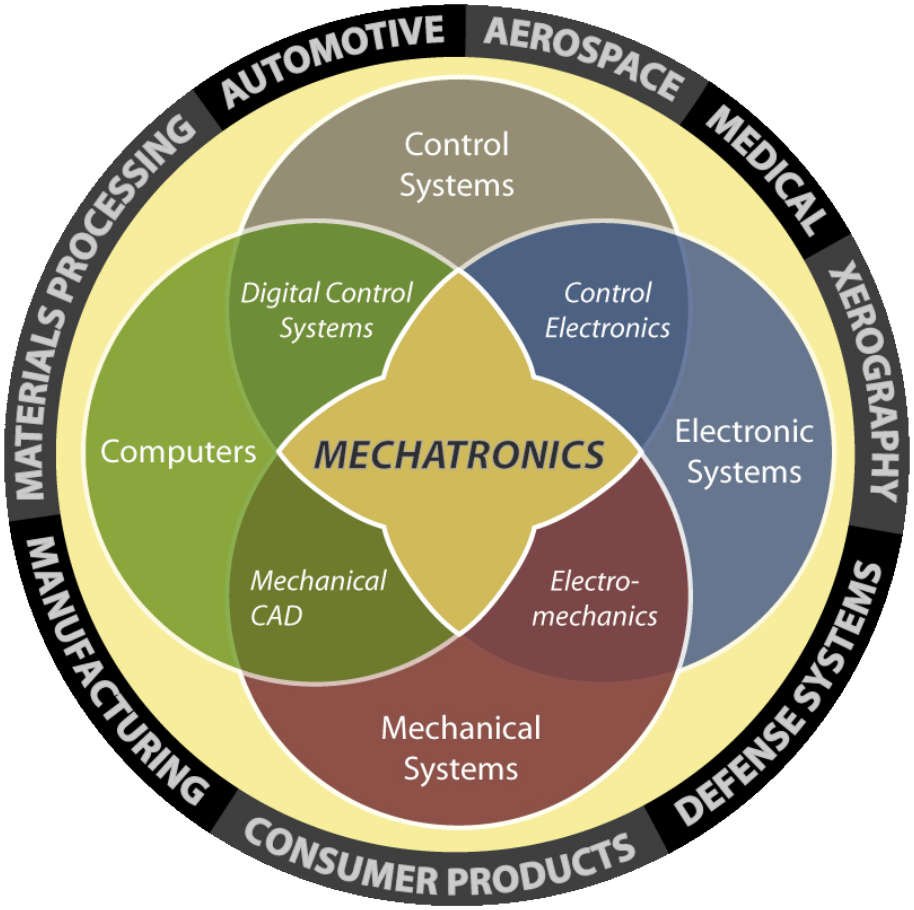
Tuning Controllers & Types of Industrial Controllers



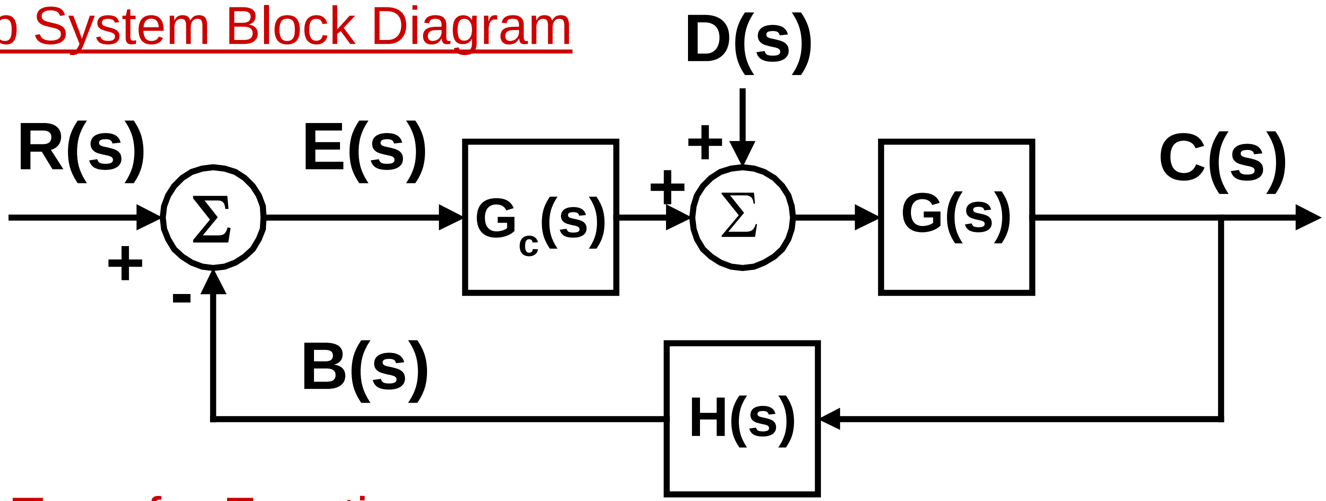
Tuning a Control System

- Tuning is the process of setting controller gains to optimize system performance.
  - It is a difficult task, as control systems have inherent limitations in response and stability.
  - High gains increase responsiveness but also move the system closer to instability.
  - All feedback systems can become unstable if improperly designed. In all real-world components there is some kind of lagging behavior between the input and output, characterized by the τ and ω_n in a transfer function. Instantaneous response is impossible in the real world!

- Instability in a feedback control system results from an *improper balance between the strength of the corrective action and the system dynamic lags*. When signals traverse a control loop, inherent phase lags in the control system can add enough delay to cause instability. Reducing delays within the control loop is a sure way to make room for higher gains; therefore, fast sampling and high-speed sensors are highly desirable.
- Instability that results from the accumulation of phase lag around the control loop usually occurs at just one frequency; this is why unstable control systems oscillate. All systems have noise and noise contains virtually all frequencies. If a system is unstable at any frequency, nature will find that frequency, usually in a few milliseconds.

- Each component in a control loop generates a certain amount of phase lag that varies with frequency: the controller, the amplifier, the plant, the feedback element. For most control systems, there will exist at least one frequency where the phase lag accumulates to 180º, but this alone does not cause instability. To cause instability, the loop gain must also be equal to unity. Similar to phase lag, each block in the loop contributes gain; the total gain through the loop is the accumulation of the block gains, i.e., the sum of the gains when measured in dB.
- Self-sustained oscillations occur when two conditions are met: the loop gain is unity (0 dB) and the phase lag around the loop is 180º.

