INTRODUCTION:
Switching power supplies use closed-loop feedback to achieve design objectives for line and load regulation and dynamic response. Fortunately, the closed-loop systems used in switching power supplies are usually not very complicated, permitting the use of simple analytical techniques to achieve loop stabilization. A simplified version of the Nyquist stability criteria can be used because unity gain crossover occurs only once in the gain vs. frequency characteristic. Bode plots provide a simple and powerful method of displaying and calculating the loop gain parameters (see Appendix B). This paper begins with a quick review of basic control loop theory.

Linear Control Loop Theory
As shown in Figure 1, a power supply feedback loop can be described in terms of small-signal linear equivalent gain blocks. The (s) appended to certain gain blocks indicates that the gain varies as a function of frequency.

\[ T(s) = K_{EA} \cdot K_{MOD} \cdot K_{PWR} \cdot K_{LC} \cdot K_{FB} \] (1)

Closed-loop gain, \( G \), defines the output vs. control input relationship, with the loop closed:

\[ T(s) = K_{EA} \cdot K_{MOD} \cdot K_{PWR} \cdot K_{LC} \cdot K_{FB} \] (1)

**Open-loop and closed-loop gain:**
The open-loop gain, \( T \), is defined as the total gain around the entire feedback loop (whether the loop is actually open, for purpose of measurement, or closed, in normal operation).

\[ T(s) = K_{EA} \cdot K_{MOD} \cdot K_{PWR} \cdot K_{LC} \cdot K_{FB} \] (1)

Closed-loop gain, \( G \), defines the output vs. control input relationship, with the loop closed:
\[ G(s) = \frac{1}{K_{FB}} \frac{T}{1 + T} \]  

(2)

At low frequencies, open-loop gain \( T \) is normally very much greater than 1, so that closed-loop gain \( G \) approaches the ideal \( 1/K_{FB} \). At higher frequencies, \( T \) diminishes, mostly because of the low-pass filter characteristic \( K_{LC}(s) \). The frequency where \( T \) has diminished to 1 (0dB) is defined as the crossover frequency, \( f_C \). Referring to Eq. 2 and Figure 2, at \( f_C \) (where \( T = 1 \), with associated 90° phase lag), the closed-loop gain \( G(s) \) is 3db down (with 45° phase lag). Thus, the open-loop crossover frequency is also the closed-loop “corner frequency”, where \( G(s) \) rolls off.

In a power supply voltage control loop, \( G(s) \) defines the power supply output vs. the reference voltage. \( K_{FB} \) is usually a simple voltage divider. For example, if \( V_{REF} \) is 2.5 V, a 2:1 divider \( (K_{FB} = 0.5, G = 2) \) results in \( V_{OUT} = 5 \) Volts. (Refer to Appendix A.)

In a two-loop system (as with current-mode control, to be discussed later) the closed-loop gain \( G(s) \) of the inner loop is one element of the open-loop gain \( T(s) \) of the outer loop.

“Gain” elements as shown in Figure 1 need not have the same units for their output and input (such as Volts/Volt). If Fig. 1 is a current mode control loop, “Output” is a current source, and \( K_{FB} \) is most likely a current sense resistor. \( K_{FB} \) “gain” is then expressed in Volts/Amp, and closed loop gain \( G(s) \) is actually a transconductance (Amps/Volt). Pulse width modulator \( K_{MOD} \) has its gain expressed as \( d/V \) (Duty cycle/Volt). This discrepancy in “gain” units is resolved in the next gain block, \( K_{PWR} \), whose characteristic is \( V/d \).

Overall open-loop gain \( T(s) \) determines how much output error results from a disturbance introduced at any point in the loop compared to the result if the loop was open. Project the disturbance forward to the output (multiply by the gain between the disturbance and the output), then divide by total open-loop gain, \( T \). For example, with no feedback (open loop, constant duty cycle), a 10% change in \( V_{IN} \) results in a 10% \( V_{OUT} \) change. With the feedback loop closed, if \( T = 100 \) at the frequency of the disturbance (DC in this example), then the \( V_{OUT} \) change is only 0.1% (10%/100). Note that the output accuracy does not depend significantly on open-loop gain accuracy. In the example above, if \( T \) was 80 instead of 100, \( V_{OUT} \) would change by 0.125% (10% \( V_{IN}/80 \)), instead of 0.1%. However, output accuracy does depend directly on the accuracy of the feedback portion of the control loop, \( K_{FB} \).

Alternatively, a disturbance can be projected back to the summing point at the input of the error amplifier. For example, the 1Volt “valley” voltage of the sawtooth ramp applied to the PWM comparator is effectively a 1Volt DC offset or “disturbance”. If the E/A gain is 1000, this 1V error is equivalent to a 1mV error in the reference voltage, and translates into the same percentage error at the output.

**Nyquist Stability Criteria:**

Referring to Figure 2, if the open-loop gain \( T \) crosses 1 (0 dB) only once, the system is stable if the phase lag at the crossover frequency, \( f_c \), is less than 180° (in addition to the normal 180° phase shift associated with any negative feedback system). Let us define the term “phase lag” to refer to any additional amount of phase lag beyond the 180° inherent with negative feedback. If the (addi-
(tional) phase lag at $f_C$ exceeds $180^\circ$, the loop will oscillate at frequency $f_C$.

The “phase margin” is the amount by which the phase lag at $f_C$ is less than the critical value of $180^\circ$. The “gain margin” is the factor by which the gain is less than unity (0 dB) at the frequency where the phase lag reaches $180^\circ$. If the phase lag at $f_C$ is only a few degrees less than $180^\circ$ (small phase margin), the system will be stable, but will exhibit considerable overshoot and ringing at frequency $f_C$. A phase margin of $45^\circ$ provides for good response with a little overshoot, but no ringing.

Note that Nyquist’s $180^\circ$ phase limit applies only at $f_C$. At frequencies below $f_C$, the phase lag is permitted to exceed $180^\circ$, even though the open-loop gain is very much greater than 1. The system is then said to be conditionally stable. But if the loop gain temporarily decreases so that $f_C$ moves down into the frequency range where the phase lag exceeds $180^\circ$, conditional stability is violated and the loop becomes unstable. This actually does occur whenever the system runs into large signal bounds, such as when a large step load change occurs. The system will then oscillate and probably never recover. So it is not a good practice to depend upon a conditionally stable loop.

**How can the loop be stable with $180^\circ$ phase lag and gain much greater than 1??**

Figure 3 shows the summing point voltage vectors at a frequency where the open loop gain is 10, for three different amounts of phase lag around the loop.

Figure 3a shows the vector relationship with zero additional phase lag. This condition usually occurs at low frequencies where there are no active poles, so that the gain characteristic slope is zero (flat). The feedback voltage $v_{FB}$ is 10 times greater than error voltage $v_E$ and $180^\circ$ out of phase. (Note that with an open-loop gain of only 10, the $v_E$ magnitude causes $v_C$ to be less than $v_{FB}$. This inequality diminishes with higher loop gain.)

Figure 3b shows the vector relationship with a gain of 10 but at a frequency where one pole is active, resulting in –1 gain slope and $90^\circ$ phase
lag. Feedback voltage $v_{FB}$ is 10 times greater than $v_E$, but lags by 270°. Note that $v_E$ now causes very little inequality between $v_C$ and $v_{FB}$ because of its phase. This situation is perfectly stable. With $v_C = 1 \text{ V}$ and open-loop gain of 10 with 90° phase lag, only this outcome is possible.

In Figure 3c, two poles are active at the frequency where the gain is 10, resulting in –2 gain slope and 180° additional phase lag. Feedback voltage $v_{FB}$ is now in-phase with $v_E$ and 10 times greater. Our intuition tells us that this should be a runaway situation. But intuition is wrong, when our thinking is restricted to this one frequency. The vector relationships in Fig. 3c are perfectly stable. They are locked in to each other. This is the only way they can exist, under the defined conditions. Note that $v_E$ now causes $v_{FB}$ to be greater than $v_C$. This does not signify instability – in fact, if the gain is increased further, $v_{FB}$ becomes smaller, reducing the error without becoming unstable.

**Why does oscillation occur only at $f_C$, where the open loop gain equals 1??**

The vectors of Figure 4a show the stable condition that exists when the gain slope is –1 as it passes through the crossover frequency. The single active pole results in 90° phase lag. Feedback voltage $v_{FB}$ is equal to $v_C$, but lags by 270°. Again, this is the only possible relationship between these vectors under the conditions defined. Note that $v_{FB}$, which represents the output, lags control voltage $v_C$ by 45° (plus 180° negative feedback), and the magnitude is down 3dB to .707 (compared with Fig. 3). This represents the closed-loop gain corner at the open loop crossover frequency, as shown in Fig. 2.

