductance differences are minimized so that the small signal gain is nearly the same in each device.
4. Thermal stability is enhanced.
5. The small signal linearity is improved.

Unfortunately, these benefits are not gained without some cost. The use of source resistors has the following disadvantages:

1. Additional components are needed: i.e. the resistors.
2. The large signal dynamic range is reduced by the voltage drop across the resistors, but since this voltage drop is usually small compared to V_{DS} this is not a severe penalty. On the other hand, the source resistors tend to extend the lower limit of the large signal linear region (Figure 16), and this compensates by extending the dynamic range.
3. The voltage gain of the stage is reduced. In the common source configuration, the voltage gain (A) is:

$$A = g_m R_L \qquad (11)$$

where R_L is the load resistance.

From Equation (6), the gain will be reduced to:

$$A = \frac{g_m R_L}{1 + R_s g_m} \qquad (12)$$

when source resistors are used.

From the foregoing discussion, the following general observations regarding parallel devices operating in the linear mode can be made:

1. This mode of operation is very different from the switching mode in most respects.
2. To avoid excessive current asymmetry and thermal instability, some positive means must be provided to stabilize the operating point and force current sharing.
3. The simplest means to achieve the above goals is to use small resistors in series with each device's source lead.
4. Good thermal coupling between devices will greatly improve thermal stability and current sharing and help to minimize the size of R_s.
5. Matching of devices, while useful and effective, is usually not necessary. Small source resistors are usually a cheaper and simpler solution. The size of the resistors can be reduced by prescreening devices to eliminate those with larger than average characteristic deviations.
6. Multiple die devices will normally (at least at Siliconix) have the die selected from adjacent positions on the same wafer and will be well matched. In addition, the thermal coupling will be very good. The user cannot, however, add individual source resistors because the die are sealed within the case. If the degree of intrinsic matching in multiple die devices is not adequate, the single die devices can be paralleled using individual source resistors.

Current Sharing During Switching Transitions

When parallel MOSFETs are used as switches, several questions arise concerning device behavior during the switch transitions from OFF to ON and from ON to OFF:

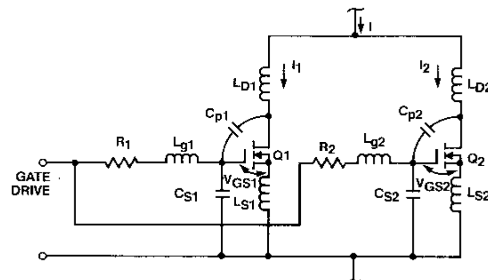
1. What is the current distribution?
2. Can current assymmetries be severe enough to damage a device?
3. Is the switching time affected by current asymmetries?

When a MOSFET makes a transition from one state to another, the device must pass through the linear region, if only momentarily. This means that much of what has been discussed concerning the linear mode of operation applies to dynamic switching. However, there are some differences:

1. The operating time in the linear region is very short (typically 10 to 100 nsec).
2. Because of the rapid transitions, thermal heating effects are usually negligible.
3. The use of source resistors is unacceptable because of the substantial increase in the effective value of $R_{DS(on)}$.
4. The external parasitic and intentional resistances, inductances, and capacitances must be taken into account.

An equivalent circuit representing two parallel FETs is given in Figure 17. Differences in current during switching can be caused by:

1. Differences in the external circuit elements.
2. Differences in capacitive or inductive device characterstics.
3. Differences in device characteristics during linear operation.

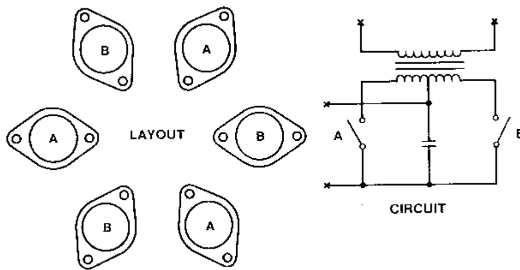


Equivalent Circuit for Two Parallel MOSFETs During Switching Transitions
FIGURE 17

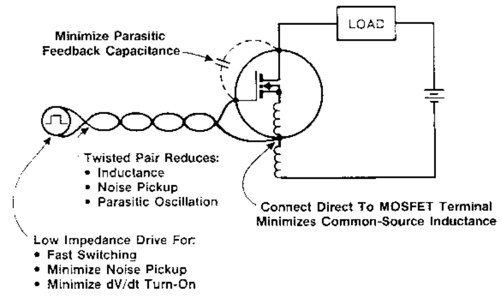
If for example, Q_1 and Q_2 are identical, but the drive circuit impedances are different, Q_1 and Q_2 will not turn on simultaneously. This problem, however, is under the control of the designer. The following steps can be taken to minimize this problem.