Closed-Loop System Block Diagram



Open-Loop Transfer Function

$$\frac{B(s)}{E(s)} = G_c(s)G(s)H(s)$$

Feed-Forward Transfer Function

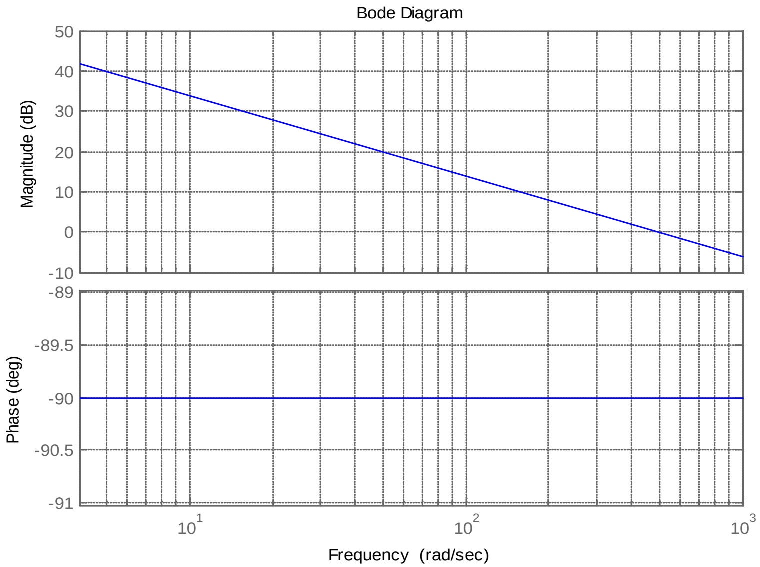
$$\frac{C(s)}{E(s)} = G_c(s)G(s)$$

Closed-Loop Transfer Functions

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1+G_c(s)G(s)H(s)}$$

$$\frac{C(s)}{D(s)} = \frac{G(s)}{1+G_c(s)G(s)H(s)}$$

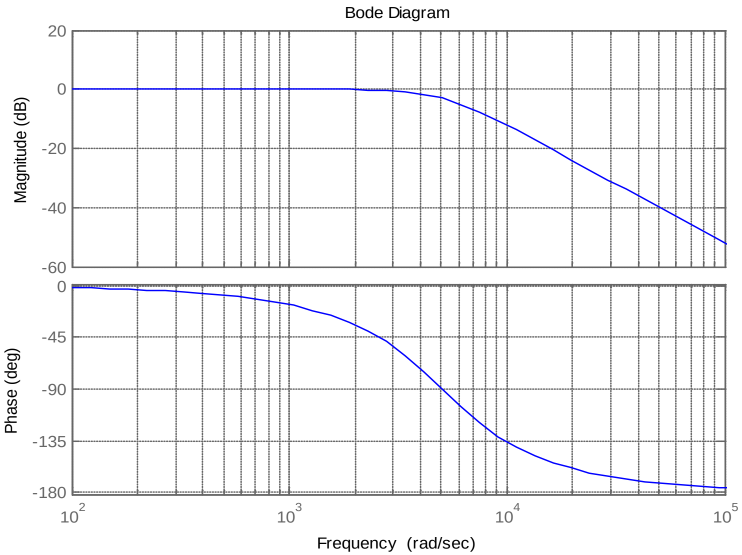
$$C(s) = \frac{G(s)}{1+G_c(s)G(s)H(s)}\left[G_c(s)R(s)+D(s)\right]$$