The vector diagram for a –2 gain slope at $f_C$ where open-loop gain equals 1 cannot be drawn, as it is unstable. Figure 4b shows the vectors at a gain of 1.2, instead. With a –2 slope, $v_E$ and $v_{FB}$ are in-phase. With a control voltage $v_C$ of 1V, a feedback voltage of 6 V with an error voltage of 5 V is required to resolve the vector diagram. As the loop gain approaches 1, it can be seen that either $v_C$ must become zero, or $v_E$ and $v_{FB}$ must become infinite. Thus, the closed loop gain, $v_{FB}/v_C$ becomes infinite, even though the open-loop gain is 1. The system is definitely unstable.

**How to design a stable loop:**

The first step in the design of a stable, high performance feedback loop is to define the gain/phase characteristic of each of the known loop elements (usually everything except the error amplifier, $K_{EA}$). Then, the characteristic of the remaining elements ($K_{EA}$) is tailored to complement the combined characteristics of the other elements in a way that will meet the overall loop stability criteria while achieving the highest possible loop gain and bandwidth.

In a switching power supply, the loop elements which actually handle the power are mostly defined by the parameters of the application. However, many options do exist, and they should be explored. (Design experience helps to narrow down the list of possible options.) Bode plots (Appendix B) are used to display the overall characteristics of all of the loop elements except $K_{EA}$. With performance objectives and stability requirements in mind, a strategy for closing the loop is developed and a tentative gain characteristic is plotted to define the goal for the entire loop. The required $K_{EA}$ characteristic (Appendix B) is then deduced from the difference between the Bode plot of the overall loop goal and the plot of the known loop elements without $K_{EA}$.

**Limitations on crossover frequency:**

![Figure 4. – Vector Relationships at Crossover](image-url)
Achieving a high $f_C$ is a worthwhile objective because the system can respond more rapidly to minimize the effects of high frequency and transient disturbances. In a purely linear feedback loop, $f_C$ is limited by cumulative phase lags in various system elements. These phase lags inevitably increase with frequency in a manner that often varies unpredictably. Compensation becomes impossible, forcing the designer to set $f_C$ at a frequency where the phase lags are still manageable.

In switching power supply loops, an additional important limitation occurs. Sampling delays inherent in any switched system introduce additional phase lags that force the crossover frequency to be well below the switching frequency. This will be discussed later.

**Transient Response:**

Transient behavior, in the time domain, is predictably related to the shape of the loop frequency domain characteristics as shown in the Bode plot.

A power supply can function without the help of a feedback loop. The duty cycle could be adjusted manually to the value that would provide the desired $V_{OUT}$. But without feedback, even small changes in $V_{IN}$ or $I_{OUT}$ (the usual disturbances in a power supply application) would send $V_{OUT}$ careening out of spec. With a functional feedback loop, when an ac disturbance at a specific frequency is introduced, the open-loop gain magnitude at the frequency of the disturbance defines how much the output disturbance is reduced compared to what would have occurred without feedback.

Figure 5a is the Bode plot of a loop having the gain characteristic of a single pole ($-1$ slope, 20dB/decade). A crossover frequency of 10kHz is shown, with the open-loop gain rising to 1000 at 10Hz. The gain shown at each frequency indicates the amount by which the feedback loop will reduce a disturbance at that frequency.

The gain vs. frequency plot can also be used to show the reduction in the Fourier components of a transient disturbance, or how the loop will respond to the Fourier components of a step change in the control signal. Fortunately, Fourier analysis is usually not required to interrelate the Bode plot characteristic, in the frequency domain, with the transient response in the time domain. For example, the initial slope of the transient response to a step change is directly related to the crossover frequency.

The simple single pole characteristic of Fig. 5a has an exponential characteristic with a time constant equal to $1/2\pi f_C$, as shown in Figure 5b. In responding to a step change, the initial slope would reach the final value in exactly one time constant (16µsec in this example), but like any exponential, it falls away to 63% of the final value at 1 time con-
stant and reaches 98% (2% error) in 4 time constants (64µsec). It takes a long time for the error to diminish ultimately to 0.1% because the loop gain reaches 1000 only for the Fourier components below 10Hz.

The single pole characteristic depicted in Figure 5 is extremely conservative. The –1 slope with its 90° phase margin results in the exponential characteristic which takes a long time to achieve good accuracy.

Figure 6 shows a less conservative approach which reduces the error much more rapidly. Two active poles provide a –2 slope below f_C raising the gain below f_C. This improves audio susceptibility at these frequencies, and improves response to the higher frequency Fourier components of a transient disturbance or control signal. As shown in Fig. 6a, the gain reaches 1000 at 300Hz, rather than at 10Hz. Note that at f_C, the –2 gain slope transitions to a single pole –1 slope. This is necessary because if the –2 slope continued above f_C, the phase margin would be too small, resulting in severe underdamped oscillations at f_C. The transition to a single pole at f_C results in an acceptable phase margin of 52°.

Figure 6b shows that the initial slope is the same as in Figure 5b, because f_C is the same in both cases. But the transient response holds up better because the gain rises more rapidly at the frequencies below f_C. However, this results in 16% overshoot, which occurs at .58/f_C (58µsec in this example).

Although the peak error with the –2 slope exceeds the error at the same time with the –1 slope, it subsequently diminishes more rapidly. What is more, the overshoot is actually beneficial in some situations.

For example, in a power supply application with an inner current control loop and an outer voltage control loop, assume Figure 5b shows the transient response of the current control loop to a step change in load current at time 0. The load current rises immediately to the final value, but the source current follows the transient response characteristic. Area “A” shows the charge deficit that results. The load draws this deficit from the output filter capacitor, whose voltage sags as a result. Ultimately, the output voltage is restored and the charge deficit made up only because the voltage loop responds to the voltage sag and calls for source current temporarily greater than the final value. However, this voltage loop intervention takes considerable additional time.

Figure 6b shows that with two active poles, not only is the charge deficit “A” reduced, but the overshoot results in a charge excess “B” which cancels all or part of the charge deficit immediately, without
requiring voltage loop intervention.

Switching Power Supply Loops

Power Circuit Design:

Just as the power supply is often the step-child in the design of the complete system, the control loop is often the step-child in the design of the power supply. The power handling circuit topology with its associated components is the most significant portion of the control loop design, causing most of the problems and complexity. The power circuit is usually defined first, attempting to implement system requirements in the most cost-effective way, with little consideration given to control loop closure. The control loop design usually must adapt to a predefined power circuit.

Before proceeding with the control loop it is necessary to examine some of the power circuit choices that must be made. This is a difficult subject to organize, because of the complex interactions between these choices.

**Choices:**
- Power Circuit Topology
- Control Method
- Transformer Turns Ratio
- Switching Frequency
- Filter Capacitor
- Filter Inductor

**Considerations:**
- Cost
- Size/Weight
- Efficiency
- Noise

**Switchmode Topologies:**

In the basic buck, boost and flyback power circuit topologies, shown in Figure 7, the inductor is the element which transfers power from the input to the output. (In the unique Cuk converter — a dual of the flyback — a capacitor is the energy transfer element.) The power switch is turned on and off during each switching period by a Pulse Width Modulator (PWM). The duty cycle, D, (the percentage of time the switch is ON) is the basis for controlling the output. An output filter averages the power pulses to obtain a DC output with acceptable ripple.

**Continuous Current Mode (CCM):**

This operating mode occurs, by definition, when inductor current flows continuously throughout the switching period. The CCM current waveforms, shown in Figure 8, apply to all three topologies. But, referring to Figure 7, input and output currents differ for each topology because of the different locations of the inductor, switch and diode. There are two operational states—Switch ON, when it carries the inductor current, or Switch OFF, when the diode carries the inductor current.

Under steady-state conditions, inductor voltage \( V_L \) must average zero during each switching period. With only two states, a specific, rigid rela-

---

**Fig. 7. - Basic Buck, Boost, Flyback Topologies**
A relationship exists between input voltage $V_i$, output voltage $V_o$ and duty cycle $D$, a relationship that is independent of load current and is unique for each topology:

Most switching power supplies are designed to operate in the continuous mode, especially at higher power levels, because filtering is easier and noise is less. Boost and flyback circuits operated in the CCM have a unique problem — their control loop characteristic includes a right half-plane zero that makes loop compensation very difficult.

**Discontinuous Current Mode (DCM):**

As shown in Figure 9, the discontinuous inductor current mode occurs when the inductor current, flowing through the diode, reaches zero before the end of the switching period. The diode prevents the current from continuing in the negative direction. Thus, the inductor current remains at zero until the switch turns on at the beginning of the next switching period. This zero current interval is a third operating state in addition to the two that exist with CCM, and the additional degree of freedom that this provides destroys the rigid $V_i$, $V_o$, and $D$ relationship. With DCM operation, the small signal gain of the power circuit is much less than in the continuous mode, and DCM gain varies considerably with load.

However, the DCM control characteristic is simpler, especially with the boost and flyback topologies because the right half-plane zero does not exist. For this reason, the flyback topology is often used in the discontinuous mode at low power levels where noise and filtering problems are not as severe.

The Pulse Width Modulator controls the duty cycle of the power switch — the fraction of time that the switch is ON during each switching period.

