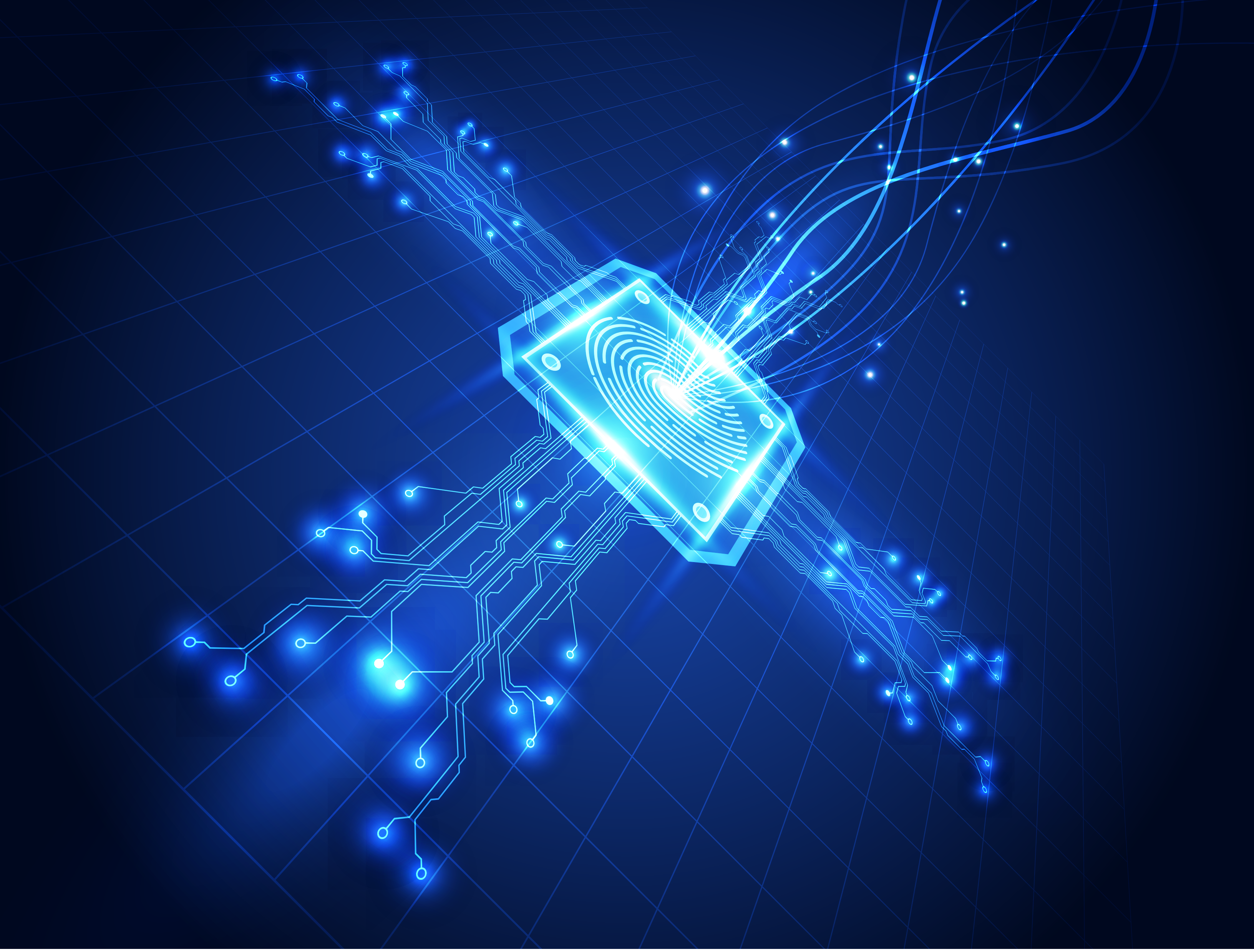

OPTIMUM FEEDBACK AMPLIFIER DESIGN FOR CONTROL SYSTEMS

H. Dean Venable ● Venable Instruments

www.venableinstruments.com

VENABLE
instruments

VENABLE TECHNICAL PAPER # 3

Optimum Feedback Amplifier Design For Control Systems

Abstract:

Many designers think of control system feedback loop compensation only in terms of stable or unstable. In reality, there is an optimum error amplifier transfer function for any given loop crossover frequency and phase margin. The optimum design is shown to be the one with the highest gain below crossover and the lowest gain above crossover. There is a unique pole-zero placement to achieve optimum performance. This pole-zero placement and resultant circuit component values can be determined without trail-and-error. In addition, there is an optimum open loop crossover frequency (loop bandwidth) or any given phase margin.

Optimizing the Performance

Optimizing the performance of a feedback control loop can be done with techniques that are simple, effective, and easy to apply. These techniques can be used on single or multiple loop systems of any type, and will yield optimum results without the usual trial-and-error process.