Example Problem

Plant $G(s) = \frac{1}{Js} = \frac{1}{0.002s} = \frac{500}{s}$

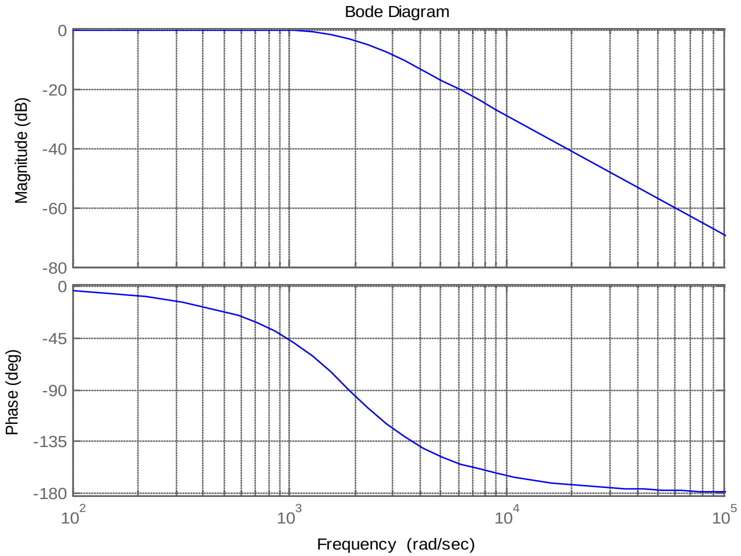
$K_t = 1 \text{ (Nm/A)}$



Amplifier $G_{amp}(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$

$\omega = 5027 \text{ (rad/s)} = 800 \text{ (Hz)}$

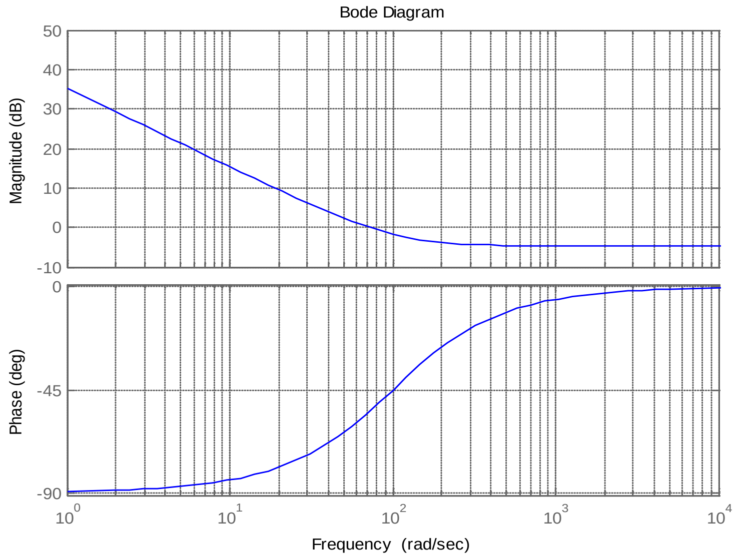
$\zeta = 0.707$



Sensor $$H(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

$$\omega = 1885 \text{ (rad/s)} = 300 \text{ (Hz)}$$

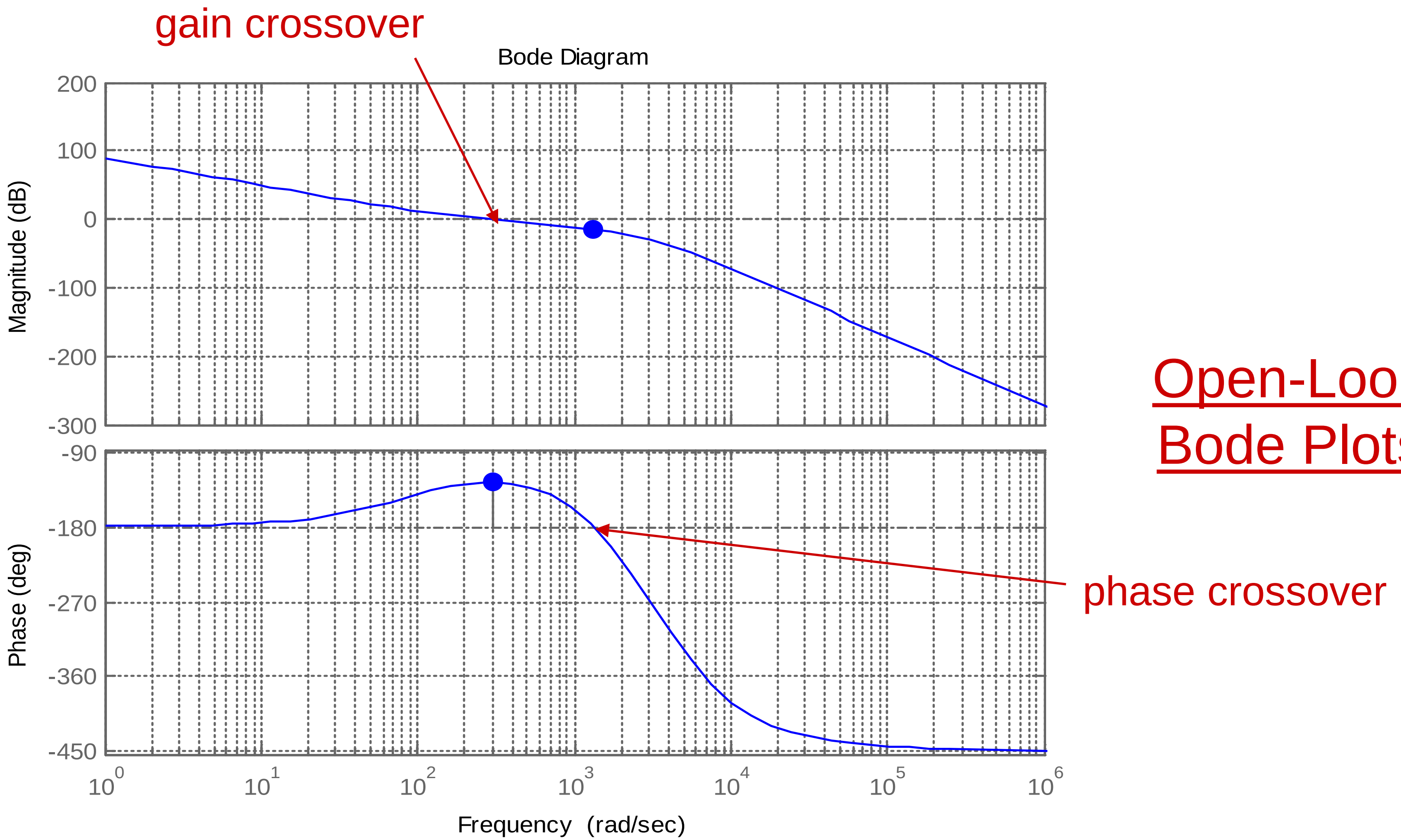
$$\zeta = 0.707$$



PI Controller $$G_c(s) = 0.58\frac{s+100}{s}$$

$$K_P + \frac{K_I}{s} = \frac{K_P s + K_I}{s} = \frac{K_P\left(s + \frac{K_I}{K_P}\right)}{s}$$