---

Fig. 8. - Continuous Mode Waveforms

![Continuous Mode Waveforms](image)

Fig. 9. - Discontinuous Mode

![Discontinuous Mode](image)
The ON/OFF action of the power circuit is averaged and filtered to provide a dc output. The output magnitude is related to the duty cycle, D, thus the pulse width modulator (PWM) provides the basis for control and regulation of the output.

There are many varieties of pulse width modulators: Fixed frequency - variable duty cycle, Fixed ON-time (Variable Frequency), Fixed OFF-time (VF), Hysteretic (VF). The choice of PWM method significantly affects power circuit behavior and small-signal characteristics and thus on the strategy for closing the feedback loop.

This paper considers only fixed frequency PWM methods, which are used in the great majority of control ICs. Fixed frequency operation is preferred because it permits the switching frequency to be synchronized with other power supplies in a system, or with video terminal horizontal sweep frequency, to prevent spurious beat frequencies and other undesirable effects. Also, fixed frequency control loops have simpler relationships which are much easier to understand and optimize.

Fixed frequency PWMs function on the basis of a latching comparator as shown in Figures 10 and 11. (Latching prevents spurious reset due to noise.) A control voltage, \( V_C \), (usually the amplified error signal from the controlled output) is compared to a fixed frequency linear sawtooth ramp, \( V_S \). The comparator output provides fixed frequency pulses of variable duty cycle which drive the power switching transistors. The duty cycle \( D \) of the power switch conduction is thereby controlled by varying \( V_C \) according to the relationship shown in Eq. 3. (\( D, V_C, V_S \) are dc values, \( d, v_C \) are small-signal ac or incremental values.)

\[
D = \frac{V_C}{V_S} \quad ; \quad d = \frac{v_C}{V_S}
\]  

The PWM waveforms of Fig. 11 can be observed only in very low bandwidth loops. In a high-performance loop with \( f_C \) near optimum, control voltage \( V_C \) is not flat, as shown, but has a superimposed triangular waveform (derived from inductor ripple current) that approaches the magnitude of sawtooth voltage \( V_S \). The superimposed triangular waveform modifies the duty cycle relationship of Eq. 3, and can also cause subharmonic oscillation. This will be discussed later. Until then, the idealized waveforms of Fig. 11 will be used.

**Modulator Phase Lag:**

Virtually all fixed-frequency PWM control ICs use the simple comparator method shown in Figs 10 and 11. The output pulse is terminated according to the instantaneous value of the feedback control voltage at the moment of pulse termination. This “naturally sampled” method of pulse width modulation ideally results in zero phase lag in the modulator and in the converter power switching stage. In practice, however, comparator delays and turn-off delays in the power switch will cause a phase lag.
directly proportional to the delay time, \( t_d \), and signal frequency, \( f \), according to the relationship:

\[
\theta_m = 360t_D/T = 360t_Df \quad (4)
\]

This additional phase lag reduces the phase margin at the unity gain crossover frequency and theoretically may contribute to control loop instability. However, the additional lag is usually negligible. For example, at an \( f_C \) of 25kHz, consistent with \( f_S = 200kHz \), a turnoff delay of 0.4 \( \mu \)sec in the IC and the power switch causes only 3.6° additional phase lag, reducing phase margin by that amount.

Most control ICs have additional “housekeeping functions” such as UVLO - UnderVoltage LockOut, HVLO - HighVoltage LockOut, and Soft Start, which are not discussed in this paper as they are not directly relevant to control loop design.

**Design Relationships – Buck-Derived Topologies:**

In addition to the basic buck regulator, transformer-coupled buck-derived topologies include the single-ended Forward Converter and a variety of push-pull converters: Center-tap, Full Bridge, and Half-Bridge.

The basic relationship governing the power circuit of all buck-derived topologies operated with continuous inductor current is:

\[
V_O = V_I D; \quad v_O = V_I d \quad (5)
\]

\[
V_{I_{min}} = V_O/D_{max} \quad (6)
\]

**Duty Cycle Range:**

It is theoretically possible for the basic buck regulator and its push-pull transformer-coupled derivatives to utilize the full 0 to 1 duty cycle range, but \( D \) close to 1 is best, as it results in the lowest primary-side current and lowest secondary voltages. (The boost topology functions most effectively with \( D \) close to 0, the flyback with \( D \) close to 0.5.)

As shown in Eq. 6, for the buck regulator, the minimum \( V_I \) at which the circuit can function is defined by \( D_{MAX} \). In transformer coupled topologies, the minimum \( V_I \) defines the transformer turns ratio.

\( D_{MAX} \) can never reach 1 because of practical limitations. Some of these limitations are: turn-on propagation delays and switch delay & rise times, resonant transition times, and reset time for the current sense transformer, if a CT is used. \( D_{MAX} \) is typically limited to between 0.85 – 0.95. Any application involving a transformer must provide time to reset the transformer core – the reverse volt-seconds must equal the forward volt-seconds to get the flux back to the starting point. Push-pull circuits automatically reset the core by driving it in opposite directions during successive switching periods. The Forward Converter has the most serious problem – it is driven in only one direction, and the subsequent voltage reversal required for core reset typically equals the time driven in the forward direc-

---

**Fig. 12. - Forward Converter**
tion, thus limiting $D_{\text{MAX}}$ to less than 0.5. This means that the minimum $V_{\text{IN}}$ referred to the secondary side must be greater than twice $V_{\text{OUT}}$.

**Minimum Duty Cycle:**

Likewise, $D_{\text{MIN}}$ cannot reach zero. Once the switch is turned ON to initiate a power pulse, the switch is committed to stay ON for a certain minimum time. This minimum pulse width at a fixed switching frequency equates to a minimum duty cycle. Some of the items that contribute to $D_{\text{MIN}}$ are: turn-off propagation delays and switch delay & fall times, resonant transition times, and noise blanking (which disables the PWM comparator for a short time after turn-on to prevent a spurious noise pulse from causing premature turn-off).

In normal operation, $D$ is always much greater than zero. Certain events will cause $D$ to approach zero temporarily, such as when load current diminishes at a rate faster than inductor current can decrease (max $\frac{dI}{dt} = \frac{V_{\text{OUT}}}{L}$). In this situation the $D_{\text{MIN}}$ value attained is not critical. The $D_{\text{MIN}}$ value does become critical when the output is short-circuited. When $V_{\text{OUT}}$ is pulled down to zero, and $V_{\text{IN}}$ is at its normal value, then $D$ must be brought to zero to maintain control and keep the current within the limit. This bleak situation is remedied by the output rectifier forward drop which acts as a minimum $V_{\text{OUT}}$. But when $V_{\text{IN}}$ is near maximum, and especially when $V_{\text{OUT}}$ is 28 V or higher and the rectifier drop has less significance, the required $D$ value may still be less than $D_{\text{MIN}}$. This is then a serious problem. Many control ICs always initiate an output pulse at the beginning of each clock cycle, relying on current limiting to turn off the power pulse quickly under overload or short circuit conditions. But “quickly” may not be quick enough.

The solution employed in many modern ICs is to skip pulses, or shift the frequency downward. Under overload conditions, if pulses are skipped entirely, the switching frequency effectively adjusts downward. The minimum pulse width does not get smaller, but $D$ does become small enough to retain control. Pulse skipping requires a control IC that has the logic to completely inhibit switch turn-on if current exceeds the limit at the beginning of the clock cycle.

**Transformer Turns Ratio:**

First, the minimum input voltage referred to the secondary side, $V_i$, is determined. Using Eq. 6, calculate $\min V_i$ based on $D_{\text{MAX}}$, then add full load switch, diode and IR drops. Allow for some additional voltage across the inductor, or its current cannot increase rapidly under min $V_i$ conditions when necessary to keep up with a load current increase. With this adjusted min $V_i$ value, and the minimum source voltage, $V_{\text{IN}}$, the turns ratio can be calculated:

$$V_i = \frac{V_{\text{IN}}}{n} ; \quad (n = \frac{N_p}{N_S})$$

In this paper, to minimize the complexity of the control loop relationships, all circuit values are referred to the secondary side. Thus, turns ratio $n$ and actual input source $V_{\text{IN}}$ do not appear, only $V_i$, the input voltage referred to the secondary.

For low voltage outputs, accuracy is improved by adding the output rectifier forward drop to the actual output voltage, using this “corrected” value of $V_o$ in the design equations.

**Inductor Ripple Current** is inversely proportional to inductance value. In buck-derived topologies a small inductor with large ripple current has these disadvantages: (1) a bigger output filter capacitor is required, (2) large ripple dictates a large minimum load current to avoid discontinuous operation. (This disadvantage is overcome by using Average Current Mode Control.) Advantages of the smaller inductor are: (1) Lower size and cost, (2) inductor current can change more rapidly in response to a sudden load change and (3) together with the larger $C_o$, reduces over/undershoot occurring with a large step load change.

The inductance value obviously plays a key role in the control loop design.

**Filter Capacitors**

Output filter capacitors are almost certainly the most troublesome element in the control loop. In their power filtering role, they typically absorb Amperes of ripple current and hold the output ripple voltage to a small fraction of a Volt. The low
impedance required usually dictates the use of electrolytic capacitors. Ceramic capacitors are not usually considered practical unless the switching frequency is well over 500kHz and/or with high output voltages.