Three new concepts are introduced: the concept of phase boost as the single variable of importance in loop stability, K-Factor as a convenient method of defining both phase boost and the shape of the Bode gain curve, and the concept of a Figure of Merit for evaluating the relative performance of various loop compensations.

Stability Criteria

A feedback loop will oscillate when there is a frequency at which the loop gain is unity and total phase lag equals 360°. Stability is usually measured by two factors: Phase margin, which is the difference between actual phase lag and 360° when the loop gain is unity (usually expressed in degrees); and gain margin, which is the amount the gain has fallen below unity when the total phase lag is 360° (usually expressed in dB).

A Bode plot is a plot of open loop gain and phase shift as a function of frequency, and is a convenient method of displaying loop characteristics such as gain and phase margins. Figure 1 shows a typical loop Bode plot and the resultant gain and phase margins.

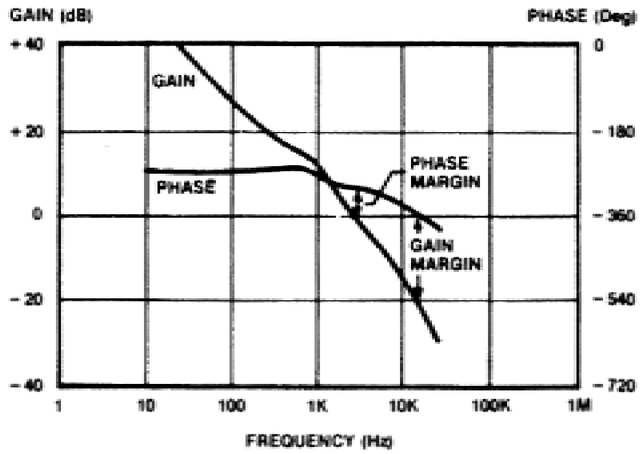

Figure 1. Typical Overall Loop Plot

Bode plots are easy to interpret when the 0 dB (unity gain) and 0° (or 360°) axes are coincident, and a 1 slope is a 45° line (i.e., the dimension which represents 20 dB of gain is the same as the dimension which represents one decade of frequency.

In classical theory, the signal polarity definition is reversed at the output of the point where the loop is broken for analysis, as shown in Figure 2. While this made life a little easier for mathematicians by making loop gain a positive number, it confuses the subject somewhat, since oscillation is discussed in terms of 180° of phase shift. In reality, oscillation occurs when the gain is unity and the signal is actually in phase (shifted by a total of 360°), therefore this paper discusses oscillation in terms of 360°.

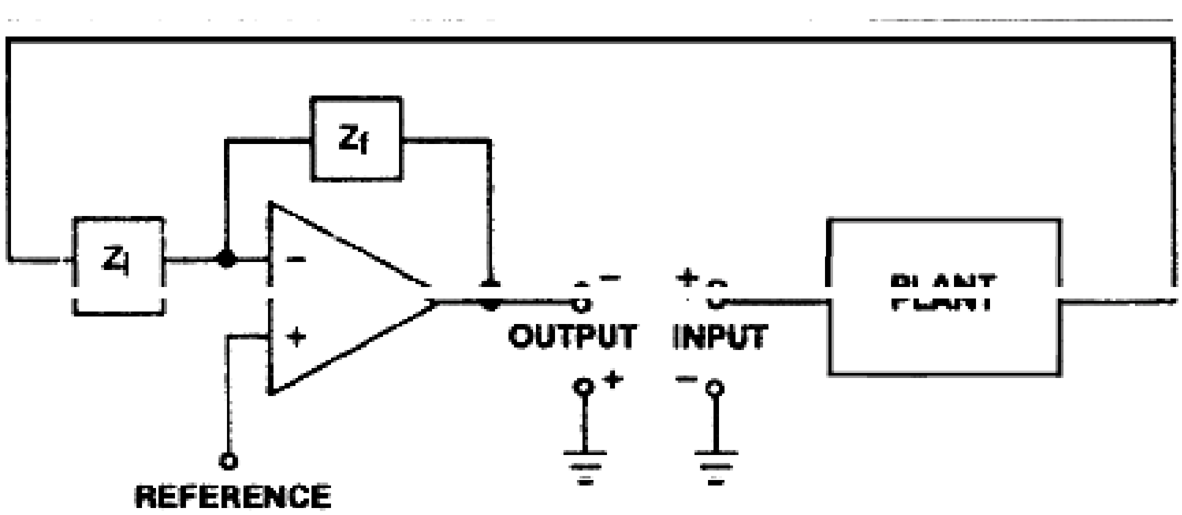

Figure 2. Classical Signal Polarity Definition

The same problem occurs for Nyquist diagrams, where the –1 point is classically defined as the critical point for stability, corresponding to 180° of phase shift at unity gain. Classical Nyquist diagrams also suffer from being plots of linear gain versus polar phase, because linear gain does not have sufficient range to be useful for most actual measurements. A far superior presentation of a Nyquist diagram has the critical point at +1, corresponding to the 360° of phase lag encountered in actual practice, and has gain plotted as a linear function inside the unity gain circle and as 1 + log (gain) outside the unity gain circle. This technique preserves the graphic benefits of a Nyquist diagram while allowing a wide range of real data to be presented.