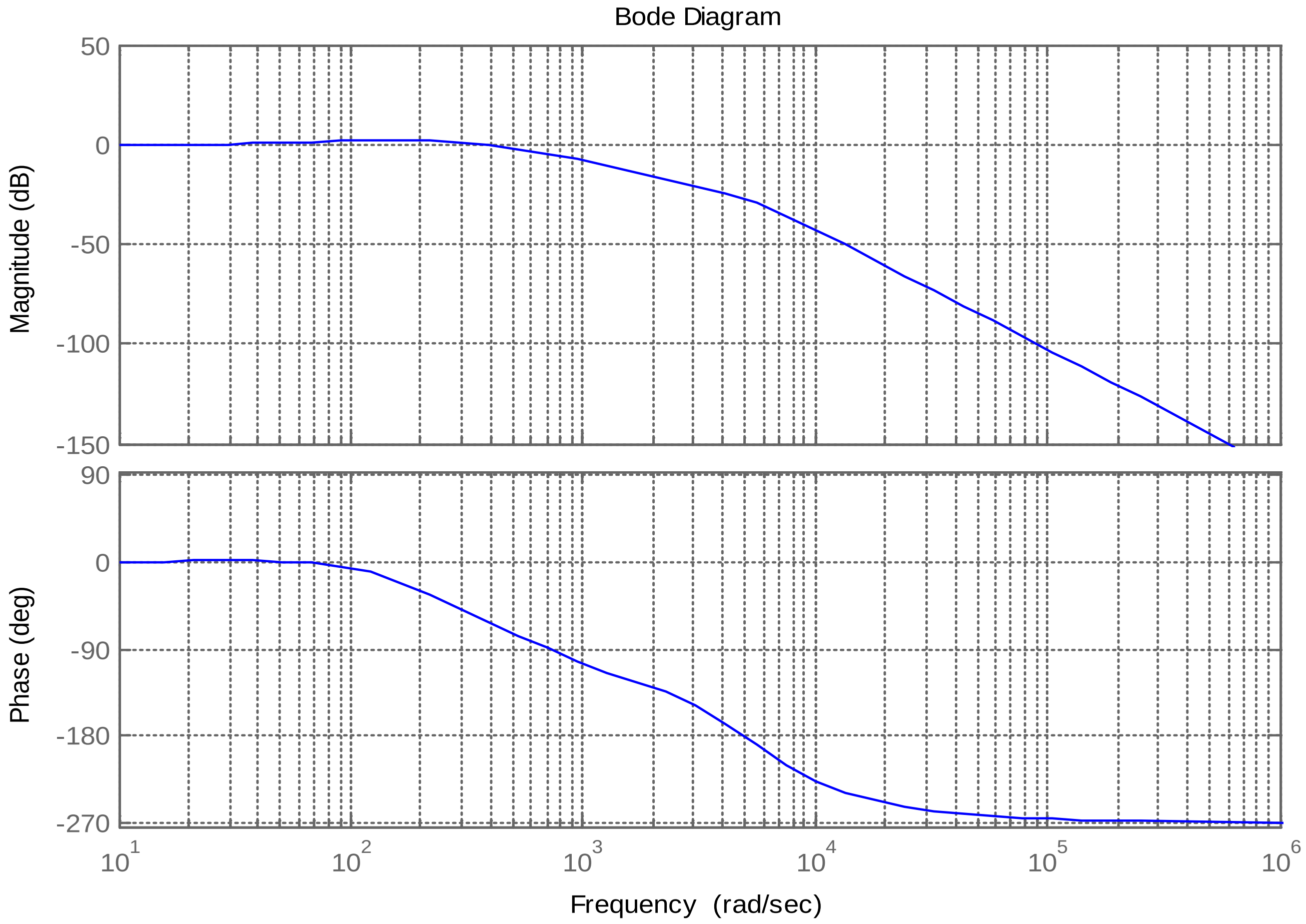
Open-Loop Bode Plots



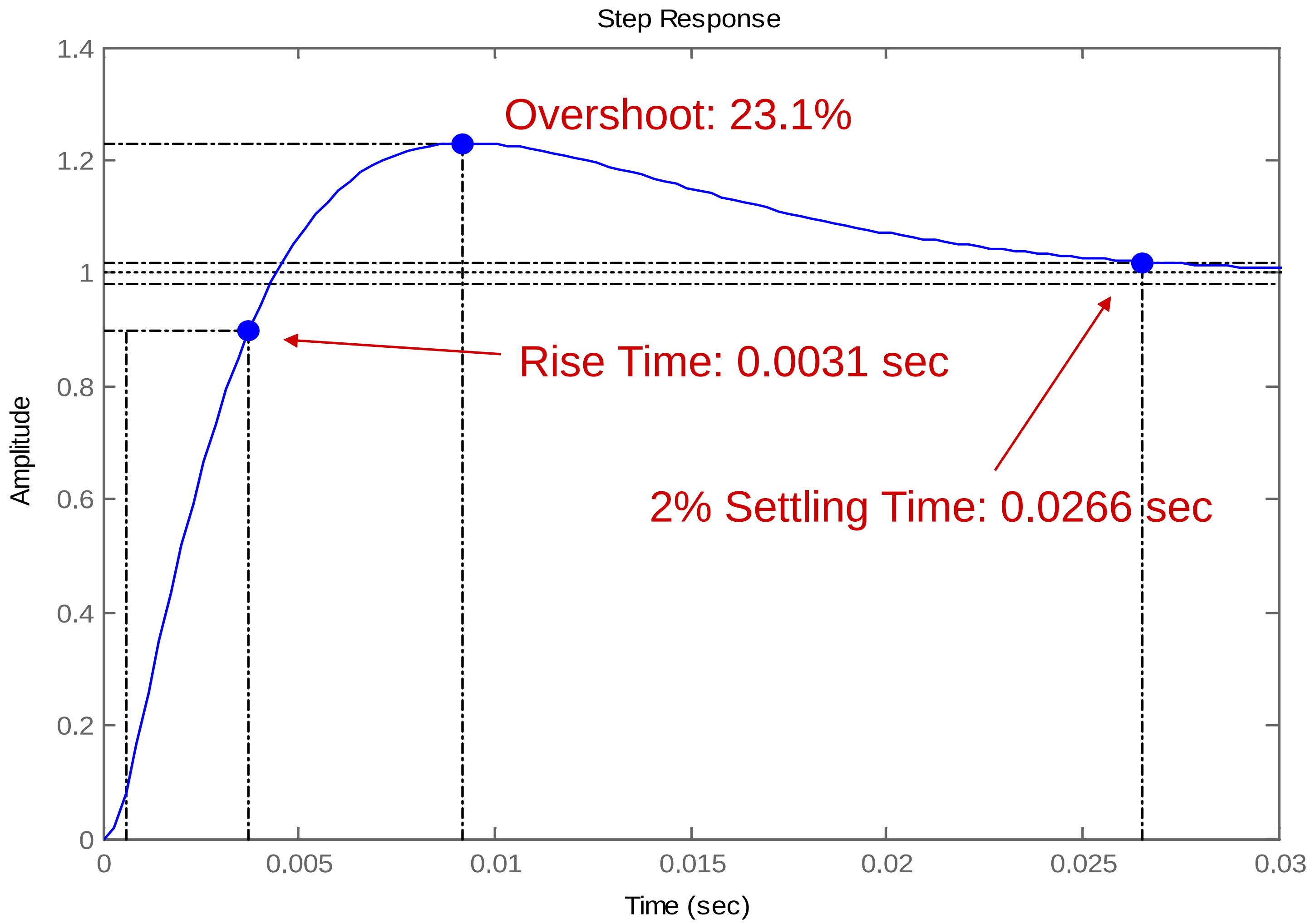
GM = 14.3 dB @ 1340 rad/s

PM = 53.7º @ 305 rad/s

Closed-Loop Bode Plots



Closed-Loop Step Response



- Two open-loop performance criteria in common use to specify relative stability are gain margin and phase margin.
- The open-loop frequency response is defined as $(B/E)(i\omega)$.
  - One could open the loop by removing the summing junction at R, B, E and just input a sine wave at E and measure the response at B. This is valid since $(B/E)(i\omega) = G_cGH(i\omega)$. Open-loop experimental testing has the advantage that open-loop systems are rarely absolutely unstable, thus there is little danger of starting up an untried apparatus and having destructive oscillations occur before it can be safely shut down.
- The utility of open-loop frequency-response rests on the Nyquist stability criterion.

- Gain margin (GM) and phase margin (PM) are in the nature of safety factors such that (B/E)(iω) stays far enough away from 1 ∠ -180° on the stable side.
  - Gain margin is the multiplying factor by which the steady state gain of (B/E)(iω) could be increased (nothing else in (B/E)(iω) being changed) so as to put the system on the edge of instability, i.e., (B/E)(iω)) passes exactly through the -1 point. This is called marginal stability.
  - Phase margin is the number of degrees of additional phase lag (nothing else being changed) required to create marginal stability.
- Both a good gain margin and a good phase margin are needed; neither is sufficient by itself.

- It is important to realize that, because of model uncertainties, it is not merely sufficient for a system to be stable, but rather it must have adequate stability margins.
  - Stable systems with low stability margins work only on paper; when implemented in real time, they are frequently unstable.
- The way uncertainty has been quantified in classical control is to assume that either gain changes or phase changes occur.
  - Typically, systems are destabilized when either gain exceeds certain limits or if there is too much phase lag (i.e., negative phase associated with unmodeled poles or time delays).
- As we have seen these tolerances of gain or phase uncertainty are the gain margin and phase margin.

- Although the measurements of PM and GM are objective, determining the desired values for these measures requires judgment.
  - Applications vary in the amount of margin they require. Systems responding to demanding commands, e.g., step command, require higher stability margins than those that must respond only to gentler commands. Some applications can tolerate more overshoot than others. Also, some control methods require more PM or GM than others for equivalent response.
- Experience teaches that GM should be between 10 and 25 dB, depending on the application and controller type; PM should be between 35º and 80º. All things being equal, more PM is better.
- Basic Rule: Eliminate unnecessary phase lag
  - Unnecessary phase lags limit the ultimate performance of the control system.