**Electrolytic Capacitors – Series Resistance:**

At the 50-400kHz switching frequencies mainly used in today’s SMPS applications, electrolytic capacitor impedance is determined by its series resistance, SR. As frequency is increased, when capacitive reactance drops below series resistance, the impedance curve tends to flatten out at the SR value. The frequency at which this occurs (the ESR zero frequency) is 1 to 10kHz for Aluminum electrolytics, 10 to 60kHz for Tantalum. Almost all power supplies today switch at frequencies well above this. Electrolytic capacitors must then be selected and specified on the basis of their series resistance. The resulting capacitance values are much greater than would be required if the SR were not dominant – often 100 times greater with aluminum electrolytics at 200kHz switching frequency.

At switching frequencies above f_{ESR}, the impedance characteristic flattens out at the SR value, so that the same capacitor is required regardless of the frequency. Going to a higher f_S does not change the filter capacitor or reduce its cost.

**SR or ESR??**

Electrolytic capacitors have both series and parallel resistance components. At low frequencies where capacitive reactance is large, the parallel resistance (leakage through the dielectric) dominates, and true series resistance (mostly in the electrolyte) is negligible. Measurements taken on a bridge cannot distinguish between actual parallel and actual series resistance. Bridge measurements lump both resistances together – the actual series resistance plus the parallel resistance converted to its series equivalent. This combination is called “Equivalent Series Resistance”, or ESR. At low frequency (50-60Hz), the converted parallel resistance dominates. Capacitive reactance, the fulcrum of the parallel to series conversion, varies inversely with frequency, which makes ESR appear to vary inversely with frequency squared.

In a switching power supply application, the actual series resistance SR is of key importance, but the parallel resistance is of little or no significance (except possibly for reliability concerns). So ESR data is very misleading until the frequency is high enough that the converted parallel resistance becomes smaller than the true series resistance. At higher frequencies, the ESR characteristic flattens out at the true SR value. Capacitors intended for high frequency application are measured and specified at 100kHz which reveals the true series resistance. Low frequency ESR measurements are totally irrelevant. However, bowing to common usage, this document uses “ESR” to refer to the actual series resistance evident at high frequency.

**Capacitance and ESR variation:**

The impedance transition from capacitive (with –1 slope) to resistive (with 0 slope) puts a zero in the control loop Bode plot. The frequency at which this occurs is called the ESR zero frequency, f_{ESR}:

\[
f_{ESR} = \frac{1}{2\pi f_S}
\]

The problem with aluminum electrolytics in the control loop is that f_{ESR} is usually near or below the desired crossover frequency. ESR variation causes a corresponding f_{ESR} variation. This results in variable loop gain and variable phase margin, making it difficult to cross over above f_{ESR}. If the supply must operate over a wide temperature range, the large ESR variation with temperature can make it impossible, forcing the design to cross over at a low frequency (probably below 1kHz).

Capacitance variation is quite small, so that below f_{ESR} the characteristic is stable and predictable. Data from Panasonic on the FA Series Aluminum Electrolytics:

**Capacitance:**

- 20°C distr.: 100%–120% of spec. value
- +10% @ 105°C; −10% @ −55°C

**ESR:**

- 20°C distr.: 60% – 85% of specified max.
A Little Trickery:
Electrolytic capacitors with the same case size and manufacture but with different voltage ratings and capacitance values all tend to have the same ESR. The dielectric oxide thickness which determines the voltage vs. capacitance tradeoff is “formed” late in the manufacturing process. The dielectric thickness does not significantly affect ESR. For example, in a 16x20 mm case size, Panasonic FA series 10V, 3300 µF and 50V, 680 µF capacitors have the same ESR: 25 mΩ max.

For SMPS ripple filtering, electrolytic capacitor selection is based entirely on the ESR requirement. A 5V output requiring 25 mΩ max. ESR could use either of the above capacitors. The 3300 µF, 10V capacitor puts a 2kHz ESR zero into the control loop, but the 680 µF, 50V puts the ESR zero at 10kHz. Thus, if it is necessary or desirable to make the loop gain crossover below fESR to avoid the problems caused by ESR variability and unpredictability, the smaller capacitance value with the higher fESR is clearly the better choice.

There is a downside to this choice, however. In the continuous conduction mode, the filter inductance prevents the inductor current from responding rapidly to a step load current change. The output filter capacitance (not the ESR) absorbs the load current change while the inductor current catches up. The extravagantly excessive capacitance value necessary with electrolytic capacitors does become very useful by providing a very low output surge impedance — it “holds the fort” until reinforcements arrive. The faster control loop does nothing to help in this situation — this is a large-signal limitation dictated by inadequate inductor current slew rate, during which the control error amplifier is driven to its limits and the loop is temporarily open and non-functional.

Ripple Current Rating:
AC ripple current flowing through the capacitor ESR generates heat. Temperature rise and reliability considerations are the basis for an rms current limit. The low ESR capacitors normally used in SMPS applications have rms current ratings that are usually adequate for their purpose. To calculate the rms equivalent of the peak-peak triangular inductor ripple-current waveform:

$$I_{\text{rms}} = \frac{I_{pp}}{2\sqrt{3}}$$

(8)

Capacitor Inductance:
The path for ac current flow within an aluminum electrolytic capacitor is quite long, simply because of their relatively large size. This results in larger series inductance than other capacitor types. The impedance characteristic is determined by ESR above fESR, but at approximately 500kHz, the impedance rises because the series inductance becomes dominant. Other capacitor types then become more advantageous.

Tantalum Capacitors:
Characteristics are similar to aluminum electrolytics, but tantalum electrolytics are better: The ESR zero frequency is 5-10 times higher than aluminum, making it easier to achieve greater loop bandwidth, with improved dynamic response. (But ESR remains the impedance determining factor for ripple filtering at SMPS switching frequencies.) The ESR has a much lower temperature coefficient, making tantalum much better suited to military and other wide temperature range applications. Size is much smaller for the same ESR. The smaller size also results in lower inductance, enabling operation up to 1MHz.

The downside for tantalum capacitors is substantially higher cost for the same ESR required. Also, the lower capacitance value associated with the necessary ESR (the reason why fESR is greater) results in a higher output surge impedance, so the output does not stand up as well to a large step load change.

Ceramic Capacitors:
Radically different from the electrolytics, ESR is negligible — an ESR zero frequency doesn’t exist. Impedance is not determined by ESR, but by capacitance (or by inductance at frequencies above 1-2MHz). Small size, surface mount packaging keeps inductance the lowest of all the alternatives.

But the cost of obtaining the necessary capaci-
itance with ceramic is excessive at switching frequencies below 500kHz. Even at higher frequencies, to achieve the required capacitance at a reasonable cost, high K dielectrics are used. The large temperature coefficients of these dielectrics make it difficult to optimize the loop over a wide temperature range. Also, the C value required to obtain the required ripple reduction is much less that the capacitance obtained by default with the electrolytics. This results in relatively high output surge impedance and little tolerance for step load changes.

**Polymer Aluminum Electrolytics:**

Similar to ceramics, these new arrivals in the capacitor catalog have negligible ESR, small size, low inductance, but high output surge impedance. But here the similarity ends. Available capacitance values are not only much greater than ceramics, but capacitance distributions are tight, and temperature coefficients are low. Polymer aluminum electrolytics approach the ideal for filter capacitors.

The limitations of the existing devices are: Low voltage ratings: 16V max. One small size surface mount package available: (8mm x 5.3mm x 3.3mm high) limits C values to the range of 6 - 33 μF. Higher cost unless the switching frequency is high enough to overcome this.

In a 200kHz power supply, one of the presently available polymer aluminum electrolytics will handle the filtering of a 5V, 25A buck regulator output at perhaps twice the cost (in Jan'96) of a competitive (but much larger) aluminum electrolytic. At f_S of 400kHz, they are probably cost-competitive.

Panasonic states that their polymer electrolytics are now being used as output filters in switching power supplies. If these devices fulfill their promise and are made available in larger sizes with greater capacitance values and/or if switching frequencies continue to rise, perhaps they will some day come to dominate this application.

**Switching Frequency:**

The rationale for the inexorable rise in SMPS switching frequency over the years has been reduced cost as well as reduced size and weight. The smaller magnetic components made possible by raising the frequency have helped the most to achieve these goals. But at frequencies above 500kHz, core losses in today's best magnetic materials (1996) rise to the point where this trend slows down and then reverses — the magnetic components start to get larger. The filter capacitor might be expected to get smaller with increased frequency, but it does not because its impedance depends on ESR, not capacitance — until the frequency is reached where ceramic capacitors become economically feasible. At higher switching frequencies, there is more high frequency noise generated, but less low frequency noise, so that conducted EMI is easier and less costly to filter. The control loop bandwidth can of course be raised proportional to f_S, but this is seldom part of the rationale for increased frequency.