How to Stabilize a Loop

Almost any loop can be divided into two major portions. One is the power processing portion, which takes a control signal as an input and outputs the variable to be controlled, which may be voltage, current, torque, speed, position, temperature, etc. This portion is typically called the modulator or plant. The other portion is the error amplifier, which compares a sample of the controlled output to a reference, amplifies the difference, and outputs a control signal to the modulator or plant. This portion is usually called the amplifier. Figure 3 shows the definition of these two blocks in a typical switch-mode power supply.

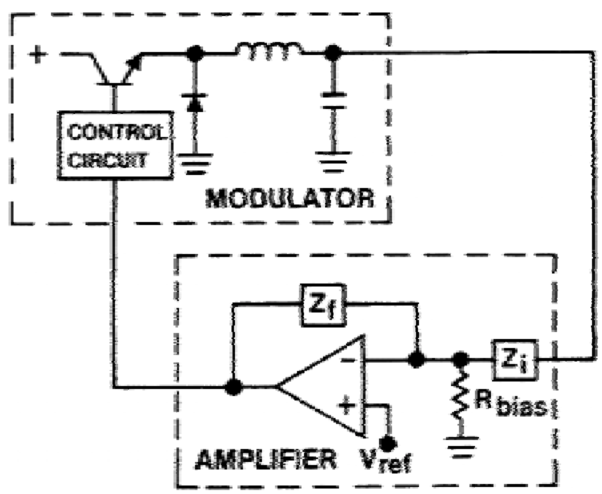

Figure 3. Typical Feedback Control Loop

Even in a multi-loop system the process is the same. The principle difference is that the reference usually comes from an outer loop. For example, in a positioning system, there may be a motor current loop (controlling acceleration), which forms a part of a velocity loop (controlling motor speed), which in turn is part of the outer loop of position, which has desired position as the reference. An error in position calls for more or less speed, and an error in speed calls for more or less acceleration (current). Each loop must be stabilized, or preferably optimized, starting with the innermost loop.

Measure the Transfer Function of the Plant

The first step in stabilizing a feedback loop is to determine the transfer function (Bode plot) of the modulator or plant. This can be done by measurement or analysis, but measurement is by far preferable since most of the hard-to-model parasitic elements are in the plant. A bode plot of a typical plant is shown in Figure 4.

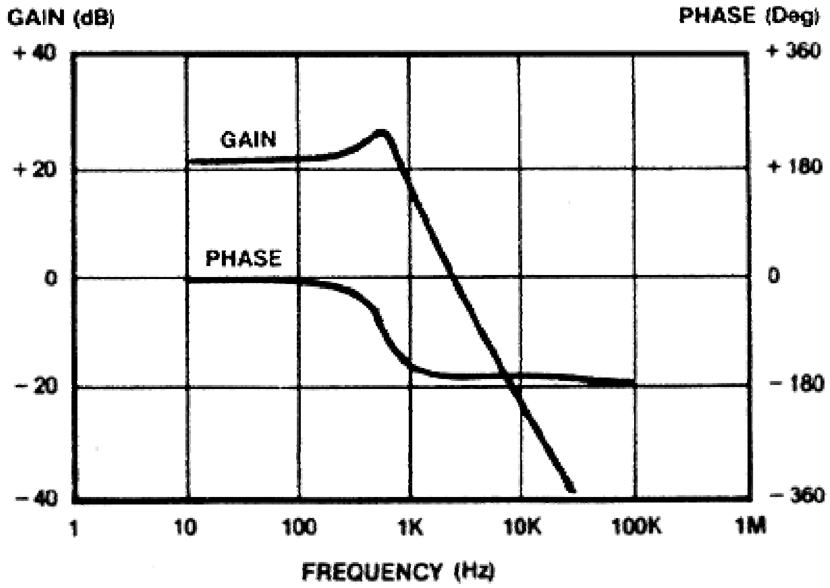

Figure 4. Typical Plant Bode Plot

Choose Crossover Frequency and Phase Margin

The next step is to choose the overall loop unity gain frequency and desired phase margin. The unity gain frequency can be chosen on the basis of desired performance, but a better method is to determine the optimum based on maximizing the Figure of Merit of the loop design. Figure of Merit is discussed below in the section on optimizing performance. The other parameter to choose is the phase margin desired in the overall loop. Typical phase margins range from 45° to 75°. Lower margins, like 45°, give good transient response at the expense of peaking of the closed-loop transfer function and output impedance. Higher margins, like 75°, give flat closed-loop transfer functions and minimum peaking of output impedance, but at the expense of speed and settling time. A good compromise is 60.