1. Minimize the values for R, L_g, L_s, L_D, C_p and C_s.
2. There will be practical limits on how small the external impedance can be. When that limit is reached, every effort should be made to equalize the values for the remaining impedances. Using symmetrical layouts is one way to do it. An example of a symmetrical layout for a parallel switch connection (push-pull) with three paralleled devices is given in Figure 18. Many other practical symmetrical layouts are possible.

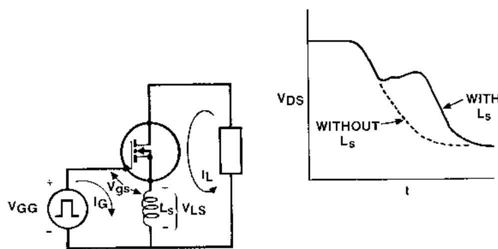
3. Good layout techniques are vital. The effect of L_S on the switching waveform is shown in Figure 19. Obviously L_S should be made as small as possible. Further examples of good layout technique are given in Figures 20 and 21.



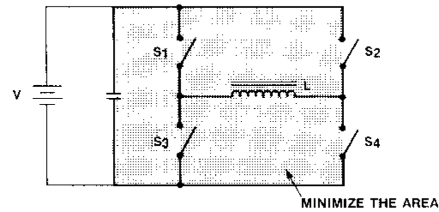
Symmetrical Layout Example
FIGURE 18



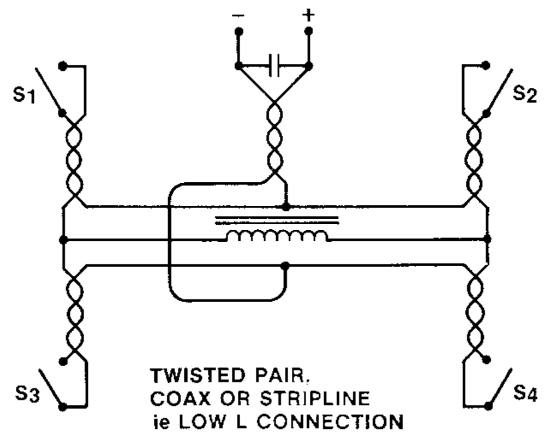
Suggested Drive Circuit Layout
FIGURE 20



The Effect of Common-Source Inductance on Switching
FIGURE 19



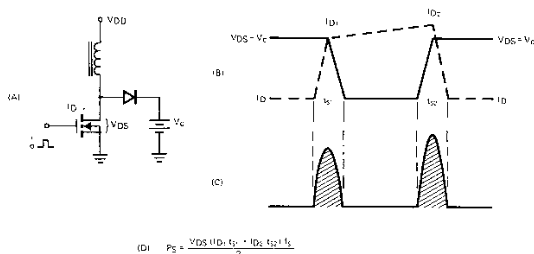
High di/dT Paths in an H-Bridge
FIGURE 21A



Low Inductance Connections for High di/dT Paths
FIGURE 21B

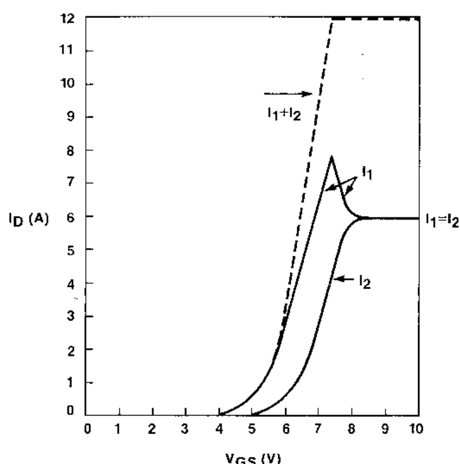
Differences in package inductance and inter-terminal capacitance are amazingly small (≈±5%) and usually can be ignored. This is fortunate since without a complex matching procedure, there is little that can be done anyway.