- The challenge in tuning is that multiple gains must be varied and each affects many of the performance measures.
  - It would be desirable if we could decouple the multiple tuning gains so that they may be adjusted individually. This can be done by considering the effects of each gain as being dominant over a certain frequency range.
- Let's consider the PI controller first applied to a simplified system.

$$G(s) = \frac{1}{Js} \qquad G_C(s) = K_P + \frac{K_I}{s} \qquad H(s) = 1$$

Closed-Loop Transfer Function

$$\frac{C(s)}{R(s)} = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)H(s)} = \frac{K_P s + K_I}{Js^2 + K_P s + K_I}$$

- While the numerator does have an impact on response, the denominator determines the overall stability and response of the system.
  - In the high-frequency range, s is large and the Js^2 term dominates the denominator.
  - In the middle-frequency range, the $K_P s$ term will dominate the denominator. Here the proportional gain dominates; the frequency is still too high for the integral gain to have much impact.
  - In the low-frequency range, as the frequency approaches zero, the K_I term will dominate.
- To apply a frequency-range-based approach to tuning, tune the highest frequency terms first.
  - Assuming the plant gain cannot be changed, tune K_P first. When optimizing K_P, you are balancing responsiveness (larger K_P) against stability (smaller K_P).

- Integral gain is important as it removes steady-state error. When tuning K_I, its value will have little bearing on the stability of the high-frequency range; changing K_I will not require returning to change K_P.

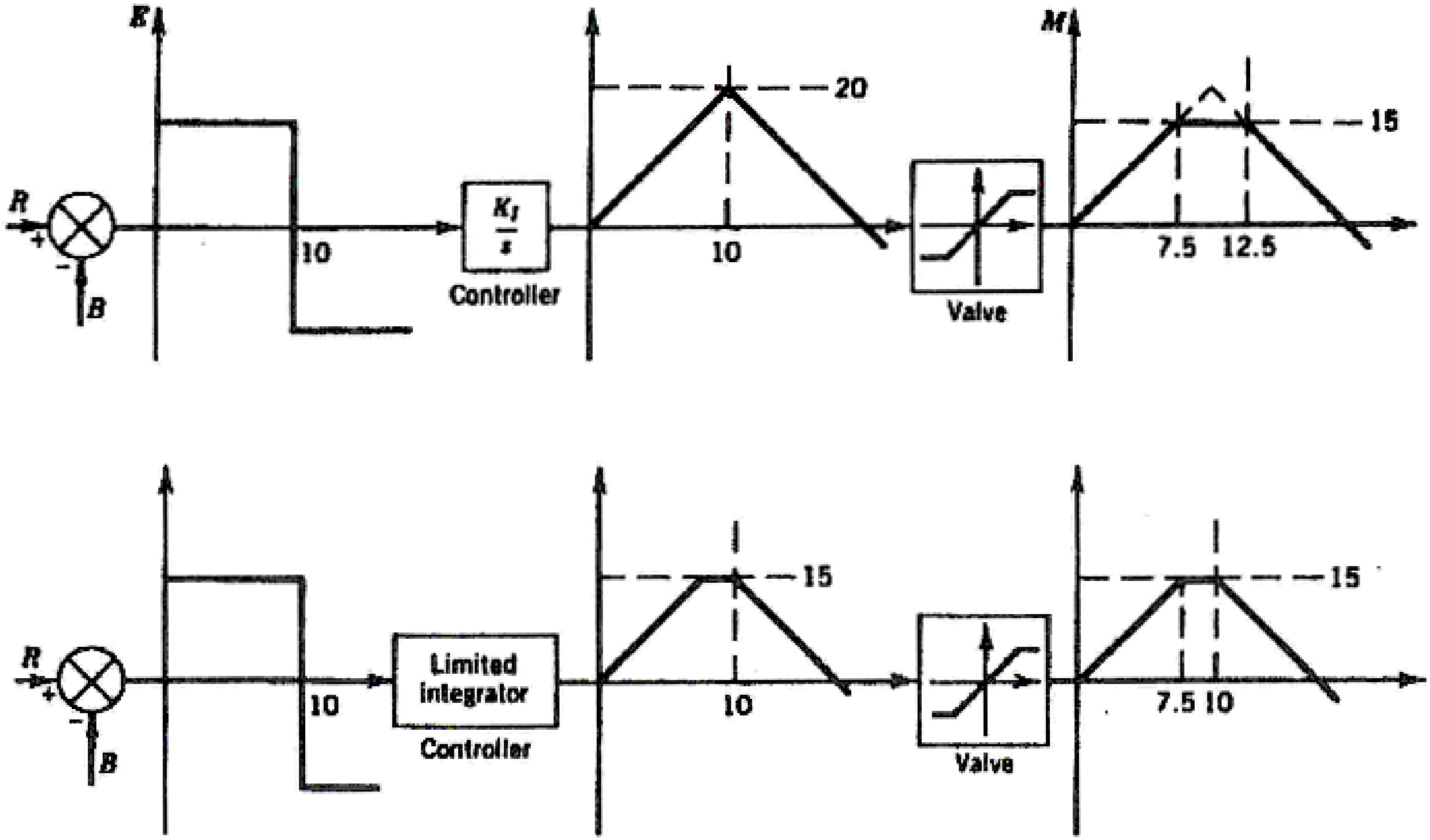
- What about changes in plant gain? The plant gain here is 1/J, which could change as J changes.
  - The plant gain is usually the most difficult to know and control.
  - The effect of changes in the plant gain is equivalent to that of changes in the loop gain as long as the system remains out of saturation.
  - Variations of ± 20% in the plant gain usually have little impact on control system performance. Larger variations may cause the system to be either unresponsive or marginally stable at the extremes. In such cases, it is difficult to find a single set of servo gains that will accommodate the range of variation.

- If changes in the plant can be anticipated with reasonable accuracy, a technique known as Gain Scheduling may be used.
  - This requires knowledge of the gain in question and a controller with gains that can be modified during operation.
- If this is not the case, it might be possible to sense the changes in the plant using the controller.
  - At the time the controller is commissioned, it can generate signals that excite the plant and then use the feedback to estimate the parameter. This technique is called Auto-Tuning. This, however, is usually not an option during normal operation, as strong excitation signals during normal operation would disturb the system. To sense gain for scheduling in this condition, a technique known as Adaptive Control is sometimes used.

- Multiple loops are common in control systems, e.g., a current loop may reside in a velocity loop which may reside in a position loop.
  - When tuning multi-loop control systems, we follow a similar procedure. The inner loops operate in the next higher frequency range to the outer loop. After an inner loop is tuned, it acts like a low-pass filter within the outer loop. Once an inner loop is tuned, there is little need to return to it, as each loop operates over a different frequency range.
  - When tuning, expect the bandwidth of the outer loop to be between 20% and 40% of that of the inner loop. Tune each loop to be as responsive as possible because it becomes the barrier for the next outer loop.