Synthesize the Error Amplifier Compensation

The final step is to synthesize or design error amplifier compensation that has gain equal to the reciprocal of the modulator or plant gain at the desired crossover frequency, and phase lag such that the sum of the modulator and amplifier phase lags plus the desired phase margin equals 360°. The other amplifier design factors to consider are to maximize the loop gain at frequencies below crossover (for better transient response and disturbance rejection), and to minimize the loop gain at frequencies higher than crossover (to improve system noise immunity). If all of these considerations are met simultaneously, optimum overall performance is the result. While meeting all these criteria may seem difficult, it is in fact surprisingly easy. A step-by-step guide is given below.

Optimum Error Amplifier Design

Since optimum performance is obtained from maximizing low frequency gain and minimizing high frequency gain, it stands to reason that an integrator is the logical starting choice for an error amplifier. The transfer function of this type of amplifier has one pole (at the origin), so for reference purposes we can call this a Type 1 amplifier. (Please do not confuse amplifier-type numbers with servo system-type numbers, which are not related.)

The transfer function of a Type 1 amplifier falls at a –1 slope at all frequencies (–20 dB per decade). The output lags the input by 270°, 180° from the inversion, and the other 90° from the pole at the origin. The schematic of a Type 1 amplifier is shown in Figure 5 and its transfer function is shown in Figure 6.

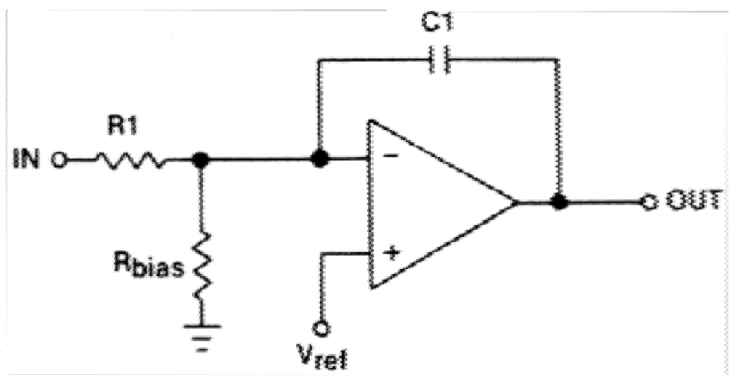

Figure 5. Schematic Diagram of a Type 1 Amplifier

The Concept of Phase Boost

Since phase lag of a Type 1 amplifier is 270° at all frequencies, no opportunity exists to control the amount of phase lag through the amplifier. By placing a zero-pole pair in the transfer function however, the phase lag can be reduced to less than 270° over some range of frequencies. A zero in a transfer function causes the slope of the gain curve of a Bode plot to break upward with increasing frequency (to a steeper up-slope or less steep down-slope), and is accompanied by a reduction of phase lag of 45° at the corner, and up to 90° at frequencies much higher than the frequency at which the zero occurs. A pole in the transfer function has the opposite effect, causing the slope of the gain curve of a Bode plot to break downward (to a less steep up-slope or steeper down-slope), and is accompanied by an increase in phase lag of 45° at the corner, and up to 90° at frequencies much higher than the frequency at which the pole occurs. A zero-pole pair introduced into the transfer function of an integrator creates a region of frequencies in which the Bode gain plot goes flat and the phase lag is reduced by an amount which is related to the frequency spread between the zero and pole. This reduction in lag through the amplifier can be viewed as a phase boost, and this concept of phase boost is very powerful in understanding the nature of loop compensation. The amount of phase boost can be varied easily by adjusting the spread of the zero-pole pair. Since the sum of the modulator phase lag, amplifier phase lag, and phase margin must equal 360° at the loop unity gain frequency, any desired phase margin can be obtained simply by adjusting the spread of the zero-pole pair.