Differences in device characteristics, on the other hand, can cause differences in current distribution during switching transitions. To illustrate what occurs, the techniques previously used for linear operation can be applied to the example in Figure 9. If a clamped inductive load is assumed, Figure 22 shows the switching waveforms. If $I_D = I_1 + I_2 = 12A$ then the current during switching will assume the value shown in



Clamped Inductive Load Switching Waveforms
FIGURE 22

Figure 23. If V_{GS} is a linear function of time (a ramp) then the graphs for I_1 and I_2 become their current switching waveforms as a function of time! The salient feature of these waveforms is that even though the two devices have very different threshold voltages, the current spike for Q_1 is small (about 1.7A). In most cases, when unmatched devices are paralleled, only relatively small current differentials are observed. A good series of actual oscillographs is given in Reference (5) to illustrate this point. Applications do exist where from 20 to 50 devices are paralleled. In these applications, it is theoretically possible (but relatively unlikely) to generate destructive current spikes. Such spikes can be avoided by prescreening devices to eliminate those devices that are radically different from the median.



Values for Drain Current during Changed Inductive Switching when I, + I_2 is limited to 12A
FIGURE 23

The following observations can be made regarding switching in parallel MOSFETs:

1. Careful circuit layout is needed to minimize and equalize parasitic impedances.
2. Minimizing differential gate impedances will equalize V_{GS} for each device.
3. Except for large numbers of devices, the current spikes during switching transitions will usually be well within the limits of the device capability.
4. If many devices are paralleled, simple prescreening which measures I_D at a given value of V_{GS} in the transition region (i.e. not fully enhanced) should be sufficient. This will eliminate the unusually different devices.

PARASITIC OSCILLATIONS IN MOSFETS

Most power MOSFETs presently available are very fast devices with appreciable gain at frequencies up to 300 MHz. This high frequency gain, coupled with the internal and circuit parasitic inductances and capacitances normally present, make it possible for unwanted parasitic oscillations to occur. The frequency of oscillation can range from 1 to 300 MHz.

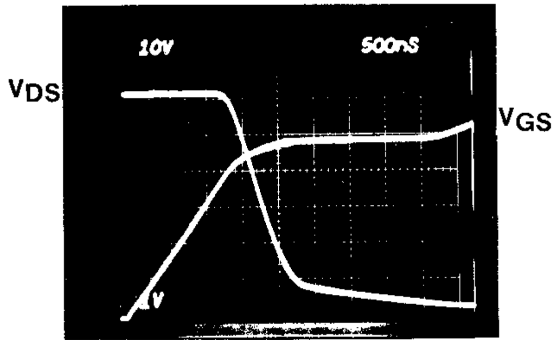
The oscillations occur while the device is in the active mode where the transconductance is large. When the device is off or when V_{GS} is large and the device is fully on, oscillations do not occur. This means that in a switching application, the oscillation will occur on a transient basis during the turn-on and turn-off transitions. In an application where the device is biased on to some fixed point in the linear region, the oscillations can be continuous.

An example of a parasitic oscillation during a turn-on transition is given in Figure 24. Figure 24A shows the normal drain-to-source (V_{DS}) and gate-to-source (V_{GS}) voltage waveforms when no oscillation is present; Figure 24B shows the same waveforms when an oscillation is present. An expanded portion of the oscillation envelope is given in Figure 25. In this case, the oscillation frequency is approximately 85 MHz. Although the oscillation amplitude shown is not very high, voltages of 100V or more are possible at the peak of the envelope.

The existence of parasitic oscillations can have the following consequences:

1. Gate rupture due to overvoltage.
2. Gate rupture due to overheating of the gate structure.
3. Increased power dissipation in the device.
4. Increased voltage and current stress in associated circuit components.
5. HF to VHF electromagnetic interference (EMI).

None of these are normally acceptable, and, in nearly all cases, the parasitic oscillations must be eliminated. Many so called "mysterious" failures in MOSFETs are due to parasitic oscillations.