The obstacles to achieving higher switching frequencies at reduced cost all seem to boil down to one thing: increased losses, which lower efficiency and raise the cost of heat removal. Ongoing improvement involves circuit topologies and innovative techniques such as the recently popular “resonant switching transitions” which reduce losses and noise. Improved high frequency magnetic materials are needed, as well as faster semiconductors. New concepts in the “wiring” and layout of high frequency circuits and magnetic components are needed to reduce parasitic inductances which increase losses, impair regulation, and radiate EMI.

**Control Methods**

**Voltage Mode Control:**

The earliest control method, implemented in most older control IC’s. This was discussed previously (refer back to Figs. 10 and 11). The fixed amplitude sawtooth ramp is usually taken from the control IC’s clock generator. VMC disadvantages are: (1) No voltage feedforward to anticipate the affects of input voltage changes. Thus, slow response to sudden input changes, poor audio susceptibility and poor open loop line regulation, requiring higher loop gain to achieve specifications. (2) In continuous mode regulators, provides no help in dealing with the resonant two pole filter characteristic with its sudden 180° phase shift.
Control changes must propagate through these two filter poles to make a desired output correction, resulting in poor dynamic response. While VMC might appear to be less costly because there is no current loop with its need for current sensing, but current limiting is almost always required, and this requires the current to be sensed. With Current Mode Control, current limiting is automatic and "free".

(Peak) Current Mode Control:
This control method (CMC) also controls the duty cycle by comparing the control voltage to a fixed frequency sawtooth ramp, but the ramp is not derived artificially from a ramp generator, as with Voltage Mode Control. The ramp is actually the inductor ripple current, as it rises while the switch is ON, translated into a voltage by a current sense resistor. This ramp, representing the inductor current, is fed back to the PWM comparator, forming an inner current control loop. When the current rises to the level of the control voltage, the switch is turned off. The control voltage (which is the amplified output voltage error), thus defines the peak inductor current. The outer voltage control loop programs the inductor current via the inner loop while the current loop directly controls the duty cycle.

In the forward converter shown in Figure 13, the inductor is on the secondary side. But since the

![Figure 13. Peak Current Mode Control](image-url)
control IC is on the primary side, it is easier to sense primary-side switch current. This works because the switch current is the inductor current (while it is rising) divided by the transformer turns ratio. This eliminates the problem of bringing the current information across the isolation boundary.

The advantages of CMC are profound. Most of the problems of Voltage Mode Control are eliminated or reduced. CMC has inherent voltage feedforward and responds instantaneously to input voltage changes. The inductor pole is now located inside the current loop. Instead of the two pole second order filter of the VMC loop the outer voltage loop now has a single pole (the filter capacitor), greatly simplifying loop compensation. The capacitor ESR with its variability remains in the voltage loop.

The CMC closed loop is part of the outer voltage control loop. The CMC closed-loop characteristic approaches an ideal transconductance amplifier. Closed-loop gain is flat up to its open-loop crossover frequency, which is optimally 1/3 to 1/6 of the switching frequency. At the CMC crossover frequency, its closed-loop gain rolls off with a –1 slope, adding a second pole into the outer voltage loop, but at a much higher frequency than the capacitor pole.

Peak current mode control does have its own set of problems: Average current is what should be controlled, but peak is controlled instead. The peak-to-average error is quite large, especially at light loads, and the voltage loop must correct for this, which hurts response time. Open loop gain of the CMC loop is already quite low (5 - 10) in the continuous current mode, but when the load diminishes to the point where inductor current becomes discontinuous, the CMC loop gain plummets and the peak-to-average error becomes huge. Operation becomes unsatisfactory in the discontinuous mode.

Subharmonic Instability:

Switching power supply control loops are all subject to subharmonic instability if the waveforms applied to the two inputs of the PWM comparator do not cross over each other at their points of intersection. This instability is observed as a tendency to oscillate (or a full-blown oscillation) at frequency \( f_S /2 \).

Figure 14 shows the subharmonic instability in a peak CMC loop. Normal operation is shown by the solid triangular waveform labeled \( i_L \). This voltage, representing the inductor current, is applied to one side of the comparator. The switch is turned on by a clock pulse, and \( i_L \) rises until it reaches control voltage \( V_C \) at the other comparator input. The switch turns off, and the current decreases until the next clock pulse occurs. (It does not matter of the current downslope is observed through the current sense resistor—referring to Fig.13—because switch turn-on is by the clock, and not dependent on the current level.)

Using perturbation analysis, a small deviation, \( \Delta \), is assumed in the inductor current. The deviated waveform has the same slopes as before, because the voltages across the inductor have not changed – just the initial current has been changed. The dash line in Fig. 14 reveals the instability. In a stable system, the perturbation gets smaller every switching period.

True subharmonic instability can be eliminated using a slope compensation technique, discussed below. Sometimes, what appears to be subharmonic instability is really noise at the comparator input. When the clock pulse turns the power switch on, much noise is generated. A noise spike at the comparator input can easily turn the switch off immediately, effectively causing one or more entire switching period to be skipped.

Latching Comparator:

When the voltages at the PWM comparator inputs intersect, and the power switch is turned off, the comparator must be designed to latch in that
state until reset by the next clock pulse. Otherwise, if the waveforms trajectories diverge without crossing over, as in Fig. 14, the switch will turn back on immediately. Even if the waveforms do cross over, a noise spike could cause the comparator to reset and turn on the power switch prematurely. The latching comparator prevents these undesired occurrences.

Slope Compensation:

Subharmonic instability is eliminated simply by forcing the waveforms at the two inputs of the comparator to cross over each other at their points of intersection. This can be accomplished by adding an artificial ramp to one of the comparator inputs. Figure 15 shows an optimum slope compensation ramp added to the control voltage comparator input, labeled “VC + VS”. The optimum ramp, as shown, causes the two waveforms at the comparator inputs to coincide during the interval when the switch is off and the inductor current is decreasing, rather than actually cross over. This is ideal, because, as shown, a perturbation is erased in the very first switching period after its occurrence!!

The compensation ramp reduces the current loop gain. If the ramp slope is increased further so that the waveforms actually cross over, the system is stable but the gain is reduced below optimum (and it actually takes longer for the perturbation to be erased). Optimum is when slopes coincide, or match.

The crossover frequency is directly related to the gain. Middlebrook has shown that for a buck-derived regulator with optimum slope compensation, the crossover frequency is:

\[ f_c = \frac{f_s}{2} \]

Thus, depending on duty cycle D, \( f_c \) ranges from 1/3 to 1/6 of \( f_S \).

Although Fig. 15 shows a ramp with a negative slope added to the control voltage waveform (because it is easier to visualize), in practice a positive ramp slope is usually added to the inductor current waveform, simply because a positive ramp is available in the IC’s clock generator.

It can be argued whether subharmonic instability results from the sampling delays inherent in a switched system, or whether it is just a geometry problem. Certainly this instability can either be generated or corrected by adding a purely artificial ramp, unrelated to the loop elements.

Linear models have been attempted so that the effects of subharmonic instability can be included in frequency domain analysis. However, these empirical models lose sight of the underlying causes and are blind to the slope manipulation techniques which can optimize bandwidth without instability. The underlying causes of instability are best demonstrated and corrected in the time domain, observing and appropriately modifying the waveform trajectories on opposite sides of the PWM comparator.

Average Current Mode Control:

The deficiencies of the Peak CMC loop basically relate to its low internal loop gain. Average CMC, as shown in Figure 16, eliminates this problem by adding an error amplifier to the current loop (in addition to the amplifier in the outer voltage loop). Inductor current is sensed through a resistor. The resulting voltage is compared with voltage \( V_{CP} \) which sets the desired inductor current. The differential, representing the current error, is amplified by \( CA \), the current error amplifier. The CA output is compared to a sawtooth ramp taken from the IC clock generator to determine the duty cycle.
the same technique commonly used with Voltage Mode Control.

Figure 17 shows the comparator voltage waveforms when the E/A gain is optimized using the slope matching criteria discussed below. Note that amplifier CA inverts the error signal, so the triangular waveform $V_{CA}$ is an upside-down representation of the inductor ripple current. The rising portion of the $V_{CA}$ waveform (coincident with sawtooth waveform $V_S$) represents falling inductor current, when the switch is OFF. As Figure 17 shows, where the waveforms intersect (near the midpoint of the sawtooth ramp) and the switch turns OFF is where the inductor current is at its peak (the waveform is inverted). Why is this called average CMC if it really functions at the peak?? Actually, average CMC when optimized is identical in its behavior to peak CMC with all of its positive attributes – it has the same crossover frequency, the same instantaneous response to a current overload, etc. But at frequencies below $f_C$, where the peak CMC loop gain flattens out at a gain of only 5 or 10, the gain of the average CMC loop keeps rising, ultimately to a gain of more than 1000 if desired. This much higher loop gain at lower frequencies eliminates the peak-to-average error and enables the average CMC loop to function well at light loads when the inductor current becomes discontinuous.

Reference (2) describes Average CMC in detail.

