

A USER'S GUIDE TO COMPENSATING LOW-DROPOUT REGULATORS

The explosive proliferation of battery powered equipment in the past decade has created unique requirements for a voltage regulator that can not be met by the industry standards like the LM340 or LM317. These regulators use an NPN Darlington pass transistor (Figure 1), and are often referred to as NPN regulators. The demand for higher performance is being met by the newer low-dropout (LDO) regulators.

THE NPN REGULATOR

The NPN Darlington pass transistor configuration requires that at least 1.5V to 2.5V be maintained from input-to-output for the device to stay in regulation. This minimum voltage "headroom" (called the dropout voltage) is given by:

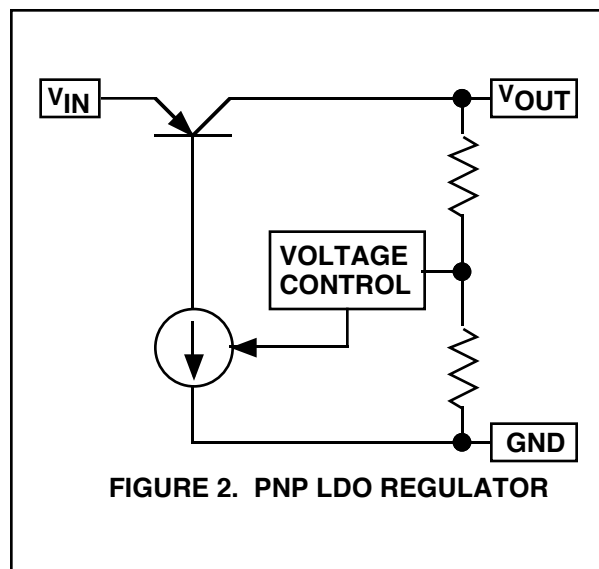
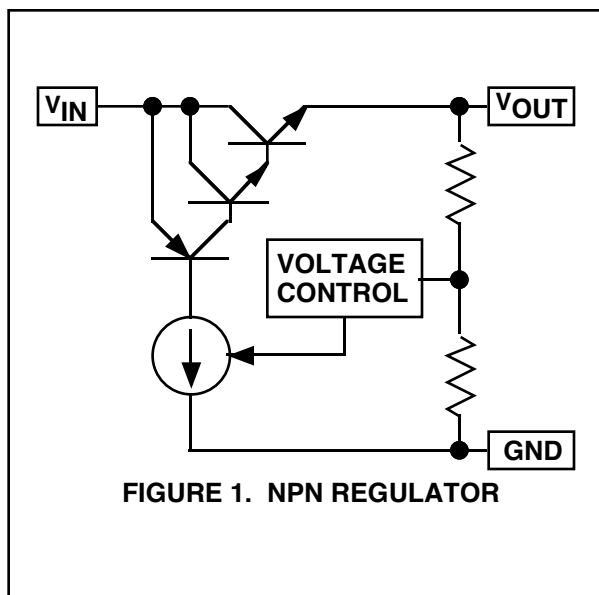
$$V_{\text{DROP}} = 2V_{\text{BE}} + V_{\text{SAT}} \quad (\text{NPN REG})$$

THE LDO REGULATOR

The low-dropout (LDO) regulator operates exactly the same as the NPN with the exception that the NPN Darlington pass transistor has been replaced by a single PNP transistor (Figure 2). The big advantage of the LDO is that the PNP pass transistor can maintain output regulation with very little voltage drop across it:

$$V_{\text{DROP}} = V_{\text{SAT}} \quad (\text{LDO REGULATOR})$$

Full-load dropout voltages < 500 mV are typical. At light loads, dropout voltages can fall as low as 10 to 20 mV.



NPN vs. LDO

The key difference between the NPN and LDO regulators is that the dropout voltage of the LDO is much lower than the NPN, which is most valuable in battery powered applications where the regulated output can be maintained farther down the battery discharge curve.

An advantage of NPN regulators is that they are unconditionally stable (they require no external capacitors). An LDO does require at least one external capacitor on the output to reduce the loop bandwidth and provide some positive phase shift (see later sections for stability analysis).

The quiescent (ground pin) current of an LDO is typically higher than an NPN, being equal to the load current divided by the beta (current gain) of the PNP pass transistor. At full load current, PNP beta values of 15 - 20 are not unusual, which means the LDO ground pin current can be as high as 7% of the load current.

FEEDBACK AND LOOP STABILITY

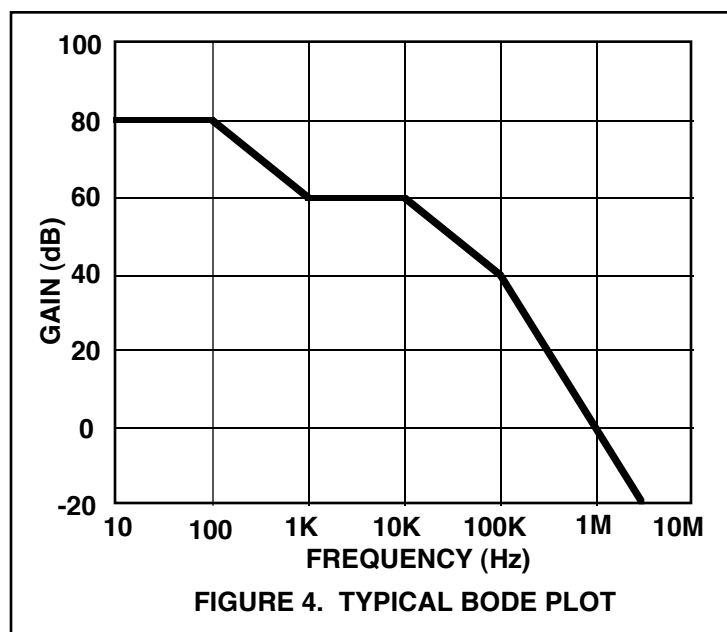
All voltage regulators use a feedback loop to hold the output voltage constant. The feedback signal experiences changes in both gain and phase as it goes through the loop, and the amount of phase shift which has occurred at the unity gain (0 dB) frequency determines stability.

BODE PLOTS

Understanding stability requires the use of Bode Plots, which show the loop gain (in dB) plotted as a function of frequency (Figure 4). Loop gain and associated terms are defined in the next sections.

Loop gain is typically measured on a network analyzer, which injects a low-level sine wave into the feedback path and then measures the gain response while the frequency of the signal is swept from DC up to the frequency where the gain drops well below 0 dB.

Bode plots are convenient tools because they contain all of the information necessary to determine if a closed-loop system is stable. However, decoding the information contained in a Bode plot requires understanding the key elements: loop gain, phase margin, poles and zeroes.