- What about Saturation?
  - Saturation is the most prevalent nonlinear behavior in industrial controllers and often causes a system to overshoot excessively, when it is actually quite stable.
  - While you should avoid nonlinear behavior in system design, all systems have limits as to how much energy they can supply to the plant. During saturation, the control system is applying all the power that is available. All systems have power converter limitations, and most will enter saturation immediately when those limits are exceeded.
  - Saturation causes a serious problem for integral controllers called wind-up.
- Avoid saturation when tuning.
  - Saturation is a nonlinear effect; it does not indicate either stability or responsiveness because the system is locked full on. When generating step response or Bode plots, ensure that the system does not enter saturation by monitoring the control signal; it should never reach its maximum value.

Integral Wind-Up and Its Correction



- Let's consider the situation of integral windup and its correction. Integral control may be degraded significantly by saturation effects.
  - For example, as seen in the figure, a large sustained error causes the integral controller to ramp its output pressure up to the 20-psig supply pressure. The diaphragm valve, sized to be wide open at 15 psig (the upper end of the 3 to 15 psig control range) saturates at 15 psig.
  - The integral signal beyond t = 7.5 seconds is really useless since it asks for a motion that the valve cannot produce. When the error reverses at t = 10 seconds, the valve cannot respond to this change until the integral signal (which has "wound up" to 20 psig) is "unwound" back to the 15-psig level at t = 12.5.
- This delayed response is called integral wind-up.

- Note that this delay is in addition to the normal lagging behavior of integral control and can cause excessive overshooting and stability problems.

- Integral windup is of course not a problem in every application of integral control. If difficulty is anticipated, the controller can be modified in different ways to give various degrees of improvement. Basically, one wants to disable the integrator whenever its output signal causes saturation in the final control element.
  - In this example, the integrator is disabled when its output pressure reaches 15 psig, preventing any windup.
  - When the error reverses at t = 10 seconds the integrator and valve immediately respond to the negative error since there is no windup that needs to first be unwound.

Types of Industrial Controllers

- The PID controller is the most common controller in general use. It can be simplified by setting one or two of the three gains to zero.
- Let's explore variations of the P, I, and D gains.
- In choosing a controller for an application, the designer must weigh complexity vs. performance. More complex controllers require more processing capability, e.g., faster processors for digital control or more components for analog control. They are also more difficult to tune.
- How much performance is worth paying for?

- The basic issues in control systems vary little between digital and analog controllers. For all control systems, gain and phase margins must be maintained, and phase loss around the loop should be minimized.
- First the focus will be on velocity control of a single-integrating loop, i.e., control of torque to produce velocity. Then we will focus on a double-integrating position loop.
- To facilitate side-by-side comparison of different controller types, we evaluate response to a square wave input, as this is the signal of choice for exposing marginal stability; testing with gentler signals may allow marginal stability to pass undetected.

- For each of the controllers discussed, we will use frequency-range-based tuning. The P and D gains combine to determine behavior in the high-frequency range and they should be set first; simultaneously tuning is required. The I gain and a command filter determine behavior in the low-frequency range.
- The high-frequency range is limited by the plant, amplifier, and sensor. The low-frequency range is limited primarily by the high-frequency range.
- Time delays, i.e., computation delay, sample-and-hold delay, sensor delay, can all be added as needed to the control system.
- We focus here on command response, although disturbance response is just as important.

- Command response is usually preferred for determining stability as commands are easier to generate. When tuning, the command should be as large as possible to maximize the signal-to-noise ratio, as this supports accurate measurements; however, the power converter must remain out of saturation during these tests.
- Proportional Control
- Virtually all controllers have a large proportional gain. While derivative gain can provide incremental improvements at high frequencies, and integral gain improves performance at lower frequencies, the proportional gain is the primary actor across the entire frequency range of operation.

- Here the manipulating variable *M* is directly proportional to the actuating signal *E*.
- We assume that the dynamics associated with the real controller are negligible relative to other system dynamics.
- The corrective effort is made proportional to system "error"; large errors engender a stronger response than do small ones. We can vary in a continuous fashion the energy and/or material sent to the controlled process.

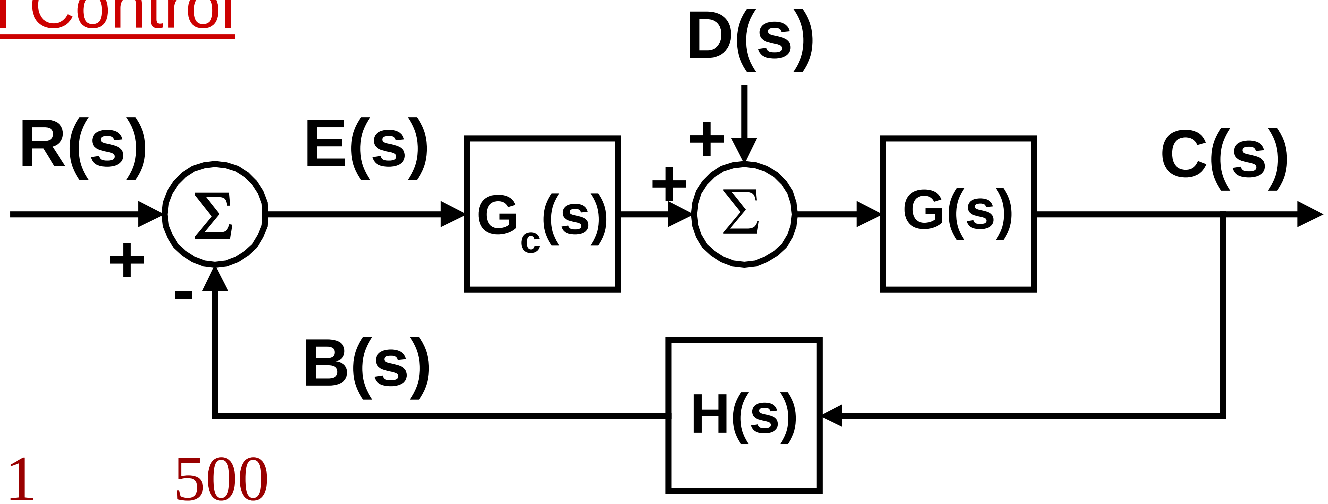
$$m(t) = K_P e(t)$$

$$\frac{M(s)}{E(s)} = K_P = \text{ Proportional Gain}$$

- Relative to on-off control, the advantage is a lack of limit-cycling behavior. The disadvantages are general complexity, higher cost, and lower reliability of hardware.
- Proportional control exhibits nonzero steady-state errors for even the least-demanding commands and disturbances.
- Why is this so? Suppose for an initial equilibrium operating point $x_c = x_v$ and steady-state error is zero. Now ask x_c to go to a new value x_{vs}. It takes a different value for the manipulated input *M* to reach equilibrium at the new x_c. When the manipulated input *M* is proportional to the actuating signal *E*, a new *M* can only be achieved if *E* is different from zero which requires $x_c \neq x_v$; thus, there must be a steady-state error.