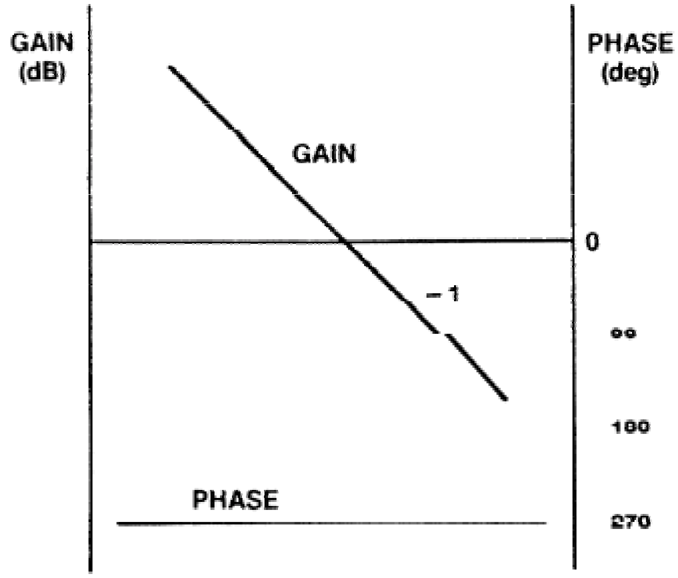

Figure 6. Transfer Function of a Type 1 Amplifier

Type 2 Amplifier

An implementation of an error amplifier with a zero-pole pair is shown in Figure 7, and the transfer function of this amplifier is shown in Figure 8. Because this amplifier has two poles, one at the origin and one from the zero-pole pair, for reference we can call this a Type 2 amplifier. Since there is only one zero, the maximum boost that can be obtained from a Type 2 amplifier is 90°. In practice, the spread between the zero and pole is adjusted to give the desired phase boost, and the zero and pole are placed symmetrically around the desired loop crossover frequency, which places the peak of the phase boost at crossover. The amplifier gain at this frequency, which is roughly the ratio of R2 to R1, is the reciprocal of the modulator or plant gain.

Type 2 amplifier characteristics are frequently referred to in the literature as "proportional plus integral compensation." While this is technically correct, since there are frequency regions where the error amplifier is acting like an integrator and a frequency region where the error amplifier is acting as a proportional amplifier, this terminology obscures the true reason for the gain characteristic — which is to create phase boost (actually a reduction in lag) in the error amplifier transfer function in the region of gain crossover to compensate for lag in the plant or modulator portion of the loop.

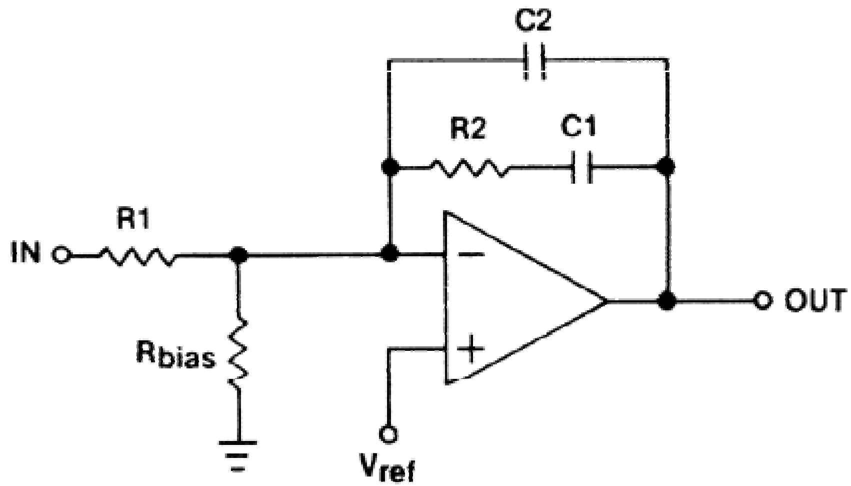

Figure 7. Schematic Diagram of a Type 2 Amplifier

The schematic shown in Figure 7 is optimum in every way. The resistor Rbias can be used to freely adjust the steady-state value of the output, yet its value has no effect on amplifier gain. R2 rather than C1 is connected to the summing node of the amplifier, so that any pickup at the R2-C1 junction is shunted into the low impedance output of the amplifier rather than the sensitive summing node. Finally, the capacitor C2 spans the R2-C1 combination rather than being placed across R2, so that the high-frequency currents it is intended to carry do not have to pass through C1 as well, which is typically much larger than C2 and may not be as good a high-frequency capacitor.

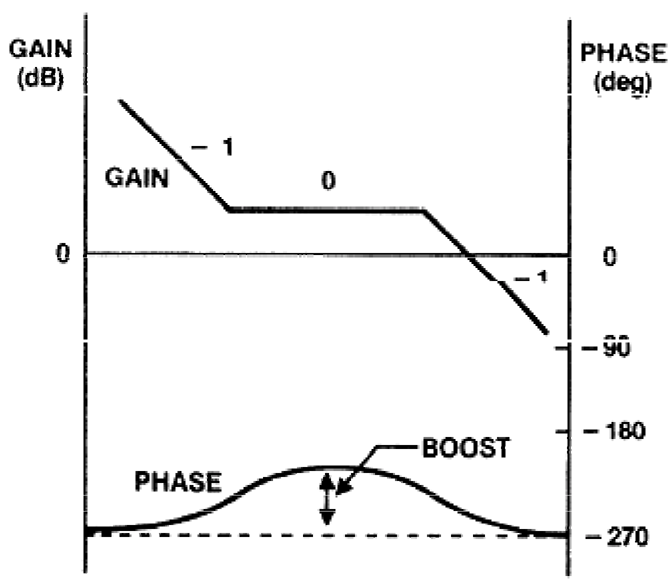

Figure 8. Transfer Function of a Type 2 Amplifier

Type 3 Amplifier

There are many instances when 90° of phase boost is not enough to compensate for the phase lag through the modulator or plant. For cases where there are two reactive elements in the plant, for example, a mechanical system driving a mass through a spring, or an electrical system with an L-C filter, the phase lag through the plant can approach 180°, and substantial boost is required in the amplifier to maintain a reasonable phase margin.