Non-Oscillating MOSFET
FIGURE 24A

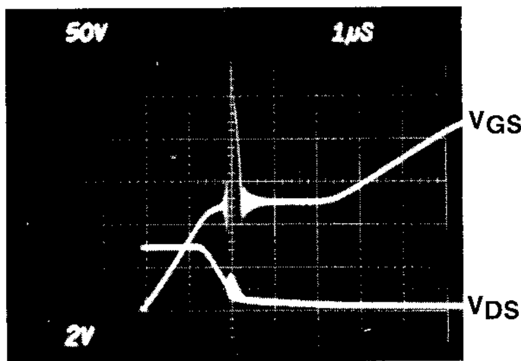
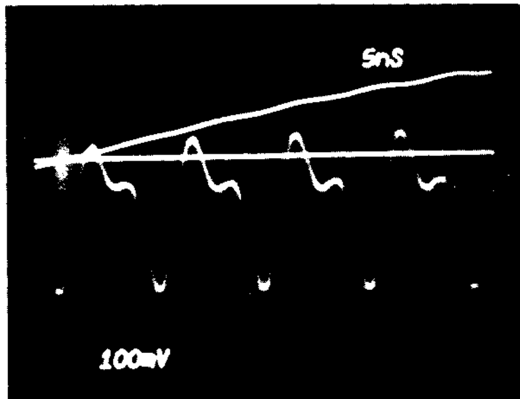


FIGURE 24B



Expanded Time Base Showing Free Running Parasitic Oscillation
FIGURE 25

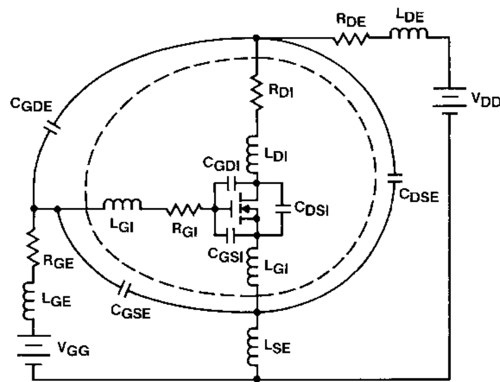
A single device can oscillate by itself. Parallel devices can also oscillate as if they were a single device; this is referred to as a "common mode" oscillation. In addition, paralleled devices can oscillate in a "differential mode." The analysis for each mode is very similar, but the parasitic circuit elements controlling the oscillations are different.

Common Mode Oscillation

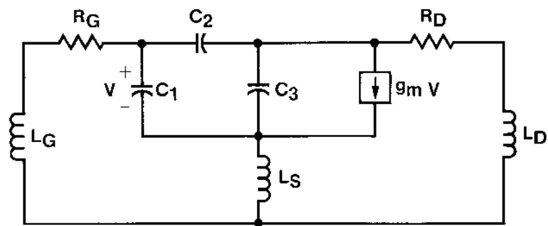
Figure 26 gives a model of the internal and external parasitic capacitances and inductances in a typical application. To analyze this circuit for oscillations, a simplified incremental model (Figure 27) can be used. Figure 27 is not exact but will still give useful answers. The following assumptions have been made and are normally valid:

1. $L_{Ge} >> L_{Gi}$
2. $L_{De} >> L_{Di}$
3. $L_{Se} >> L_{Si}$
4. $R_{Ge} >> R_{Gi}$
5. $C_1 = C_{gse} + C_{gsi}$
6. $C_2 = C_{gde} + C_{gdi}$
7. $C_3 = C_{dse} + C_{dsi}$

A more complex model could, of course, be used at the cost of greatly increased computational difficulty.



Equivalent circuit
FIGURE 26



Incremental Model
FIGURE 27

The characteristic equation for the incremental model is (7):

$$a_1S^4 + a_2S^3 + a_3S^2 + a_4S + 1 = 0 \quad (13)$$

Where:

$$a_1 = C_e^2L_e^2 \quad (14)$$

$$a_2 = C_e^2[R_G(L_D + L_S) + R_D(L_G + L_S)] + g_m L_e^2C_2 \quad (15)$$

$$a3 = R_DR_GC_e^2 + g_m C_2[R_G(L_D + L_S) + R_D(L_G + L_S)] + L_G(C_1 + C_2) + L_D(C_2 + C_3) + L(C_1 = C_3) \quad (16)$$

$$a_4 = g_mR_DR_GC_2 + R_G(C_1 + C_2) + R_D(C_2 + C_3) + g_mL_S \quad (17)$$

$$L_e^2 = L_DL_G + L_DL_S + L_GL_S \quad (18)$$

$$C_e^2 = C_1C_2 + C_1C_3 + C_2C_3 \quad (19)$$

g_m = transconductance of the MOSFET

Note that equation (13) is fourth order even though there are six reactive elements in the model. The reason for this is that C_1, C_2, and C_3 form a ring of capacitors which means that the voltages on any two determine the voltage across the third. This reduces the equation from sixth to fifth order. L_D, L_G, and L_S form a cut set of inductors where the current in any two inductors determines the current in the third. This further reduces the order from fifth to fourth.