- How do you tune a proportional controller?
  - Set K_P low.
  - Apply a square wave command at about 10% of the desired loop bandwidth. Use large amplitude, but avoid saturation.
  - Raise K_P for little of no overshoot.
  - The loss of stability is a consequence of phase loss in the loop, and the proportional gain will rise to press that limit. Be aware, however, that other factors, primarily noise, often ultimately limit the proportional gain below what the stability criterion demands.
  - Independent of its source, noise will be amplified by the high-frequency gains in the controller, such as the proportional gain.
  - Setting the proportional gain requires balancing the need for performance and the elimination of noise.

Proportional Control



$$G(s) = \frac{1}{Js} = \frac{1}{0.002s} = \frac{500}{s}$$

$$G_{amp}(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

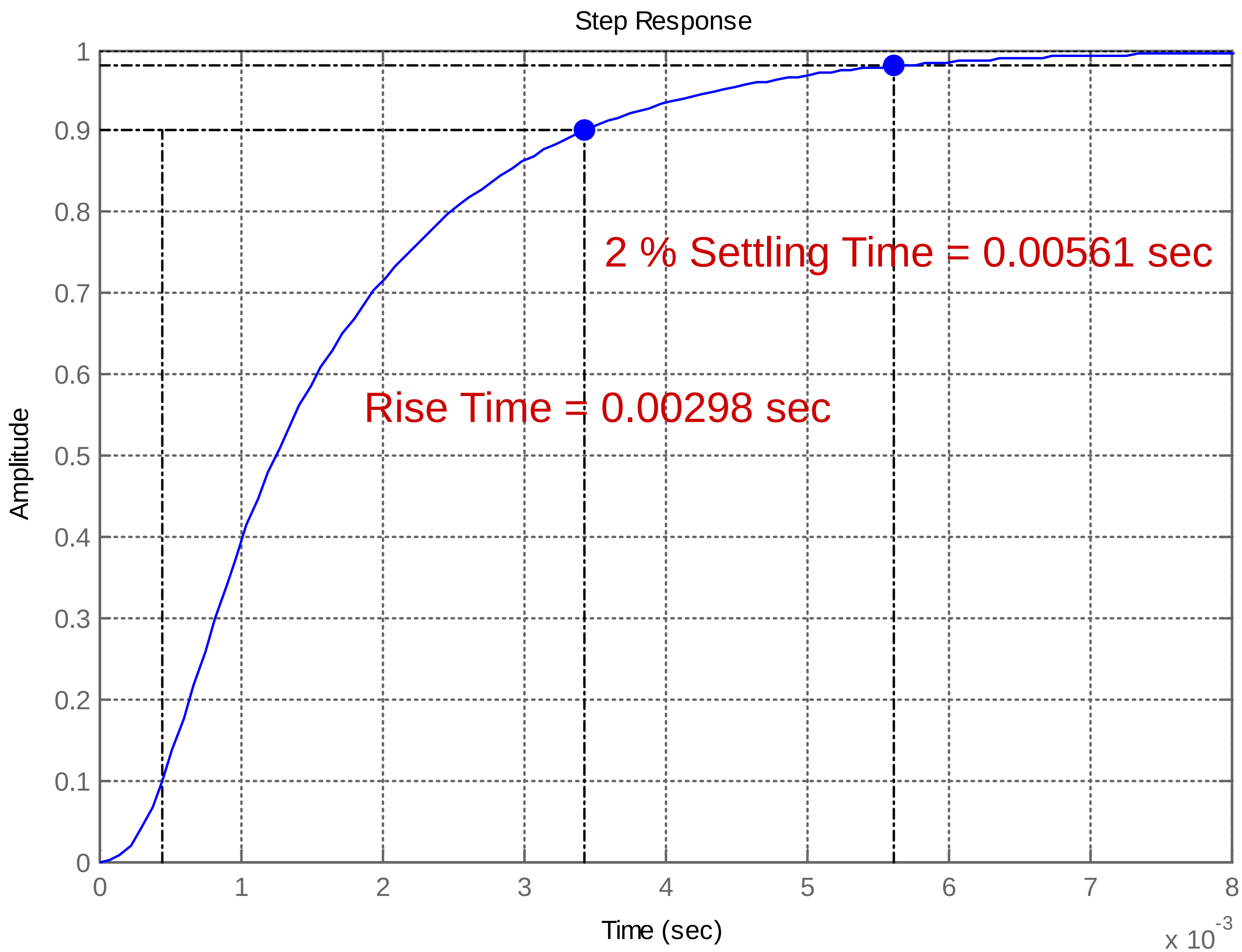
$$\omega = 5027 \text{ (rad/s)} = 800 \text{ (Hz)}$$