As described above, a zero can provide phase boost up to 90°. This concept can be expanded to two zeroes, each providing up to 90°of boost for a total of 180°. An implementation of this concept is shown in Figure 9, with the resultant transfer function shown in Figure 10. The addition of two extra components, C3 and R3, allow a range of frequencies where the gain is increasing at a +1 slope (+20 dB per decade), and a larger phase boost, which can be as high as 180°. There are two coincident zeroes and two coincident poles, again placed symmetrically around crossover.

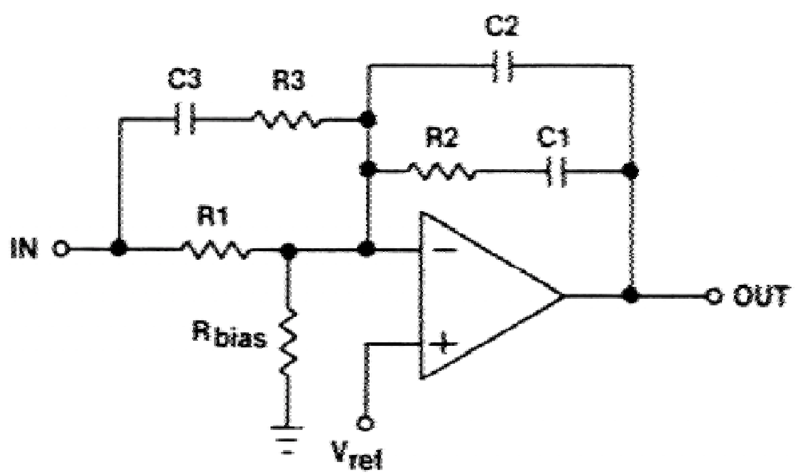

Figure 9. Schematic Diagram of a Type 3 Amplifier

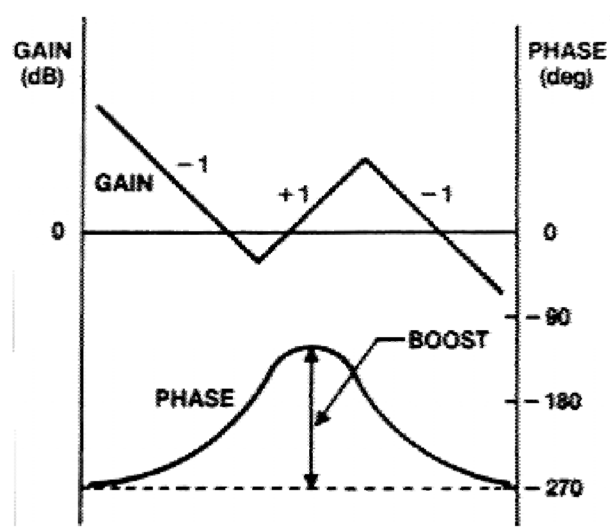

Figure 10. Transfer function of a Type 3 Amplifier

As with the Type 2 amplifier, the schematic shown in Figure 9 is optimum in every way. Rbias has no effect on AC gain of the amplifier. The R-C series combinations are arranged so that any pickup at the connection nodes is shunted to relatively insensitive points rather than the error amplifier summing node. High frequency currents through C2 do not have to pass through C1. And the final optimum feature unique to Type 3 amplifiers is that the DC input current does not have to flow through R3. (An alternative circuit arrangement with the same transfer function has C3 in parallel with R1, and R3 in series with this combination.) Because input current does not flow through R3, its value does not affect the steady-state value of the output. The output value is determined only by R1, Rbias, the constantness of the reference, and the bias current and offset voltage of the op amp.

Coincident Zero-Pole Pair Placement

For Type 2 amplifiers, there is little question about placement of the zero-pole pair. Required phase boost sets the spread, and the zero and pole are symmetrically placed on either side of desired unity gain frequency of the overall loop. For a Type 3 amplifier, optimum zero-pole placement is not so obvious. To understand why the two zeroes need to be at the same frequency and the two poles need to be at the same frequency, consider the Bode gain plot of Figure 10 and the optimum design consideration that the low frequency gain should be as high as possible. If the two zeroes or the two poles are spread apart in frequency, the phase boost curve in Figure 10 becomes wider and the amount of phase boost is reduced. To obtain the same amount of phase boost (required to maintain phase margin), the two –1 gain slope regions must be spread farther apart, reducing the low frequency gain of the amplifier and increasing the high frequency gain. Since the overall loop gain is the product of the amplifier gain and plant or modulator gain, the overall loop gain will be affected in the same way.