In some applications, the model may be further simplified if either L_D or L_S is very small. The coefficients of the characteristic equations for these cases are:

$L_D = 0$

$$L_e^2 = L_GL_S \quad (20)$$

$$a_1 = L_GL_SC_e^2 \quad (21)$$

$$a_2 = C_e^2(L_SR_G + L_GR_D + L_SR_D) + g_m L_GL_SC_2 \quad (22)$$

$$a3 = R_DR_GC_e^2 + g_mC_2(L_SR_G + L_GR_D + L_SR_D) + L_G(C_1 + C_2) + L_S(C_1 + C_3) \quad (23)$$

$$a_4 = g_mR_DR_GC_2 + R_G(C_1 + C_2) + R_D(C_2 + C_3) + g_mL_S \quad (24)$$

$L_S = 0$

$$L_e^2 = L_DL_G \quad (25)$$

$$a_1 = L_DL_GC_e^2 \quad (26)$$

$$a_2 = C_e^2(L_DR_G + L_GR_D) + g_m L_DL_GC_2 \quad (27)$$

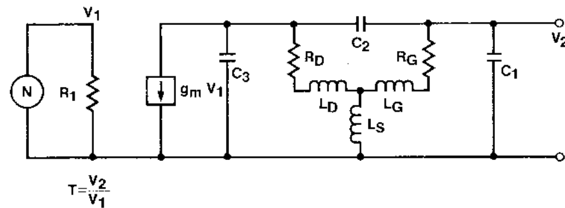
$$a_3 = R_DR_GC_e^2 + g_m(L_DR_G + L_GR_D)C_2 + L_G(C_1 + C_2) + L_D(C_2 + C_3) \quad (28)$$

$$a_4 = g_mR_DR_GC_2 + R_G(C_1 + C_2) + R_D(C_2 + C_3) \quad (29)$$

The practical question which must be answered is: "Does the circuit oscillate or not?" The characteristic equation can answer this question. If any of the roots of the polynomial are negative (i.e. lie in the right half-plane or have the term s-α) then the circuit will oscillate. There are several ways this can be determined.

1. Examine the polynomial for negative coefficients. In this particular case, all of the known quantities have positive signs, and therefore, all of the coefficients will be positive.
2. Use a calculator or computer, with a root finding program, to see if any negative roots exist. For example, the HP41-CV with a math pack can find the roots for up to fifth order polynomial and would be very useful.
3. The Routh-Hurwitz procedure (7) can be used to determine if negative roots exist. This procedure can conveniently be done on a programmable calculator.

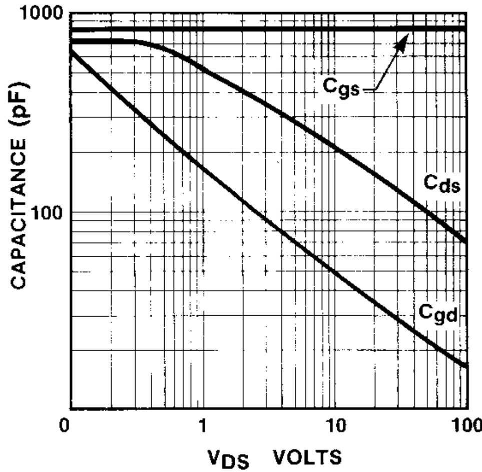
Giadomenico (8) has suggested an alternate means for determining stability by treating the incremental model as a feedback system and then modeling the circuit using SPICE. If the equivalent open loop gain and phase shift at 0° and a gain of 1 coincide, the circuit will oscillate at that frequency. The SPICE model suggested by Giadomenico is shown in Figure 28. This approach would be particularly useful if more complex models were used such as the complete high frequency RF model (9).