Slope Matching:

In the basic PWM system used with Voltage Mode Control (Fig. 10 and 11), and with peak CMC (Fig. 14 and 15), the error signal applied to one side of the comparator is usually thought of as a dc level crossing over the sawtooth ramp, as shown in Fig. 11. This is only true if the open loop bandwidth, $f_C$, is extremely low — at least a factor of 10 below optimum. As the error amplifier gain is increased (and bandwidth along with it), the triangular inductor ripple current becomes evident at the output of the error amplifier. In Figure 17, where gain and bandwidth are optimum, the inductor ripple current (seen as $V_{CA}$) has become quite large. Optimum error amplifier gain is achieved when the slopes of the two waveforms coincide as shown in Figure 17 during the interval preceding the next clock pulse. In this case, it also happens to be the...
interval following switch turn-off when the inductor current is falling (the amplifier inverts the waveform).

Note that when the slopes coincide, the peaks also must coincide. Also, a perturbation applied to the VCA waveform is eliminated in the very first switching period, just like with optimum slope compensation with peak CMC (Ref. Fig. 15).

If the amplifier gain is increased beyond this optimum condition, two bad things happen:

1. The triangular waveform \( V_{CA} \) increases, making its positive peak exceed the positive peak of sawtooth \( V_S \). Depending upon the IC design, the E/A output may clamp \( V_{CA} \) at a voltage not much larger than the \( V_S \) peak. (The amplifier should be designed to clamp at this level. Otherwise during large signal events when the amplifier is “in the stops”, the E/A output would rise substantially, increasing the time required to recover from such an event.) If the waveform becomes clamped, the gain will suddenly appear to drop. Slope matching is consistent with the \( V_C \) waveform not exceeding the sawtooth \( V_S \).

2. Even if clamping does not occur, the increased triangular amplitude means the waveforms do not cross over or coincide after the switch turns off, and a tendency toward subharmonic instability begins.

It should be obvious that slope matching and slope compensation are closely related. In fact they are two sides of the same coin — the problems are identical, the optimization criteria are identical, and the benefits are identical. The only difference is that with Peak CMC, the triangular voltage representing inductor current is fixed, and a sawtooth compensating ramp is introduced whose magnitude is adjusted to obtain coincident slopes. With Average CMC, the sawtooth ramp is fixed, and the triangular voltage representing inductor current is adjusted (by varying the E/A gain) to obtain coincident slopes. In both systems, when the slopes are made to coincide, their crossover frequencies will not only be optimum, they will be the same.

**How to Implement Slope Matching:**

The inductor current downslope is translated into a voltage downslope by a current sense resistor, \( R_S \). The gain of the Current amplifier, \( CA \), (at the switching frequency \( f_S \)) is set so that the slope at \( CA \) output equals the ramp slope at the other input of the PWM comparator. For buck and boost topologies, the inductor current downslope is \( V_O/L \). The ramp slope is \( V_S/T_S \), or \( V_S f_S \). Therefore:

\[
G_{CA} = \frac{V_S f_S}{V_O R_S} \]

**Slope Matching with Voltage Mode:**

The slope-matching criteria for loop bandwidth optimization applies not only to the Average CMC loop, but to any system that uses a similar PWM technique. For example, the single-loop Voltage Mode Control described earlier benefits from the same strategy. With VMC, an electrolytic output capacitor appears resistive at \( f_S \), so the triangular inductor current waveform appears across the capacitor ESR, just as it does across the Average CMC current sense resistor. The voltage error amplifier gain is adjusted until its output slope coincides with the sawtooth ramp slope. The comparator waveforms look exactly like the Avg. CMC waveforms in Figure 17. The result is that, when optimized by slope matching, the lowly single-loop Voltage Mode Control not only has (a) the same crossover frequency as Current Mode Control,[3] the optimized VMC loop has (b) constant gain, independent of \( V_{IN} \). Even more importantly, the optimized VMC control loop (c) responds instantly to changes in \( V_{IN} \), just like CMC. The advantage of CMC remains that it is easier to implement, because the frequency dependent elements are apportioned between the two loops, thus are easier to deal with.

**Slope Matching Effect on PWM Gain:**

It was not recognized until recently that the optimized triangular waveform applied to the PWM comparator causes a change in the PWM gain characteristic. The relationship given in Eq. 3 is correct for low-gain, low-bandwidth loop whose amplified error signal appears as a dc level, as
shown in Fig. 11. But when the E/A gain is optimized, the slope of the initial portion of the triangular waveform, when the switch is ON, varies as a function of duty cycle. As shown in Figure 18, when the slope is relatively flat with D almost 1, an incremental change in the control voltage, $v_c$, causes a large incremental change in the duty cycle, d. When D is near zero, the initial slope is steep, so the same incremental $v_c$ change causes a much smaller change in d. By inspection, it can be seen that with slope-matched waveforms, the PWM “gain” is directly proportional to the duty cycle D:

$$d = \frac{v_c}{V_s} D$$

(11)

Whereas the PWM characteristic with a low-bandwidth “flat” control voltage from Eq. 3 does not change with D.

This modified PWM characteristic was not known at the time several earlier papers on Average CMC were written, and some of their gain expressions are in error. For example, with a Buck-derived regulator, duty cycle $D = \frac{V_o}{V_i}$. In Reference (2), “Average Current Mode Control of Switching Power Supplies,” Eq. (2), the expression for the power circuit plus PWM gain is:

$$\frac{v_{RS}}{v_{CA}} = \frac{R_S}{sL} \frac{V_O}{V_N}$$

Corrected Ref. (2), Eq. (2)

Replacing $V_N$ by $V_{OUT}$ may seem like a minor correction, but $V_N$ changes, $V_O$ is fixed. Thus in the optimized version, gain is constant, with the original version, gain varies directly with $V_N$.

Also, in Reference (2), Eq. (3) changes:

$$f_c = \frac{f_s}{2\pi D} \text{ becomes: } f_c = \frac{f_s}{2\pi}$$

(12)

With the loop gain and crossover frequency now constant and independent of $V_N$, closing the current loop and the outer voltage loop become much easier.

**Interaction in Two-Loop Systems:**

In a two-loop system, the inner current control loop determines the response to input voltage changes, while the outer voltage control loop determines response to load current change. These loops do interact, especially if their respective crossover frequencies are close to each other.

If both loops must be optimized for fast response, interaction is involved in the slope-matching process. There is only one PWM in a two-loop system. The triangular waveform $v_{CA}$ at the output of the current error amplifier CA actually has two components – the inductor ripple current seen across the current sense resistor and fed through CA, and the inductor current seen across the output capacitor ESR and fed through voltage error amplifier VA and CA. These two triangular waveforms are in-phase. VA and CA gains must be adjusted so the combined slopes match the saw-tooth waveform, but this can be accomplished in different ways. For example, in a buck regulator with a single loop optimized by slope matching, $f_s$ equals $f_c/2\pi$. But with two loops, if VA and CA gains are adjusted so that each loop contributes $1/2$ 

![Figure 18. – Modulator Gain vs. Duty Cycle](image)
of the total slope-matched triangular waveform, each will have \( f_S = \frac{f_C}{4\pi} \). However, if the current loop has more gain and the voltage loop less, the current loop contributes more than \( \frac{1}{2} \) of the total triangular waveform so its crossover frequency \( f_{CI} \) will be greater than \( f_{CV} \) of the voltage loop.

The closed-loop gain of the current loop is part of the open-loop gain of the voltage loop. The current loop closed-loop gain rolls off at its crossover frequency, \( f_{CI} \), adding an additional \(-1\) slope to the voltage loop above \( f_{CI} \). It is best to have \( f_{CV} \) below \( f_{CI} \) to minimize this interaction.

**The Right Half-Plane Zero:**

In the boost and flyback topologies, the output is driven through a diode, as shown in Figure 7. The inductor current flows to the output only when the power switch is off and the diode conducts. If load current increases, the duty cycle must be increased temporarily to make the inductor current rise. But operating in the continuous inductor current mode, when \( D \) is increased the diode conduction time decreases, before the slowly rising inductor current has time to change. The result is that the average diode current decreases at first, then as inductor current rises, the diode current ultimately reaches the proper value. This action, where the average diode current must actually decrease before it can finally increase, results in the small-signal phenomenon known as a right half-plane zero.[4]

\[
\omega_{RHPZ} = \frac{R_{LOAD}}{L} \frac{(1-D)^2}{D} \tag{13}
\]

A “normal” zero occurs in the left half of the complex s-plane, and has a gain characteristic that rises with frequency, with \( 90^\circ \) phase lead (+1 slope). The right half-plane zero also has a rising gain characteristic, but with a \( 90^\circ \) phase lag (-1 slope). This combination is almost impossible to compensate within the control loop, especially as the RHP zero frequency varies with load current. So most designers give up and cross over the voltage control loop below the lowest RHP zero frequency. One argument in favor of Average CMC is that it can operate in the discontinuous inductor current mode, which permits the use of a smaller inductor value. This not only saves size, weight and cost, it raises the RHP zero frequency to permit greater bandwidth for these topologies.