LOOP GAIN

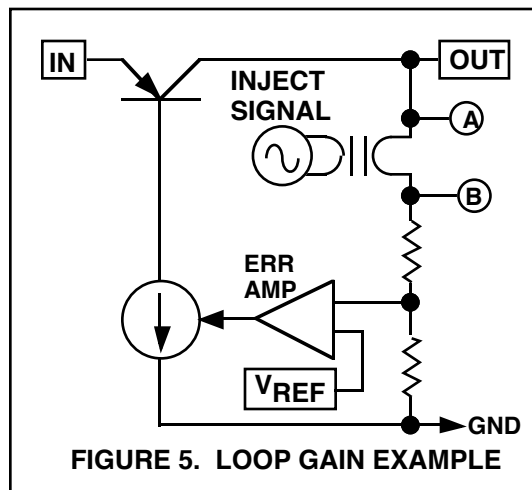
Every closed-loop system has a characteristic called loop gain. In this analysis of voltage regulators, loop gain will be defined as the magnitude of the voltage gain that the feedback signal experiences as it travels through the loop. The block diagram of the LDO in Figure 2 will be re-drawn to illustrate this concept (Figure 5).

A transformer is used to inject an AC signal into the feedback path between points “A” and “B”. Using this transformer, a small-signal sine wave is used to “modulate” the feedback signal. The AC voltages at “A” and “B” are measured and used to calculate loop gain.

The the loop gain is defined as the ratio of the two voltages:

$$\text{Loop Gain} = V_A / V_B$$

It is important to note that the signal starting at the V_B point has a phase shift introduced into it as it travels through the loop (eventually arriving at the V_A point). The amount of phase shift is critical in determining stability.

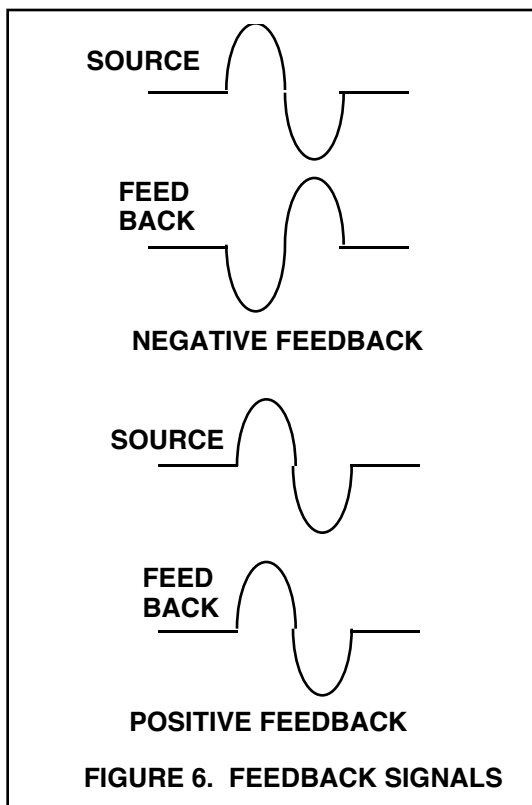


FEEDBACK

Feedback is used in all voltage regulators to hold the output voltage constant. The output voltage is sampled through a resistive divider (Figure 5), and that signal is fed back to one input of the error amplifier. Since the other input of the error amplifier is tied to a reference voltage, the error amplifier will supply current as required to the pass transistor to keep the regulated output at the correct DC voltage.

It is important to note that for a stable loop, negative feedback must be used. Negative feedback (sometimes called degenerative feedback) is opposite in polarity to the source signal (see Figure 6).

Because it is opposite in polarity with the source, negative feedback will always cause a response by the loop which opposes any change at the output. This means that if the output voltage tries to rise (or fall), the loop will respond to force it back to the nominal value.



Positive feedback occurs when the feedback signal has the same polarity as the source signal. In this case, the loop responds in the same direction as any change which occurs at the output. This is clearly unstable, since it does not cancel out changes in output voltage, but amplifies them.

It should be obvious that no one would intentionally design positive feedback into the loop of a linear regulator, but negative feedback becomes positive feedback if it experiences a phase shift of 180° .

PHASE SHIFT

Phase shift is defined as the total amount of phase change (referred to the starting point) that is introduced into the feedback signal as it goes around the loop.

Phase shift (expressed in degrees) is most often measured using a network analyzer.

Ideal negative feedback is 180° out of phase with the source (Figure 7), so its “starting point” is at -180° . This “ 180° offset” can also be seen in Figure 6, as the negative feedback waveforms are exactly one half cycle shifted with respect to each other.

It can be seen that starting at -180° , an additional phase shift of 180° (positive or negative) brings the signal back to zero, which is the phase of the source signal and would cause the loop to be unstable.

PHASE MARGIN

Phase margin is defined as the difference (in degrees) between the total phase shift of the feedback signal and -180° at the frequency where the loop gain is equal to 0 dB (unity gain). A stable loop typically needs at least 20° of phase margin (more is better).

Phase shift and phase margin can be calculated using the poles and zeroes present in the Bode plot.

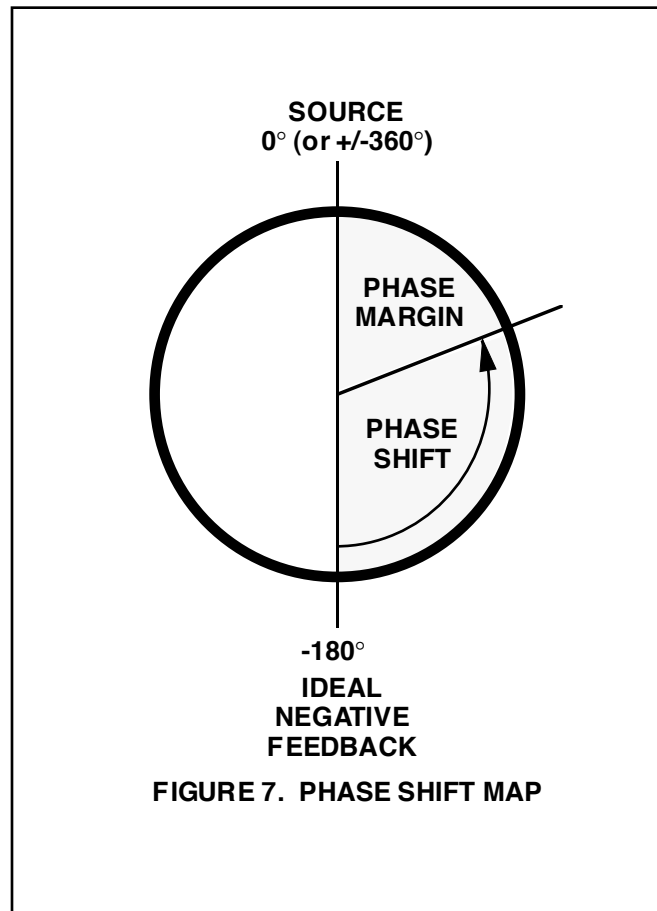


FIGURE 7. PHASE SHIFT MAP

POLES

A pole (Figure 8) is defined as a point where the slope of the gain curve changes by -20 dB/decade (with reference to the slope of the curve prior to the pole). Note that the effect is additive: each additional pole will increase the negative slope by the factor “ n ” X **-20 dB/decade**, where “ n ” is the number of additional poles.