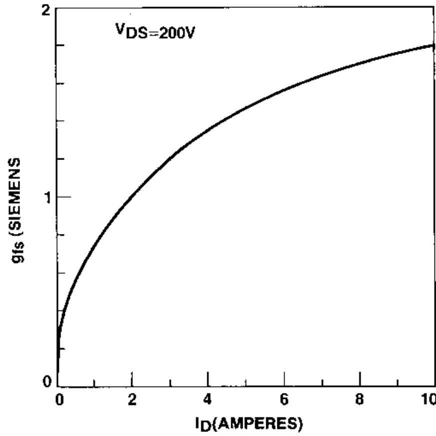
SPICE Model for Stability Analysis
FIGURE 28

Whatever approach is used, appropriate values for the model elements must be known. The values for the internal and external parasitic inductances, the external parasitic capacitances, and the parasitic resistances can be measured or estimated from simple calculations (10). These elements do not change value during the operating cycle, but g_m, C_{gd}, and C_{ds} do vary.

Typical variations for intra-terminal capacitances are shown in Figure 29. Clearly C_{ds} and C_{dg} will change dramatically as the device is switched. The variations in g_m during switching transitions are shown in Figures 30 and 31. g_m is not very sensitive to V_{DS} except for low values, but it does vary as I_D changes. From these graphs, for the particular device being used, the simultaneous values for C_{gd}, C_{ds}, and g_m can be determined from the I_D/V_{DS} load line. The stability analysis can then be performed at several points on the load line to determine if oscillations are possible.

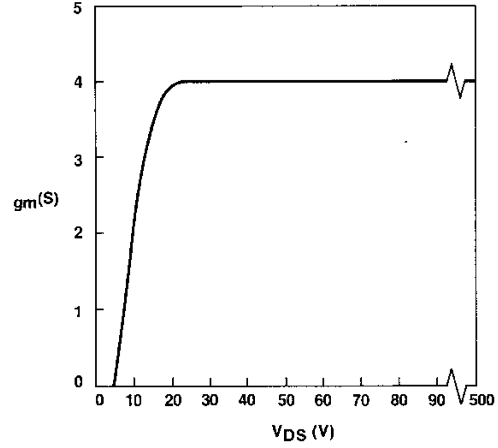


Intraterminal Capacitance Variation
FIGURE 29



Transconductance as a Function of I_D
FIGURE 30

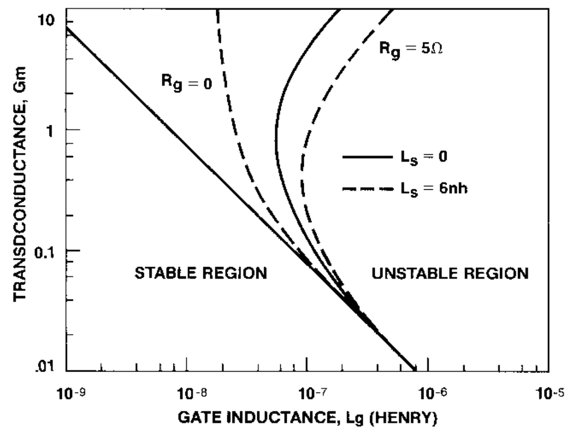
This analysis procedure is relatively complex, and unless the whole business is pre-programmed into a computer, it is unlikely that most designers would use this for solving day to day design problems. The equations can be solved for a few examples, however, and from these examples, general trends can be identified. From these, appropriate means for suppressing oscillations and estimates of the proper values for damping elements can be found.



Transconductance Variation as a Function V_{DS}
FIGURE 31

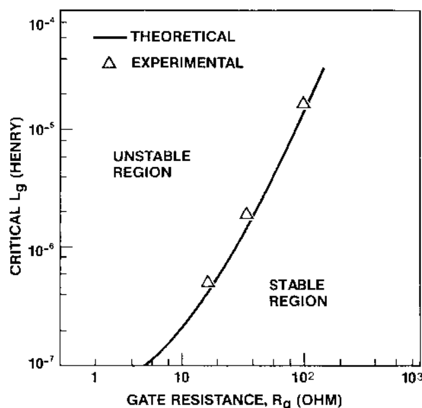
A typical example of the effect of g_m, L_G, L_S, and R_G on stability is shown in Figure 32. Examining this graph, several trends can be clearly seen:

1. The larger the value of g_m, the smaller the region of stability.
2. The smaller the value of L_G, the larger the region of stability.
3. Even a very small value for L_S will greatly improve the region of stability.
4. Small increases in R_G greatly increase the region of stability.
5. As the I_D rating is increased and the BV_{DSS} rating decreased, g_m will increase. This implies that large, low voltage devices or multiple devices in parallel will be more prone to oscillation.