**Loop Design Procedure:**

Normally, the power circuit topology is decided upon and the power circuit values are determined, based on the application requirements, before control loop design begins. Occasionally, problems encountered in the control loop design process may force a rethinking of these power circuit decisions. The steps in the control loop design process will generally proceed as follows:

1. Define the control loop strategy and plot the tentative goal.
2. Plot the known part of the loop.
3. Define the crossover frequency, \( f_S \).
4. Try to meet the goal — Define and plot the error amplifier and overall loop characteristics.

Examples given in Appendix C should help to clarify this process.

**Step 1. Define the Control Loop Goal and Strategy:**

Based on application requirements for line and load regulation and transient response, output filter capacitor type. Define and crudely plot a tentative goal for the overall loop characteristic. The ideal goal is shown in Fig. 6a (two active poles below crossover, one above). Several strategies for practical situations are outlined below. Implementation is shown in Appendix C.

**Strategy #1 – The Easiest but not the Best:** For a buck-derived topology with aluminum electrolytic capacitor. Line and load variations are small and/or slow. Use single loop Voltage Mode Control. Cross over well below 1kHz, don’t worry about slope matching. The only problem to deal with is that the loop gain varies with \( V_{IN} \). The result of this short cut stabilization method is poor dynamic response, but if this is acceptable, who can argue.

**Strategy #2 — How to handle large step changes in load:** Output regulation in the face of a large step load change depends heavily on the output filter capacitor by itself, backed up by the
voltage control loop. In a two-loop system, the current loop does not provide any help in responding to a load change. In this situation with a continuous mode buck-derived topology, it will take several switching periods for the inductor current to slew to the new value (especially for a current rise at low $V_{IN}$). While the inductor current is slew-rate limited, the control circuit is non-functional because the amplifiers have been driven into their limits.

The salvation of this problem is an electrolytic output filter capacitor, especially an aluminum electrolytic whose C is huge because of the ESR requirement. The aluminum electrolytic does such a good job of “holding the fort”, that the voltage loop bandwidth does not need to be pressed to the limit. Thus, the current loop can be designed with slope matching for optimum $f_C$, and the voltage loop designed on a strictly linear basis to cross over at or below the capacitor ESR zero frequency.

Ceramic or polymer capacitors make a very poor showing with large rapid load changes — ESR is negligible but the C value used to achieve the desired output ripple is orders of magnitude less than an electrolytic. A really big help is to make the inductor smaller and the capacitor bigger – lower the surge impedance $\text{L}/\text{C}$. The smaller L can slew the current faster, the larger C will hold the fort longer. The increased ripple current will raise the minimum load where discontinuous operation begins, but Average CMC can cross the mode boundary nicely.

Strategy #3 – Large ESR Variation: An automotive application must operate over a wide temperature range and must have rapid response to input surges and load changes. Optimizing $f_C$ by slope matching, along with the input voltage feed-forward that slope matching provides, would provide a satisfactory solution. However, an ESR variation of 6:1 including initial distribution and temperature coefficient causes a 6:1 variation in loop gain and crossover frequency. The triangular ripple waveform which is the basis for slope matching varies by the same amount.

In this difficult situation, it is best to use Current Mode Control. The current loop does not contain the ESR, so it will be very stable and can be designed with slope matching to optimize bandwidth and input transient response. The voltage loop should then be designed to cross over at a lower frequency than the current loop. Then, the voltage loop will not significantly affect slope matching, and the roll-off of the closed current loop will be above the range of concern for the voltage loop. The voltage loop is definitely simplified, but the ESR is still there. If the variable ESR zero is in the vicinity of the desired $f_C$, it may be necessary to reduce $f_C$ to below the ESR zero frequency. The response of the voltage loop will not be excellent, but the aluminum electrolytic’s huge C value will probably handle this problem better than the best control loop, if the control loop gets knocked out of action by the inductor current slew rate.

The original single-loop VMC approach would be much more workable with Tantalum electrolytics, which have much smaller ESR temperature variation. If the frequency is high enough for economic viability, polymer capacitors might be worth considering.

As demonstrated above, many of the problems encountered while developing a control strategy lead back to the power circuit components or even to complete replacement of the original power circuit topology. This is to be expected, but this is clearly an area where experience can help to make the right choices the first time (or maybe the second time!).

Step 2. Plot the Known Part of the Loop:

After the power circuit topology and the control method have been at least tentatively defined, and the power circuit values established according to application requirements, Make a Bode plot of the entire loop but not including the error amplifier, $K_{EA}$. This plot must include the control-to-output characteristic plus feedback $K_{FB}$. The characteristic of the PWM, power circuit and filter must be known – see examples in Appendix C. In a two loop system, do the complete design of the inner loop first, before starting outer loop design.

Step 3. Define the Crossover Frequency:

If slope-matching to optimize $f_C$, the slope matching process defines the E/A gain at $f_C$. The crossover frequency will be optimum, but the spe-
specific frequency will not be know until the next step. If slope matching in a two-loop system, remember that each loop will contribute its share of the total slope. The relative share contributed by each loop determines the relative crossover frequency of each loop. It is best to have the current loop contribute most of the slope. This will result in current loop crossover frequency greater than the voltage loop which is desirable because the closed current loop is contained within the voltage loop.

If \( f_C \) is put at a frequency less than optimum to avoid problems, or just because there is no need for high bandwidth and fast response time in this application, then subharmonic instability will not occur, and loop stability can be totally handled with Bode plots. Steps 3 and 4 meld together. Again, in two-loop systems, there is loop interaction. The closed current loop within the voltage loop adds a pole to the voltage loop at the current loop crossover frequency, so it is best to have the current loop cross over at a higher frequency than the voltage loop.

**Step 4: Try to Meet the Goal — Define the E/A and Overall Loop:**

Since \( f_C \) has been determined, E/A gain at \( f_C \) is by definition the complement of the gain from step 2. Starting at \( f_C \), work up and down in frequency, combining E/A characteristic with gain from step 2, to obtain overall loop gain. Tailor E/A gain as needed to shape the overall loop gain characteristic working toward the ideal defined in Step 1. The examples in Appendix C should help explain this process.

**Error Amplifier Compensation Circuits:** Two circuit models given in Appendix A will handle most E/A compensation requirements. In most applications, fewer poles and zeros are required, and the circuit models can be simplified accordingly by omitting components. One of the two models has a current sense resistor input and is intended for an Average Current Mode Control loop. The “reference” voltage is actually the E/A output of the voltage loop, which sets the current level for the inner loop.

The other circuit model has a voltage divider input and is intended for use either as the outer voltage control loop of a two-loop CMC system, or as a single-loop Voltage Mode Controller.

Note that both circuit models include the feedback loop gain element, \( K_{FB} \), in addition to \( K_{EA} \), the E/A gain element. This is because with the voltage divider input, it is difficult to separate \( K_{FB} \) from \( K_{EA} \). The divider resistors in series form \( K_{FB} \), but their parallel combination forms all or part of the E/A input resistance, which determines \( K_{EA} \). The only problem this causes is mental – \( K_{FB} \) is part of the Step 2 Bode plot, \( K_{EA} \) is defined in Step 4.

**Problems Preventing Optimization:** There are many problems that can get in the way of achieving optimum \( f_C \). Such things as excessive ESR variation or excessive gain variation with \( V_{IN} \), or with RHP zero, or insufficient amplifier bandwidth can all add tremendous uncertainty in both gain and phase in the region near crossover, especially if several factors are at play. The effects of these variable elements must be examined at their extremes. Either the uncertainties must be reduced to manageable proportions, or \( f_C \) must be shifted to a much lower frequency. But some of these problems can be reduced or eliminating by making different choices, including rethinking some of the decisions made regarding the power circuit:

If the RHP zero is a problem in a continuous mode boost or flyback circuit, making \( L \) smaller raises the RHP zero frequency. If the smaller inductor means crossing into the discontinuous mode at light loads, where peak CMC or VMC falls apart because of the large drop in loop gain (which changes the crossover frequency!), consider using Average CMC which adds enough gain to make discontinuous operation feasible. And the smaller inductor cost less. The penalty is increase ripple and noise.

ESR variation over a wide temperature range with aluminum electrolytics is a tough problem to get around. Tantalum capacitors have much less variation, but they cost a lot more. Polymer aluminum electrolytics have no ESR, but limited capacitance and low voltage rating make them unsuitable for most applications.

Gain variation due to wide swings in \( V_{IN} \) can be eliminated using peak or average CMC.
Insufficient Amplifier Bandwidth: As switching frequencies rise, error amplifier bandwidth may not be sufficient for slope matching or optimization of \( f_C \). If the amplifier bandwidth is not enough for the desired compensation scheme, there are some alternatives other than backing down on the crossover frequency: (1) use an IC with a better amplifier. (2) In the current loop, use a larger current sense resistor or a current transformer with a larger turns ratio. Two cascaded amplifiers can provide a very large gain increase at frequencies well below their crossover frequencies.

Control Problems your Mother Never Told You About

A tremendous amount of effort has been put into the development of small-signal techniques and linear models of the various switching power supply topologies. Hundreds, if not thousands of papers have been written over the years. Your academic “mother”, whoever “he” may be (note the PC sexual ambiguity), typically focuses on new topologies and/or linear modeling.