The phase shift introduced by a single pole is frequency dependent, varying from 0 to -90° (with a phase shift of -45° at the pole frequency). The most important point is that nearly all of the phase shift added by a pole (or zero) occurs within the frequency range one decade above and one decade below the pole (or zero) frequency.

NOTE: a single pole can add only -90° of total phase shift, so at least two poles are needed to reach -180° (which is where instability can occur).

ZEROES

A zero (Figure 9) is defined as a point where the gain changes by $+20$ dB/decade (with respect to the slope prior to the zero). As before, the change in slope is additive with additional zeroes.

The phase shift introduced by a zero varies from 0 to $+90^\circ$, with a $+45^\circ$ shift occurring at the frequency of the zero.

The most important thing to observe about a zero is that it is an “anti-pole”, which is to say its effects on gain and phase are exactly the opposite of a pole.

This is why zeroes are intentionally added to the feedback loops of LDO regulators: they can cancel out the effect of one of the poles that would cause instability if left uncompensated.

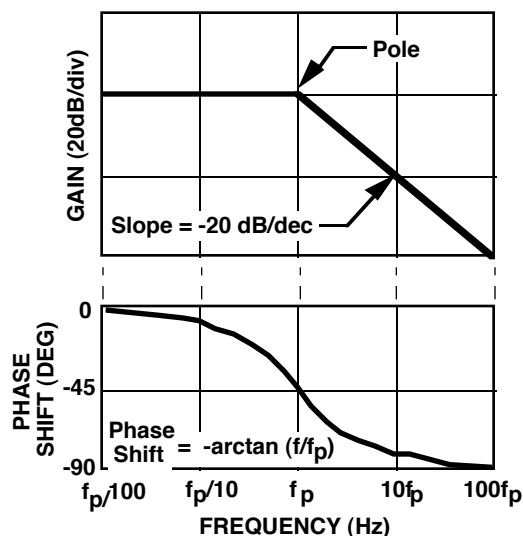


FIGURE 8. POLE GAIN/PHASE PLOT

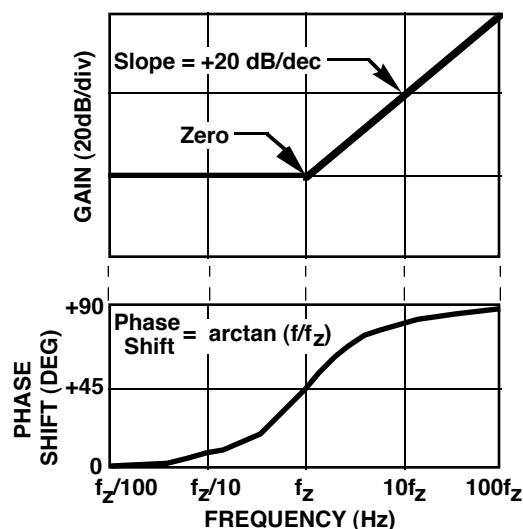


FIGURE 9. ZERO GAIN/PHASE PLOT

BODE PLOT ANALYSIS

A Bode plot which contains three poles and one zero (Figure 10) will be analyzed for gain and phase margin.

The DC gain is assumed to be 80 dB, with the first pole occurring at 100 Hz. At that frequency, the slope of the gain curve changes to -20 dB/decade.

The zero at 1 kHz changes the slope back to 0 dB/decade until the second pole at 10 kHz, where the gain curve slope returns to -20 dB/decade.

The third and final pole at 100 kHz changes the gain slope to the final value of -40 dB/decade.

It can also be seen that the unity-gain (0 dB) crossover frequency is 1 MHz. The 0 dB frequency is sometimes referred to as the loop bandwidth.

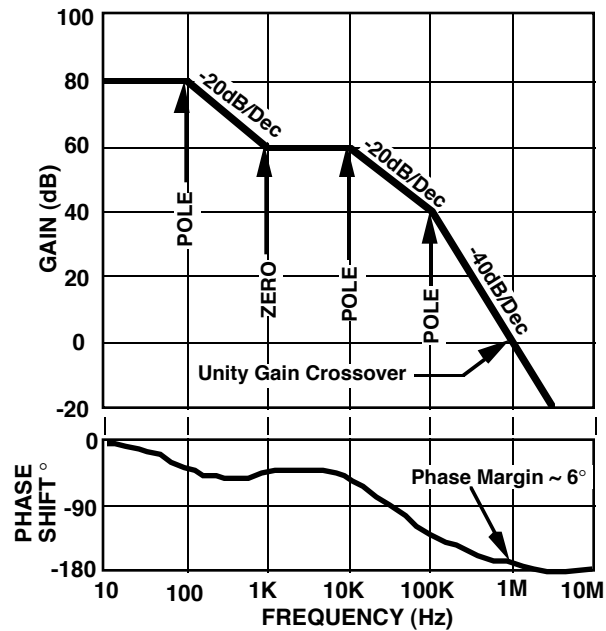


FIGURE 10. BODE PLOT WITH PHASE INFO

The plot of Phase Shift shows how the various poles and zeroes contribute their effect on the feedback signal. To produce this plot, the phase shift at each frequency point was calculated based upon summing the contributions of every pole and zero at that frequency. The phase shift at any frequency “f” which is caused by a pole located at frequency “f_p” can be calculated from:

$$\text{Pole Phase Shift} = -\arctan(f/f_p)$$

The phase shift resulting from a zero located at frequency “f_z” can be found using:

$$\text{Zero Phase Shift} = \arctan(f/f_z)$$

Is this loop stable? To answer that question, we need only know the phase shift at 0 dB (which is 1 MHz in this case). Finding this does not require complex calculations:

As stated in the previous sections, a pole or zero contributes nearly its full phase shift in the frequency range one decade above and below the center frequency of the pole (or zero). Therefore, The first two poles and the first zero contribute their full phase shifts of -180° and +90°, respectively, resulting in a net phase shift of -90°.