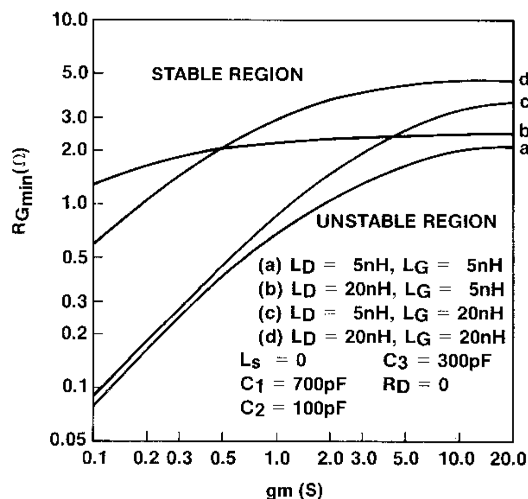


Transconductance as a Function of
Gate Inductance
FIGURE 32

A typical relationship between the maximum value of L_G which provides stability for a given value of R_G is shown in Figure 33. Again the message is clear, minimize L_G and then use a small value of R_G to stabilize.



Critical Gate Inductance as a Function of Gate Resistance
FIGURE 33



Minimum Value of Gate Resistance for Stability
FIGURE 34

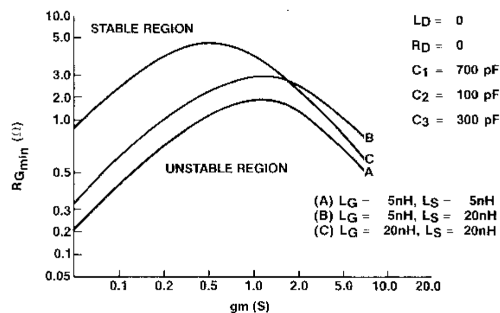
Differential Mode Oscillations

Oxner (11) has demonstrated that two or more parallel MOSFETs can oscillate in a differential or "push-pull" mode. Kassakian (4) has shown how to analyze this case. The following discussion uses his method.

When two devices are paralleled, the equivalent circuit, including parasitics, becomes double that shown in Figure 26. The circuit is quite complex and more than a little intimidating. The circuit can, however, be greatly simplified by recognizing that in the *differential* mode, the circuit can be simplified by assuming that all of the nodes on the plane of symmetry are at incremental ground. This means that all elements common to *both* devices can be eliminated. In addition, the method of half-circuits can be used to further reduce the equivalent circuit complexity. The result of these simplifications is topologically identical to Figure 27. This means that we can use equations (13) through (29) without modification. We *cannot*, however, make the same assumptions regarding the relative values of the parasitic elements. The following assumptions apply:

1. All elements *common* to both devices are ignored.
2. The internal inductances (L_{Gi}, L_{Di}, L_{Si}) cannot be ignored. In fact, in devices using multiple internal chips, the only differential inductances present are the internal bond wires!

With representative values for the circuit elements, we can again generate some informative examples as shown in Figures 34 and 35.



The Effect of Changing L_S and L_G on $R_{G(min)}$
FIGURE 35

When $L_S = 0$ (Figure 34), the following points can be made:

1. The higher the value for g_m, the larger R_G must be.
2. The larger L_D, the larger R_G required.
3. Increasing L_G increases R_G, but this is not a strong function.

When $L_D = 0$ (Figure 35), which is more often the case than $L_S = 0$, the relationships change a bit:

1. R_G no longer increases monotonically with g_m. This is in agreement with Figure 32.
2. For values of g_m above 1.5S, increasing L_G actually improves stability!
3. Typical values for R_G needed to assume stability are small (in the range of 0.5 to 5 ohms) even in large devices (high g_m).

How to Recognize Oscillation Problems

It is important that the designer recognize when oscillation problems might appear. Normally the first indication of this problem is the failure of devices with gate-to-source shorts or with high leakage (I_{GSS} > 100nA). The best means for detecting oscillations is to place a scope probe between the gate and source terminals, directly on the device. The scope should have a bandwidth of at least 200 MHz since it is very difficult to detect VHF oscillations with an instrument that responds to only a few MHz. A low capacitance probe should be used, and as little stray inductance as possible should be introduced. It is possible for the measurement process itself to alter the circuit operation and to suppress or induce the oscillation!