The K-Factor Concept

A convenient mathematical tool for defining the shape and characteristics of a transfer function is the K-Factor. When a flat or +1 gain slope region is introduced into an amplifier transfer function to reduce phase lag, there is a corresponding reduction in low frequency gain and increase in high frequency gain. The reduction at low frequency is equal to the increase at high frequency; by a factor we can call K. Because the Bode gain plot is logarithmic in nature, K also controls the location of the zero and pole. In a Type 1 amplifier, K is always 1 because there is no boost or corresponding decrease or increase in gain. In a Type 2 amplifier, the zero is placed a factor of K below loop crossover, and the pole is place a factor of K above. In the case of a Type 3 amplifier, the two zeroes are placed below crossover a factor of the square root of K, and the two poles above crossover by the same amount.

Figure 11 shows the Bode gain plot of a Type 1 amplifier designed to have gain G at frequency f. A system stabilized with this amplifier would be designed to have a loop crossover frequency of f, and the modulator or plant gain at that frequency would be 1/G. If phase boost were required in the amplifier to compensate for lag in the modulator or plant, a Type 2 or Type 3 amplifier would be used instead of a Type 1. The transfer functions of these two amplifiers are shown in Figures 12 and 13, respectively, with the same gain G at frequency f, which would yield the same loop crossover frequency when used to compensate the same modulator or plant.

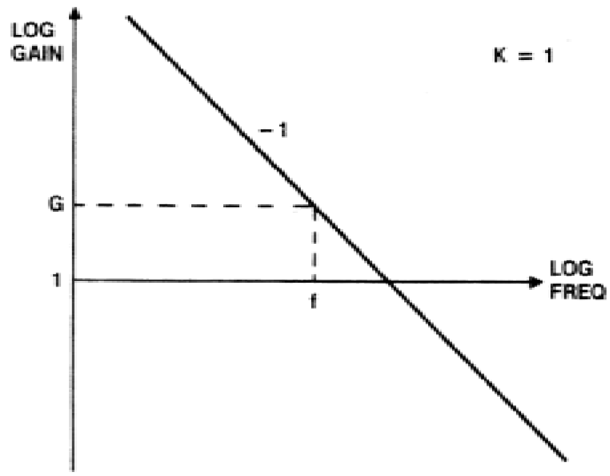

Figure 11. Type 1 Amplifier

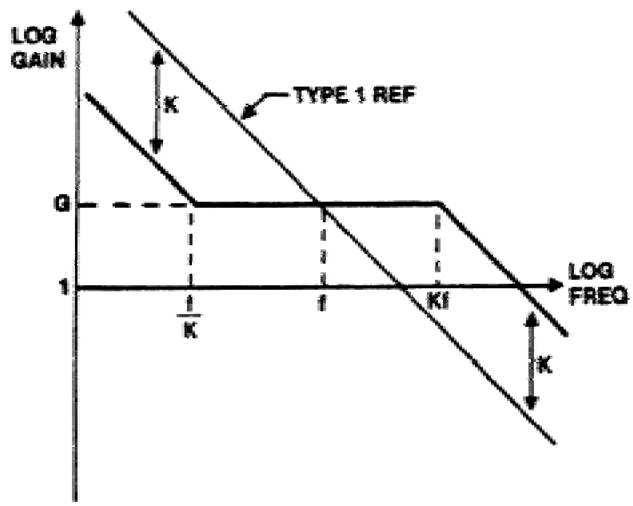

Figure 12. Type 2 Amplifier

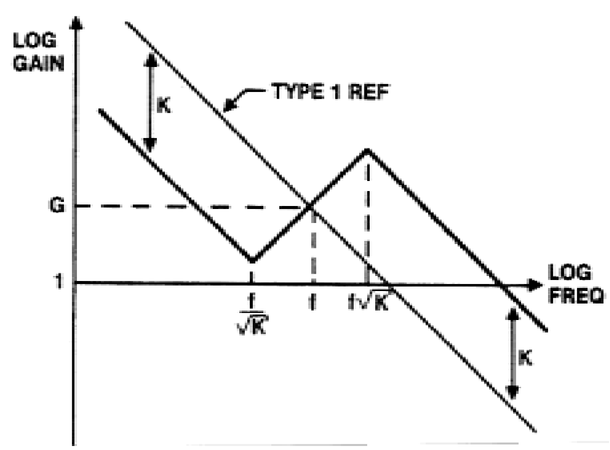

Figure 13. Type 3 Amplifier

K as a Function of Boost

The expression for the amount of phase boost from a zero-pole pair is given by the inverse tangent of the ratio of the measurement frequency to the frequency at which the zero or pole is located. The principle of superposition applies, that is, the total amount of phase shift can be determined by summing the individual phase shifts of each zero and pole taken individually. The boost at frequency f from a zero at frequency f/K and a pole at frequency Kf is given by the equation

(1) Boost = Tan–1(K) – Tan–1(1 / K)

From equation (1), a plot of phase boost as a function of K for Type 2 and Type 3 amplifiers can be computed, which yields a graph like that shown in Figure 14.