The final pole is exactly one decade below the 0 dB frequency. Using the formula for Pole Phase Shift, this pole will contribute -84° of phase shift @ 1 MHz. Added to the -90° from the two previous poles and the zero, the total phase shift is -174° (which means the phase margin is 6°). This loop would either oscillate or ring severely.

NPN REGULATOR COMPENSATION

The pass transistor of the NPN regulator (see Figure 1) is connected in a circuit configuration known as common collector. An important characteristic of all common collector circuits is wide bandwidth, which means the inherent pole that it places in the loop gain occurs at a very high frequency.

A simple compensation technique called dominant pole compensation can be used in the NPN regulator because it has no inherent low-frequency poles.

In this case, a capacitor is built into the IC which places a pole in the loop gain at a low frequency (Figure 11).

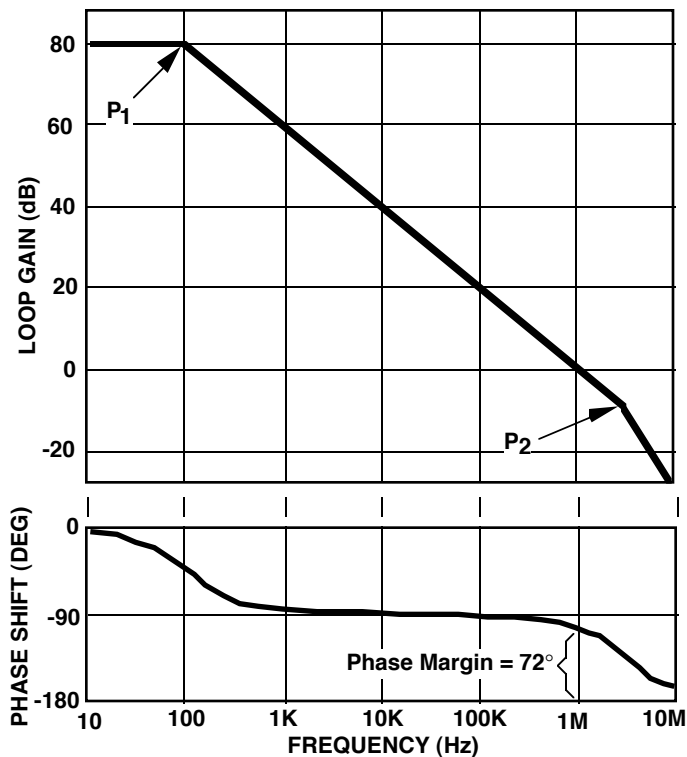


FIGURE 11. BODE PLOT FOR NPN REGULATOR

This dominant pole (shown as P1) for a typical NPN regulator is set at about 100 Hz. The 100 Hz pole causes the gain to decrease at a rate of -20 dB/decade until the second pole (P2) which is located at 3 MHz. At that point, the slope of the gain plot changes to -40 dB/decade.

The frequency of P2 is primarily due to the NPN power transistor and associated drive circuitry, so it is sometimes referred to as the power pole. Since P2 occurs at a frequency where the loop gain is -10dB, its contribution to phase shift at the 0 dB frequency (1 MHz) will be small.

To determine stability, it only requires that the phase margin at 0 dB be calculated:

The first pole (P1) will contribute -90° of phase shift, but the second pole (P2) will add only -18° of negative phase shift @ 1 MHz (0 dB). This means the total phase shift @ 0 dB is -108°, which yields a phase margin of 72° (which is very stable).

It should also be noted that simple observation would clearly show this loop is stable, since reaching -180° of phase shift (the point of instability) would require the full contribution of -90° (each) from both poles, and P2 is too high in frequency to contribute significant phase shift at the 0 dB frequency (1 MHz).

LDO REGULATOR COMPENSATION

The PNP power transistor in an LDO regulator (Figure 2) is connected in a configuration called common emitter, which has a higher output impedance than the common collector configuration in the NPN regulator. This adds an additional low-frequency pole whose frequency is dependent both on load resistance and output capacitance. The frequency of this pole (which will be designated P_L for load pole) is found from:

$$f(P_L) = 1 / (2\pi \times R_{LOAD} \times C_{OUT})$$

The presence of the frequency-variable load pole P_L means that the simple dominant pole compensation method used in the NPN regulator will not work in an LDO unless additional compensation is added. To illustrate why this is true, the loop gain of a 5V/50 mA LDO regulator will be illustrated using these assumptions:

1) At maximum load current, the load pole (P_L) occurs at a frequency given by:

$$P_L = 1 / (2\pi \times R_{LOAD} \times C_{OUT}) = 1 / (2\pi \times 100 \times 10^{-5}) = 160 \text{ Hz}$$

2) The internal compensation will be assumed to add a fixed pole (P_1) at 1 kHz.

3) A 500 kHz power pole (which will be designated P_{PWR}) is present due to the PNP power transistor and driver.

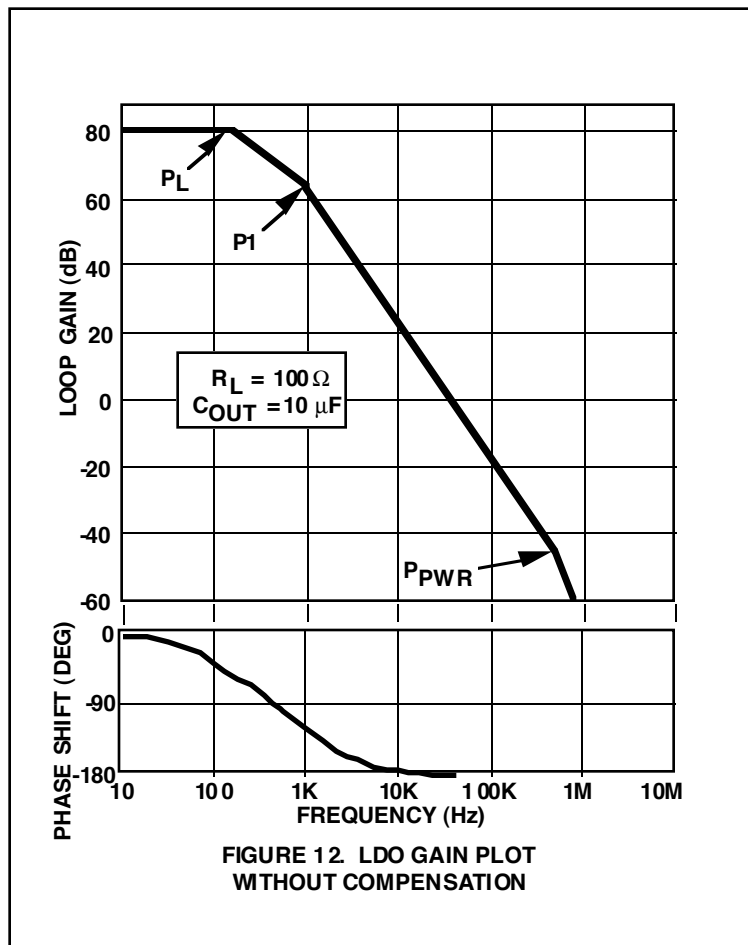
4) The DC gain is assumed to be 80 dB.