Prevention of Oscillation in MOSFETs

From the theoretical and experimental work on this problem, it is clear that preventing parasitic oscillations is not a major problem and can be accomplished by observing the following guidelines:

1. Minimize the parasitic inductances and capacitances. In particular L_G, L_S, and C_{gd} should be made as small as possible. Making the parasitic elements smaller raises the resonant frequency. As the frequency is increased, the gain of the device will decrease, and the resistive damping present will become more effective. The net result is a reduced likelihood of oscillation.
2. Use small (1-5 ohm) differential resistors in the gate lead of each device. Because of the silicon gate structure of Siliconix devices, most of the needed resistance will already be present. R_G should, of course, be non-inductive. Carbon composition resistors are particularly good.
3. Any resonant circuit has a non-zero value of Q. The higher Q is, the slower the oscillation will build up. The time constant (τ) will be:

$$\tau = \frac{2Q}{\omega} \qquad (30)$$

   If the switch transition time is short compared to τ, then the oscillation will not appear even though the circuit is potentially unstable.
4. Minimize the differential values of L_S and L_D.
5. For the differential mode of oscillation, ferrite beads provide both R_G and increased L_G and can be very effective in suppressing oscillations.

CONCLUSIONS

In this discussion, a wide variety of possible problems has been considered. In each case, the potential problems have been shown to be either no problem at all or curable through some simple circuit means. In particular, it has been shown that matching of devices is rarely needed. From this discussion, it is quite clear that paralleling MOSFETs is relatively easy if a few simple rules are followed:

1. Provide good thermal coupling between devices.
2. Use good circuit layout practices.
3. Use small series gate resistors to suppress oscillation.
4. Other than the gate resistors needed for oscillation suppression, minimize the differential gate impedances.
5. In linear applications, use small differential source resistors to stabilize the operating point and force current sharing.
6. Carefully examine the circuit waveforms for any signs of parasitic oscillation or current spiking.

For those readers desiring more information, the articles referenced in the bibliography should prove useful.

BIBLIOGRAPHY

(1) R. Severns, "Safe Operating Area and Thermal Design for MOSPOWER Transistors," Siliconix Applications Note AN83-10, November 1983.

(2) W. Schultz, "High Current FETs, A New Level of Performance," Power Conversion International, March 1984, pp. 43-46.

(3) J. Kassakian, "Some Issues Related to the Behavior of Multiple Paralleled Power MOSFETs,"

(4) J. Kassakian and D. Lau, "An Analysis and Experimental Verification of Parasitic Oscillation in Paralleled Power MOSFETs," IEEE. Tran. Electron Devices, Vol. ED-31, NO7, July 1984 pp. 959-963.

(5) K. Gauen, "Match Power — MOSFET Parameters for Optimum Parallel Operation," EDN, February 23, 1984, pp. 249-267.

(6) K. Gauen, "Paralleling Power MOSFETs In Linear Applications," Proceedings of PCI 1984, April 1984, Atlantic City, pp. 144-151.

(7) R. Saucedo and E. Schiring, "Introduction to Continuous and Digital Control Systems," MacMillan, 1968, p. 96.

(8) D. Giadomenico, "Anomalous Oscillations and Turn-off Behavior In a Vertical Power MOSFET," University of California Research Project, December 1983; Later published in the proceedings of POWERCON 11, April 1984.

(9) E. Oxner, "Meet the Power-MOSFET Model," Siliconix Applications Note AN80-2, 1980.

(10) F. Terman, "Radio Engineers Handbook," McGraw-Hill, 1943, pp. 47-64.

(11) E. Oxner, "Controlling Oscillation in Parallel Power MOSFETs," Siliconix TA83-2, September 1983; Also published in the proceedings of POWERCON 10, 1983.

Siliconix 2201 Laurelwood Road, Santa Clara, CA 95054 (408) 988-8000 TWX: 910-338-0227
The information contained in this data sheet is current at the time of publication. Siliconix incorporated reserves the right to make changes in the circuitry or specifications at any time without notice, henceforth, assumes no responsibility for the use of any circuits described herein and make no representations that they are free from patent infringement.
Readers are cautioned to verify currency of information in this publication prior to placing orders. Printed In U.S.A.