While not disparaging any of these efforts – far from it, these contributions have been immense and totally necessary – there has been a lack of balance and a tendency to try to force behavior that is uniquely related to switching phenomena into linear equivalent models (with sometimes uncertain results). Many of the major significant problems with switching power supplies do not show up in the frequency domain, or in the time domain using averaged models, unless these problems are anticipated in advance and provided for in the models. Simulation in the time domain using switched models, although slower, reveals these problems that would have been hidden:

- Modulator gain, \( \frac{dV_C}{dt} \), varies with duty cycle \( D \) when E/A gain is adjusted to optimize \( f_C \). This makes buck regulator gain independent of \( V_{IN} \) provides input voltage feed-forward (Fig. 16). This is a geometry problem dealing with the ripple waveform at the E/A output.
- Subharmonic instability and the slope compensation / slope matching solution.
- Leakage inductance leading edge delay causes dc cross-regulation problems.

Large signal problems involve changes that are so large or so rapid that the control loop cannot keep up. Error amplifier outputs are driven to their limits, and the loop(s) become temporarily open. Large signal events include: start-up, input voltage drop-outs, rapid input voltage changes, rapid load current changes.

All energy storage elements within the loop are likely to either become the cause of large signal problems, or to behave badly as a result. This includes not only the filter inductor and filter capacitor, but even the small compensation capacitors around the error amplifier.

- The inductor is the main cause of large signal problems, because of its limited ability to slew the current rapidly to accommodate a large, rapid load change, or during start-up or after a line voltage drop-out. In a buck regulator with increasing current demand, \( \frac{di}{dt} = (V_{IN}D - V_O) \). If min \( V_{IN} \) times max \( D \) is only marginally greater than \( V_O \), it will take forever for the inductor current to rise. Even if the loop bandwidth is 1MHz — The loop is open! Under normal operating conditions, it will still take several switching periods. Once the inductor slewed to the proper current, the filter capacitor, whose voltage has sagged, takes more time to recharge, further delaying loop recovery. Soft start is helpful only during start-up. A smaller inductor certainly helps at the expense of greater noise and output filtering. If the inductor is small enough for discontinuous operation, the slew rate problem disappears.
- A unique overshoot problem can occur at startup unless soft start is used. Without soft-start, \( L \) and \( C \) will start to charge resonantly toward \( 2xV_{IN} \), but as soon as the current limit is reached, inductor current stabilizes at this value. The capacitor voltage now rises linearly toward the desired \( V_{OUT} \). But the inductor current is at the current limit, and if the load current happens to be minimal, there is way too much current. It takes time for the inductor current to slew back down to the load current demand. During this time the capacitor voltage keeps rising, above the required \( V_{OUT} \) value. The
overshoot is probably only a few percent with the huge C value of aluminum electrolytics, but with the much smaller C values that would be used with ceramic or polymer capacitors (if frequency is high enough to make this viable) the overshoot can be quite large – 30-50% — requiring soft-start to prevent this from happening. A lower L/C ratio helps here, as well.

- It may be tempting to add a zero to the error amplifier to boost low frequency gain and improve accuracy. Even though the resulting small signal plot appears optimum, adding this capacitor in the E/A feedback can hurt more than it helps. If a situation arises where a rapid load current increase causes inductor current to become slew-rate limited, the power supply output will sag, and the E/A output will be driven into its positive limit. The feedback loop is temporarily non-functional. The compensation capacitor will charge to an abnormal voltage which later will delay recovery of the loop to normal operation. The lower the frequency of the pole or zero involving the compensation capacitor, the longer it will take to recover.

Many IC’s in wide use today have error amplifiers whose outputs can swing from 0 to +VCC. If the sawtooth voltage against which the E/A output will be compared ramps from 1 to 4 Volts, what is the virtue of allowing the E/A output to swing to 18 V whenever the E/A is temporarily driven into its limits by a large signal event? IC designers should include clamps from the E/A output to its input to prevent the output from being driven significantly beyond the useful range. Not only might this hasten recovery of the amplifier itself, the amplifier input will always remain at its normal operational level, and external feedback capacitors will not charge to abnormal voltages and thus will not delay recovery from large signal events.

It may seem paradoxical, but the VOUT tolerance band can be cut in half by reducing the voltage loop gain. Figure 19 shows the output voltage waveforms that result when the load current changes suddenly and then, at some later time, changes back. The magnitude of VPK is a function of fC and the loop gain at high frequency. VSS, the steady-state voltage deviation or error from light to heavy load, is a function of the loop gain at low frequency. Fig. 19a demonstrates what happens with very high low frequency gain resulting in minimal steady state error, VOUT starts near the nominal value and returns there following each load change. Thus, the total swing is twice the peak value, exceeding the permissible tolerance band. Figure 19c shows what happens with the high frequency loop gain unchanged, but with low frequency gain reduced to the amount necessary to result in VSS much larger but within the tolerance band. The initial shape and amplitude is the same in Fig 19a and 19b, but the voltage never returns to nominal because of the deliberately large dc error. The required gain is easy to calculate, especially with current mode control inner loop. If the current sense resistor is .02 Ω, then 0.2 V E/A output swing is required for a 10 A current

Figure 19. – VOUT Tolerance
change. If the desired $V_{OUT}$ swing is 0.1 V (within a 0.15 V tolerance band), then the E/A gain must be set at $0.2V/0.1V = \text{a gain of } 2.0$. In Strategy #2 above, this technique would be helpful.

**Triangle vs. Sawtooth PWM Waveform:**

In a fixed frequency PWM using a sawtooth waveform, a switching decision is made once per switching period based on the control signal level at that instant when the decision is made. A second switching action is taken at the clock pulse at the beginning of the sawtooth ramp, but this is not influenced by the control signal. Thus, the duty cycle is modulated according to a single control signal sample per switching cycle. In order for the duty cycle to be modulated effectively by a small AC signal, it is obvious that a minimum 2 samples must be taken during the period of the signal in order to define its amplitude. From this point of view, the highest signal frequency that can pass through the PWM is one-half of the switching (sampling) frequency.

With a triangular waveform, decisions to switch ON and OFF are each made on the basis of separate intersections of the control signal vs. the triangular waveform. The argument in favor of the triangular waveform is that since two control signal samples are taken per switching period, the PWM should be able to handle twice small signal frequency as the sawtooth PWM. Therefore a higher crossover frequency, greater bandwidth and improved performance should be attainable.

Some counter arguments are: Although there are two decisions per switching period, they affect only one power pulse per period. Also, the two points of decision converge upon each other at duty cycle extremes. However, the real limitation on crossover frequency is not related to the number of samples taken. The real limitation on $I_C$ is the subharmonic instability which start to occur when the slopes of the inductor ripple current waveform seen at the error amplifier output exceed the slopes of the sawtooth (or triangular) waveform at the other comparator input. This slope-matching criteria limits the E/A gain, and thereby limits the gain and the crossover frequency of the entire loop. In this regard, the triangular waveform is no better and perhaps worse than the sawtooth. Certainly it is more difficult to implement. Slope matching is much more difficult to optimize when there are two slopes to consider. The PWM comparator must latch in both states, and unlatch at each subsequent peak of the triangular waveform, otherwise the comparator will false trigger on the tiniest noise pulse and will not function at all if the control signal slope exceeds the triangular waveform at any point.

**REFERENCES:**

IMPORTANT NOTICE

Texas Instruments and its subsidiaries (TI) reserve the right to make changes to their products or to discontinue any product or service without notice, and advise customers to obtain the latest version of relevant information to verify, before placing orders, that information being relied on is current and complete. All products are sold subject to the terms and conditions of sale supplied at the time of order acknowledgment, including those pertaining to warranty, patent infringement, and limitation of liability.

TI warrants performance of its products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are utilized to the extent TI deems necessary to support this warranty. Specific testing of all parameters of each device is not necessarily performed, except those mandated by government requirements.

Customers are responsible for their applications using TI components.

In order to minimize risks associated with the customer’s applications, adequate design and operating safeguards must be provided by the customer to minimize inherent or procedural hazards.

TI assumes no liability for applications assistance or customer product design. TI does not warrant or represent that any license, either express or implied, is granted under any patent right, copyright, mask work right, or other intellectual property right of TI covering or relating to any combination, machine, or process in which such products or services might be or are used. TI's publication of information regarding any third party's products or services does not constitute TI's approval, license, warranty or endorsement thereof.

Reproduction of information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations and notices. Representation or reproduction of this information with alteration voids all warranties provided for an associated TI product or service, is an unfair and deceptive business practice, and TI is not responsible nor liable for any such use.

Resale of TI’s products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties for the associated TI product or service, is an unfair and deceptive business practice, and TI is not responsible nor liable for any such use.

Also see: Standard Terms and Conditions of Sale for Semiconductor Products, www.ti.com/sc/docs/stdterms.htm

Mailing Address:

Texas Instruments
Post Office Box 655303
Dallas, Texas 75265

Copyright © 2001, Texas Instruments Incorporated