5) $R_L = 100\Omega$ (which is the value at maximum load current).

6) $C_{OUT} = 10 \mu\text{F}$

Using the conditions stated above, a Bode plot (Figure 12) is drawn. It is immediately obvious that this loop is not stable: the two poles P_L and P_1 will each contribute -90° of phase shift, causing the total phase shift to reach -180° at the 0 dB frequency (which is about 40 kHz in this example).

To reduce the negative phase shift (and prevent oscillations), a zero must be added to the loop. A zero can contribute as much as $+90^\circ$ of positive phase shift, which will cancel out the effects of one of the two low frequency poles.



All monolithic LDO regulators require that this zero be added to the loop, and they derive it from a characteristic that is inherent in the output capacitor: equivalent series resistance (usually referred to as ESR).

LDO COMPENSATION USING ESR

Equivalent series resistance (ESR) is a characteristic that is present in every capacitor. It can be modeled electrically as a resistance that is placed in series with the capacitor (Figure 13).

The ESR of the output capacitor puts a zero in the loop gain which can be used to reduce excess negative phase shift.

The frequency where the zero occurs is directly related to the value of the ESR and amount of output capacitance:

$$F_{ZERO} = 1 / (2\pi \times C_{OUT} \times ESR)$$

Using the example in the previous section (Bode plot shown in Figure 12), we will assume that the value of $C_{OUT} = 10 \mu F$ and the output capacitor ESR = 1Ω , which means a zero will occur at 16 kHz.

Figure 14 shows how this added zero will change the unstable plot into a stable one:

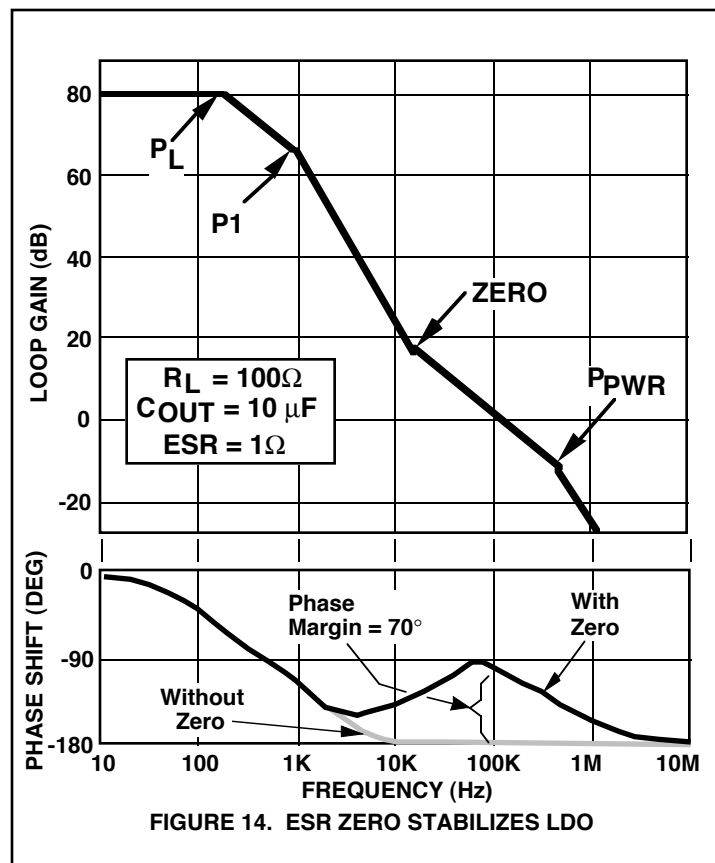
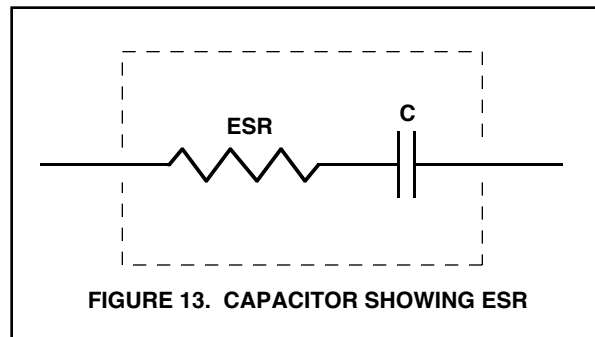
The bandwidth of the loop is increased so that the 0 dB crossover frequency moves from 30 kHz to 100 kHz.

The zero adds a total of $+81^\circ$ positive phase shift at 100 kHz (the 0 dB frequency). This will reduce the negative phase shift caused by the poles P_L and P_1 .

Since the pole P_{PWR} is located at 500 kHz, it adds only -11° of phase shift at 100 kHz.

Summing all poles and zeroes, the total phase shift at 0 dB is now -110° . This corresponds to a phase margin of $+70^\circ$, which is extremely stable.

This illustrates how an output capacitor with the correct value of ESR can generate a zero that stabilizes an LDO.



ESR AND STABILITY

Virtually all LDO regulators require that the ESR of the output capacitor be within a set range to assure regulator stability.

The LDO manufacturer provides a set of curves which define the boundaries of the stable region, plotted as a function of load current (Figure 15).

To explain why these boundaries exist, the effects of low and high ESR on phase margin will be illustrated using the example previously developed.