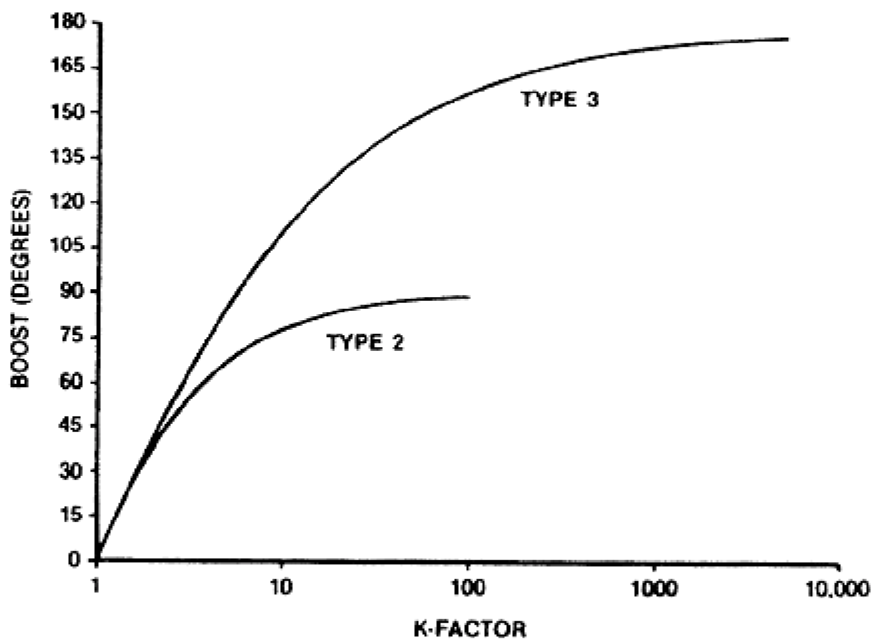

Figure 14. Phase Boost as a Function of K-Factor

From desired boost, K-Factor can be determined. From K-Factor, the location of the zeroes and poles of the transfer function can be determined. And from location of poles and zeroes, the circuit values can be determined. In order to calculate circuit values, one must know the location of the poles and zeroes of the transfer function of Type 2 and Type 3 amplifiers as a function of the component values. There is a pole at the origin that causes the initial –1 gain slope region. The frequency at which this line crosses (or would have crossed) unity gain is called the unity gain frequency, or UGF. The other terms in the equations below are obvious.

Type 2 Amplifier Components

The components of a Type 2 amplifier can be calculated from the following equations:

(2) UGF = 1 / (2 pi R1 (C1 + C2))

(3) Zero = 1 / (2 pi R2C1)

(4) Pole = 1 / (2 pi R2 (C1C2 / (C1 + C2)))

Type 3 Amplifier Components

The components of a Type 3 amplifier can be calculated from the following equations:

(5) UGF = 1 / (2 pi R1 (C1 + C2))

(6) Zero = 1 / (2 pi R2C1)

(7) Zero = 1 / (2 pi (R1 + R3) C3)

(8) Pole = 1 / (2 pi R3C3)

(9) Pole = 1 / (2 pi R2 (C1C2 / (C1 + C2)))

From the above equations, it is relatively easy to calculate the component values required to achieve desired system performance.

fG/K as a Figure of Merit

In actual practice, most designers want the performance of their feedback loops to be as good as possible, where "good" means minimum response time with a reasonably small amount of overshoot and ringing. This can be systematically achieved using the techniques described above by maximizing the value of the expression fG/K, where f is the desired overall loop crossover frequency, G is the gain required in the error amplifier at that frequency to achieve crossover, and K is the variable determining the shape of the Bode gain curve as described above.

When pushing a system to the limits of its performance, there are usually factors causing large increases in phase lag of the plant or modulator at frequencies not far above the maximum achievable loop crossover frequency. A Type 3 amplifier gives the most phase boost, but large amounts of boost (required to offset large amounts of lag in the plant) require large K-Factors, and this depresses the low frequency gain of the feedback loop. Increasing the crossover frequency and increasing the gain of the amplifier (assuming the transfer function of the plant is fixed) both cause improved system performance, while increasing K-Factor simply to maintain a given phase margin decreases system performance. If, with increasing crossover frequency, the product fG increases faster than the K required to maintain the same phase margin, then performance is enhanced. At some frequency there will be a peak of fG/K, above which K increases faster than the fG product. The frequency at which fG/K peaks is the optimum crossover frequency for that particular loop given the desired phase margin as a goal. The value of fG/K can then be used as a Figure of Merit to evaluate the relative performance of various choices of crossover frequency.

Conclusion

A simple, straightforward method exists for achieving optimum control system performance without trial-and-error. A new concept, the K-Factor, can be used to conveniently describe the shape of the transfer function of any error amplifier. K-Factor can also be used to calculate actual component values to use for system frequency response compensation. And finally, K-Factor can be used in a Figure of Merit calculation to determine the optimum bandwidth for any control system.