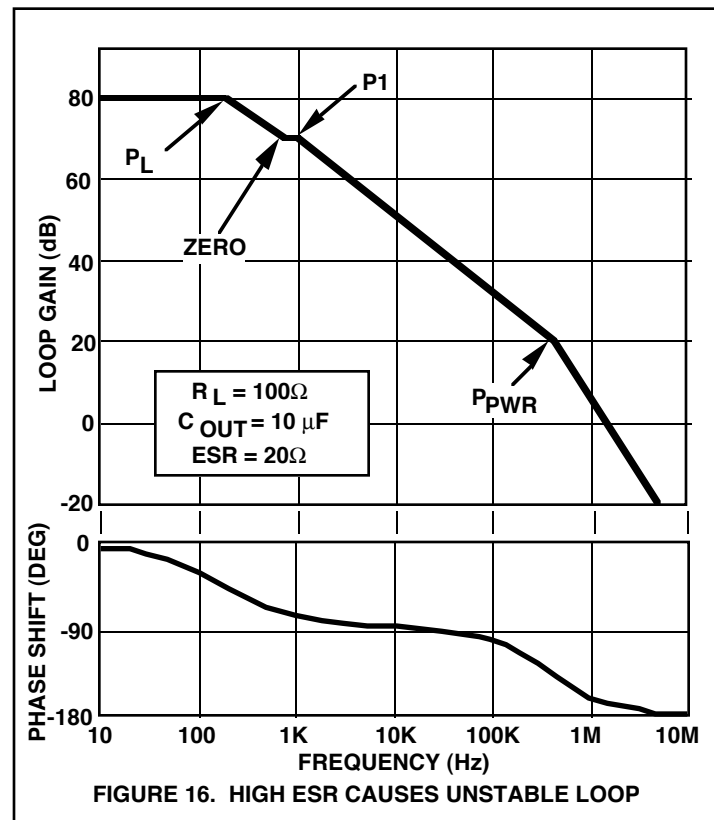
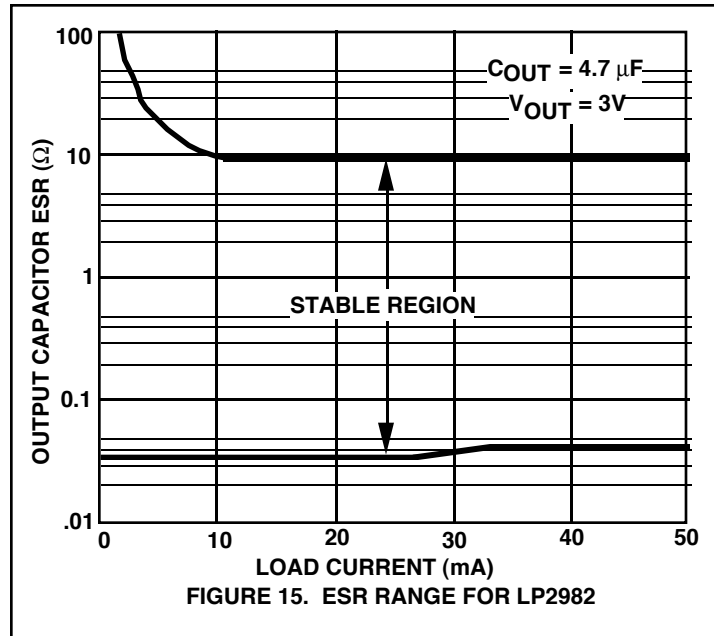
HIGH ESR

Using the example developed in the previous sections, we will change the conditions and assume the ESR of the $10\ \mu\text{F}$ output capacitor is increased to $20\ \Omega$. This will decrease the frequency of the zero to $800\ \text{Hz}$ (Figure 16). Reducing the frequency of the zero causes the loop bandwidth to increase, moving the 0dB crossover frequency from $100\ \text{kHz}$ to $2\ \text{MHz}$.

This increased bandwidth means that the pole P_{PWR} occurs at a gain value of $+20\ \text{dB}$ (compared to $-10\ \text{dB}$ in Figure 14).

Analyzing the plot (Figure 16) for phase margin, it can be assumed that the zero cancels out either P_1 or P_L . This means the loop has a two-pole response with the low frequency pole contributing -90° of phase shift and the high frequency pole P_{PWR} contributing about -76° of phase shift.

Although this appears to leave a phase margin of 14° (which might be stable), bench test data shows that ESR values $> 10\ \Omega$ usually cause instability because of phase shifts contributed by other high-frequency poles which are not shown in this simplified model.



LOW ESR

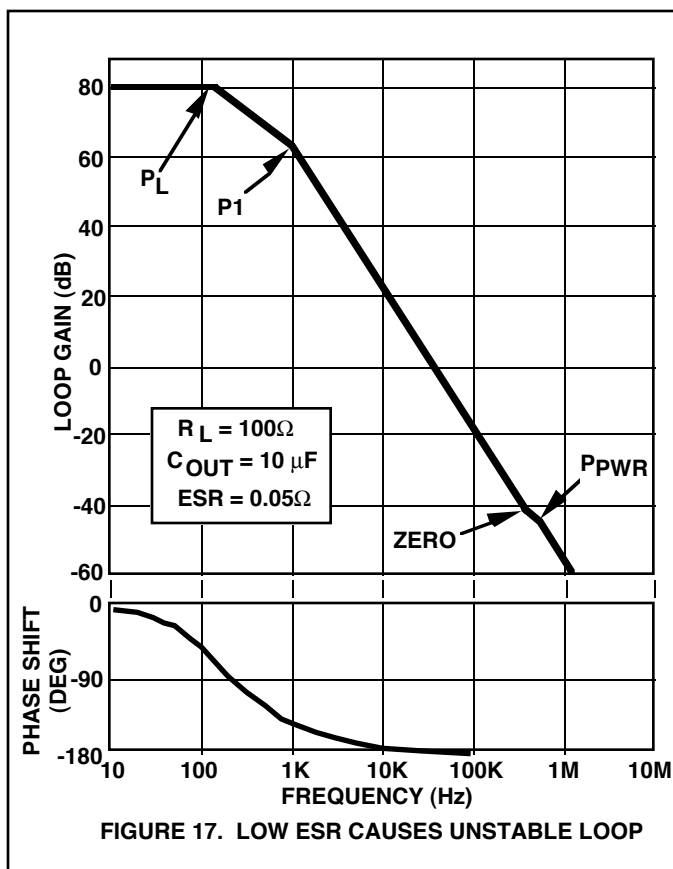
An output capacitor with a very low ESR value can cause oscillations for a different reason.

Continuing the example developed in the previous section, we will now reduce the ESR of the 10 μF output capacitor to 50 m Ω , increasing the frequency of the zero to 320 kHz (Figure 17).

When the plot is analyzed for phase margin, no calculations are required to see that it is unstable.

The -90° phase shift from each of the two poles P_1 and P_L will produce a total phase shift of -180° at the 0 dB frequency.

For this system to be stable, a zero is needed that would provide positive phase shift before the 0 dB point. However, since the zero is at 320 kHz, it's too far out to do any good (and is cancelled out by P_{PWR}).



OUTPUT CAPACITOR SELECTION

Since the output capacitor is the user's tool for compensating a monolithic LDO regulator, it must be selected very carefully. Most cases of oscillations in LDO applications are caused by the ESR of the output capacitor being too high or too low.

When selecting an output capacitor for an LDO, a solid Tantalum capacitor is always the best choice. Tests performed on an AVX 4.7 μF Tantalum showed an ESR of 1.3 Ω @ 25°C, a value that is almost perfectly centered in the stable region (Figure 15).

Also very important, the ESR of the AVX capacitor varied less than 2:1 over the temperature range of -40°C to $+125^\circ\text{C}$. Aluminum electrolytic capacitors are notorious for exhibiting an exponential increase in ESR at cold temperatures, and are not suitable for use as an LDO output capacitor.

It must be noted that large ($\geq 1\mu\text{F}$) ceramic capacitors typically have very low ESR values ($< 20\text{ m}\Omega$), and will cause most LDO regulators to oscillate if connected directly to the output. A ceramic capacitor can be used if some external resistance is added in series with it to increase the effective ESR. Large value ceramics also have a poor tempco (typically Z5U) which means the capacitance will drop in half as the temperature is increased or decreased to the operating limits.

P-FET LDO REGULATORS

THE FET ADVANTAGE

An LDO regulator can be built using a P-FET as the power transistor (see Figure 18).

To see why using a P-FET in an LDO would be advantageous, it should be noted that all of the base current required by the power transistor in a PNP LDO (Figure 2) flows out of the ground pin and back to the negative input voltage return. Therefore, this base drive current is drawn from the input supply but does not drive the load, so it generates wasted power that must be dissipated within the LDO regulator:

$$\text{PWR (Base Drive)} = V_{IN} \times I_{BASE}$$

The amount of base current required to drive the PNP is equal to the load current divided by the beta (gain) of the PNP, and beta may be as low as 15 - 20 (at rated load current) in some PNP LDO regulators. The wasted power generated by this base drive current is very undesirable (especially in battery-powered applications). Using a P-FET solves this problem, since the Gate drive current is very small.

Another advantage of the P-FET LDO is that the dropout voltage can be made very small by adjusting the ON-resistance of the FET. In monolithic P-FET devices, the ON-resistance is adjusted by increasing (or decreasing) the die area allotted to the FET. In discrete P-FET LDO designs, you just select the FET with the desired ON-resistance for your application.

In summary, the P-FET LDO is most attractive in high-current applications, where the saturation voltage of the PNP (and the base current required to drive it) make the PNP LDO an impractical choice.

THE FET DIS-ADVANTAGE

Although the P-FET LDO may seem like a design breakthrough (compared to a PNP LDO regulator), it has inherent drawbacks when integrated into IC form. Only a small percentage of monolithic LDO products being sold in the marketplace are P-FET designs, and some of them have had limited commercial and technological success.

The CMOS process used to fabricate a P-FET LDO will typically yield worse performance for output noise and bandwidth (loop frequency response). Stability and compensation can also be troublesome in a monolithic P-FET LDO, since the gate capacitance of the FET adds another pole to the loop. This may result in strangely shaped ESR curves unmatched by the characteristics of any real-world capacitor.

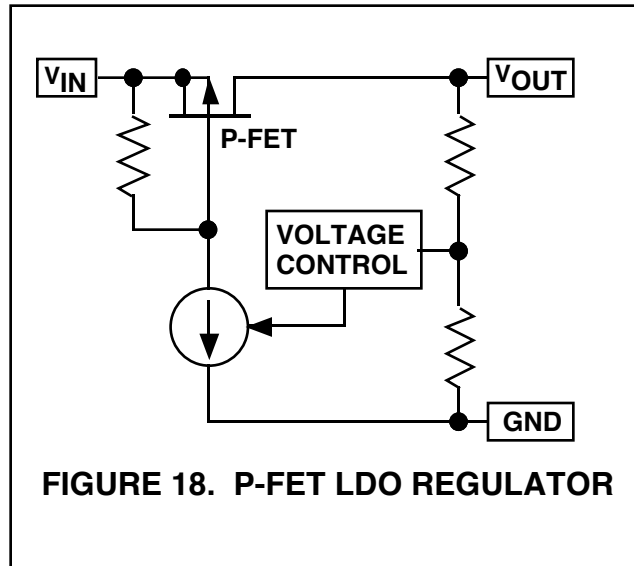


FIGURE 18. P-FET LDO REGULATOR

P-FET CONTROLLER IC

The best way to bypass the inherent drawbacks of a monolithic P-FET IC is to use a controller IC (fabricated on a bipolar process) driving an external discrete P-FET. In higher-current LDO applications, this is superior to a fully-integrated approach for these reasons:

Flexibility: the designer can select any P-FET which yields the optimum dropout voltage for the load current in his application.

Cost savings: Higher current ($> 1\text{A}$) FET's are typically much cheaper in discrete form than when they are integrated into the IC, thus lowering overall design cost.

To address this need, National Semiconductor has just released the **LP2975**: packaged in a mini SO-8, this bipolar controller IC makes it easy to build a P-FET LDO with adjustable current limiting and logic-level ON/OFF.

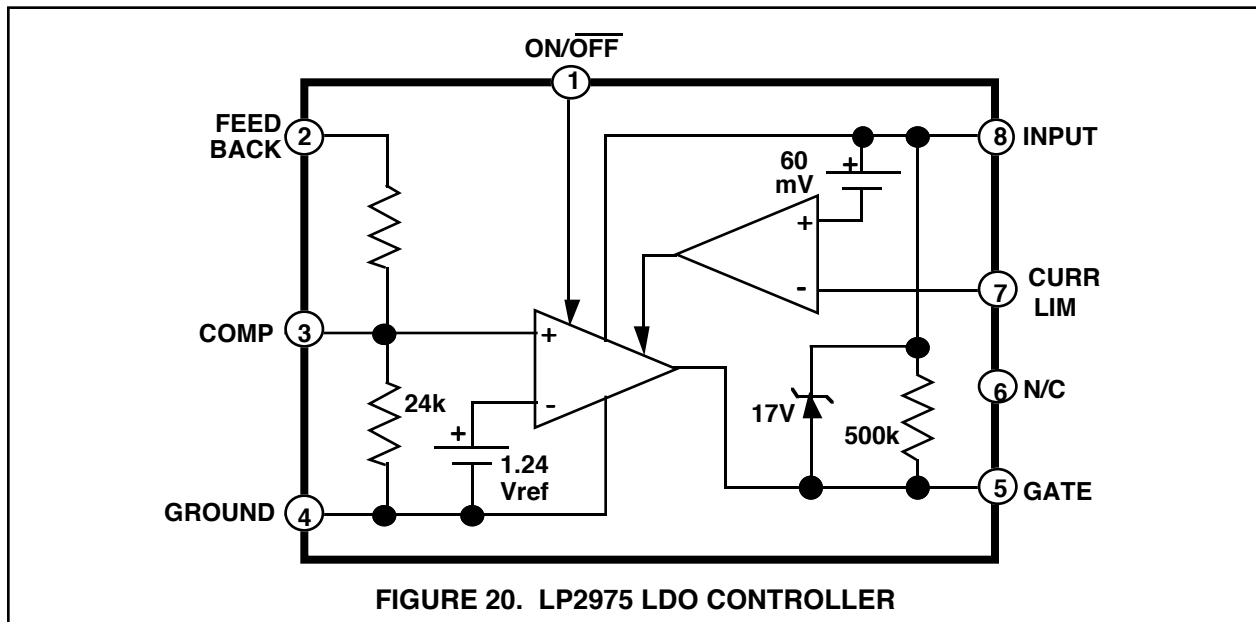
12V, 5V, and 3.3V versions are available as standard products. Output voltage values other than the 12V, 5V, and 3.3V standard values can be obtained by paralleling one or both of the resistors in the internal divider (see Figure 22).

The LP2975 uses an on-board resistive divider (factory trimmed) to set the output voltage. An internal 1.24V bandgap reference is connected to the error amplifier, whose output adjusts the voltage at the gate of the P-FET.

Current limiting is easy to program with just a single external resistor, which may be omitted for lowest possible dropout voltage.

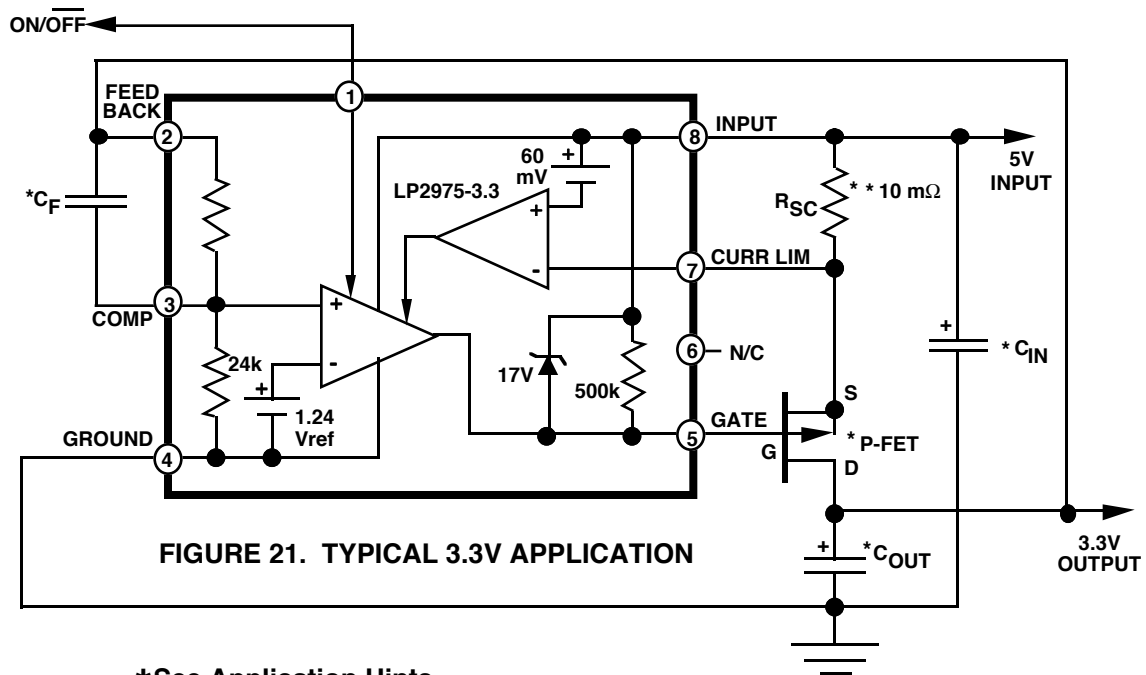
The regulated output can be turned ON and OFF using a logic-level signal applied to the ON/OFF input (pin 1).

Detailed information on LP2975 performance as well as reference designs (with test data) can be found on the data sheet. Information is also provided about compensation and general application information as applied to LDO regulators.



TYPICAL APPLICATIONS (LP2975)

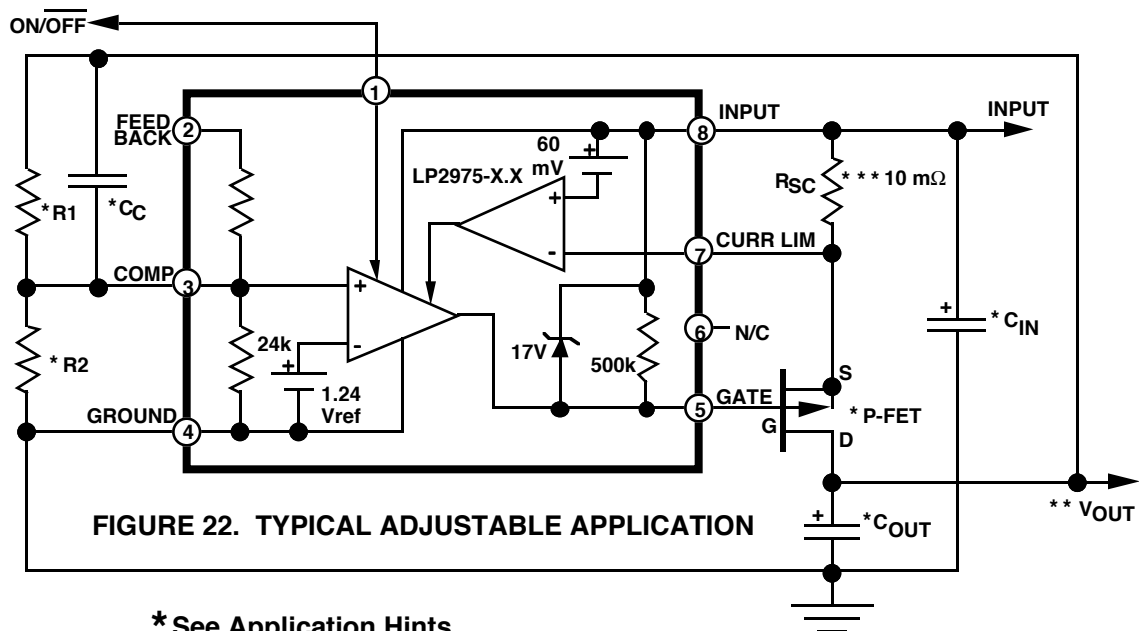
5V - 3.3V @ 5A LDO REGULATOR



*See Application Hints.

If current limiting is not required, short out this resistor.

ADJUSTABLE VOLTAGE 5A LDO REGULATOR



* See Application Hints.

** $V_{OUT} = 1.23 \times (R1/R_{EQ} + 1)$, where $R_{EQ} = (R2 \times 24k) / (R2 + 24k)$

*** If current limiting is not required, short out this resistor.