$$\zeta = 0.707$$

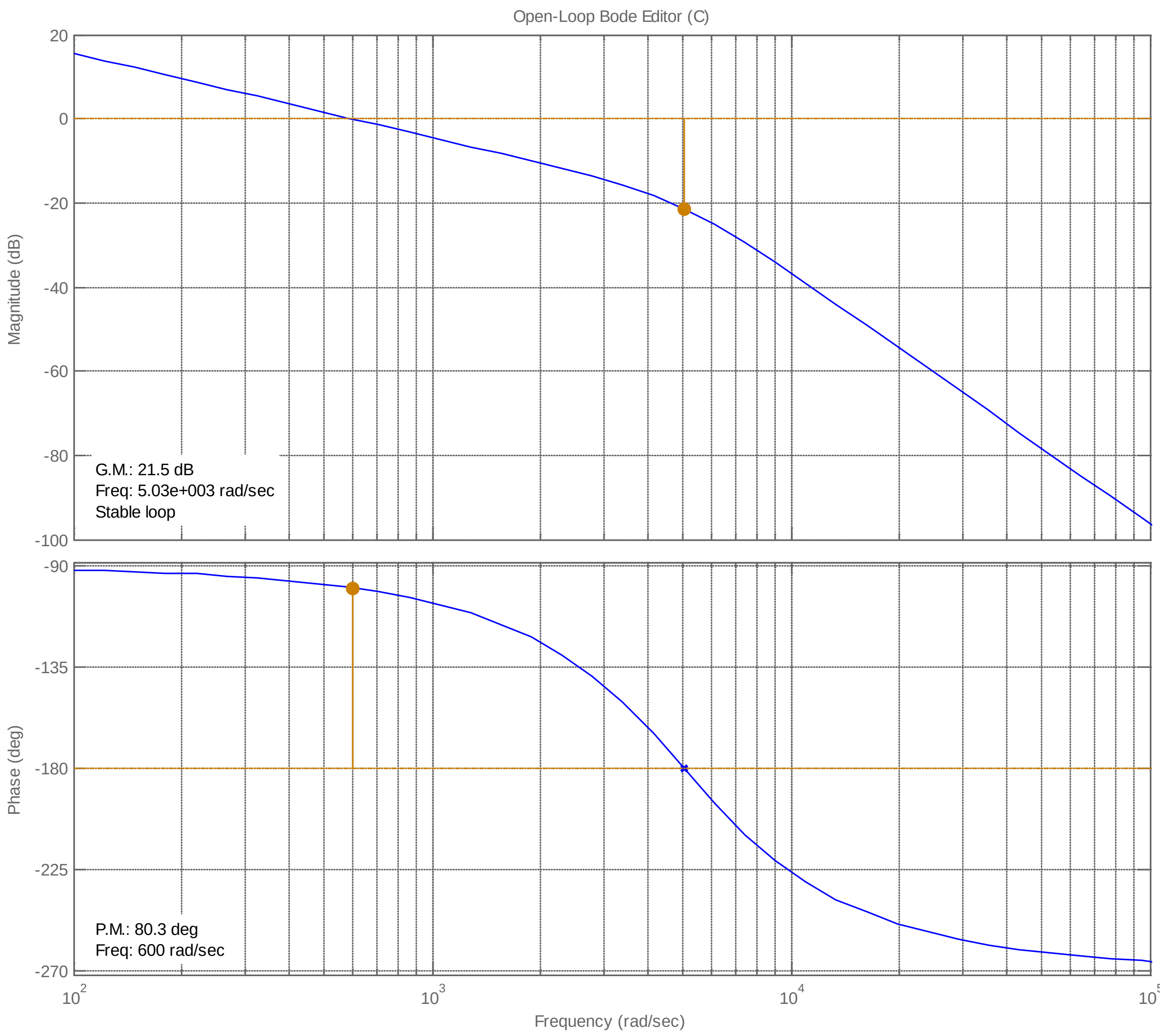
$$K_t = 1 \text{ (Nm/A)}$$

$$H(s) = 1 \qquad G_C(s) = K_P = 1.2$$

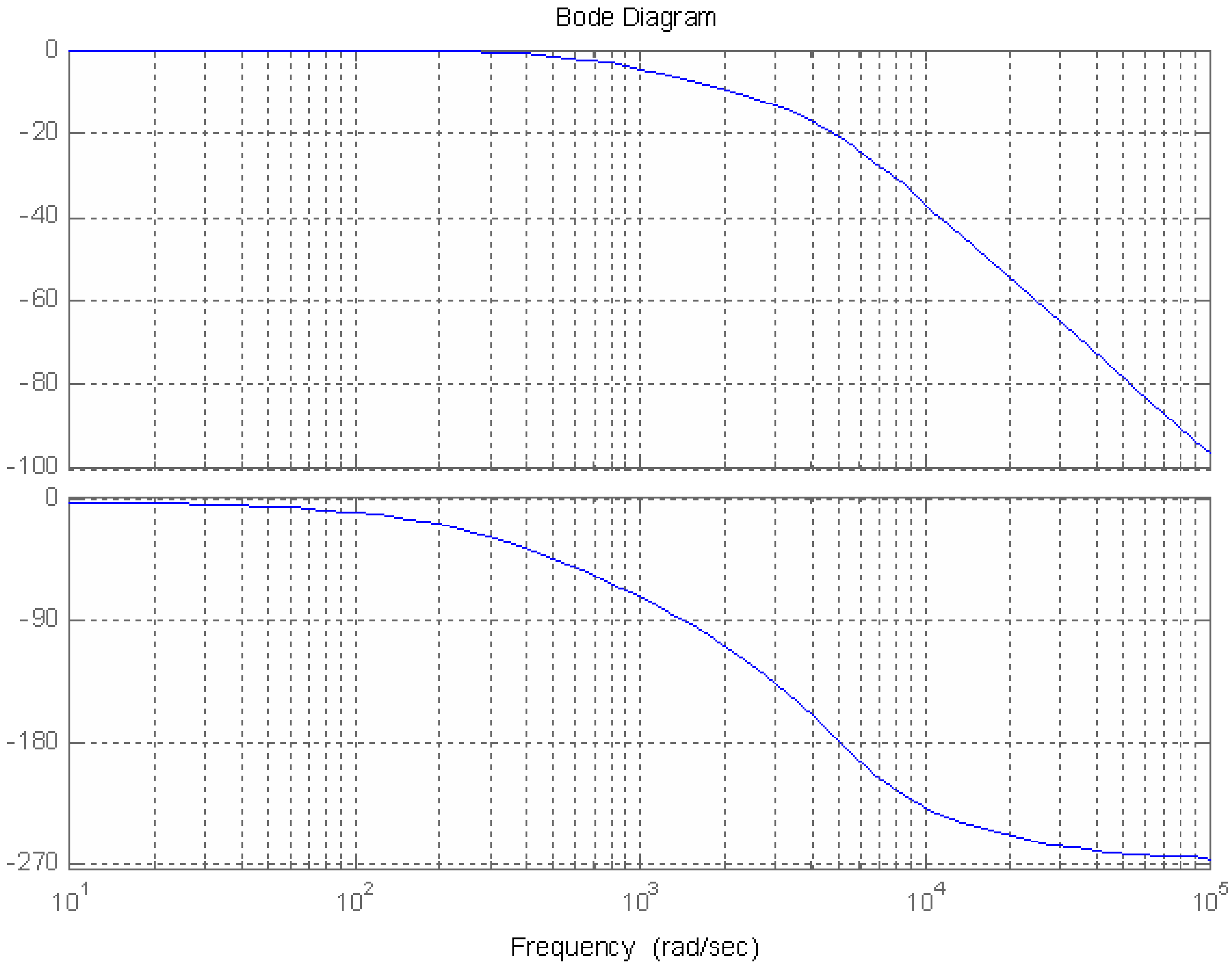
Closed-Loop Step-Response Plot



Open-Loop Bode Plots: PM = 80.3º and GM = 21.5 dB



Closed-Loop Bode Plots: Closed-Loop Bandwidth = 726 rad/s



- Integral Control
- When a proportional controller can use large loop gain and preserve good relative stability, system performance, including those on steady-state error, may often be met.
- However, if difficult process dynamics such as significant dead times prevent use of large gains, steady-state error performance may be unacceptable.
- When human process operators notice the existence of steady-state errors due to changes in desired value and/or disturbance they can correct for these by changing the desired value ("set point") or the controller output bias until the error disappears. This is called *manual reset*.

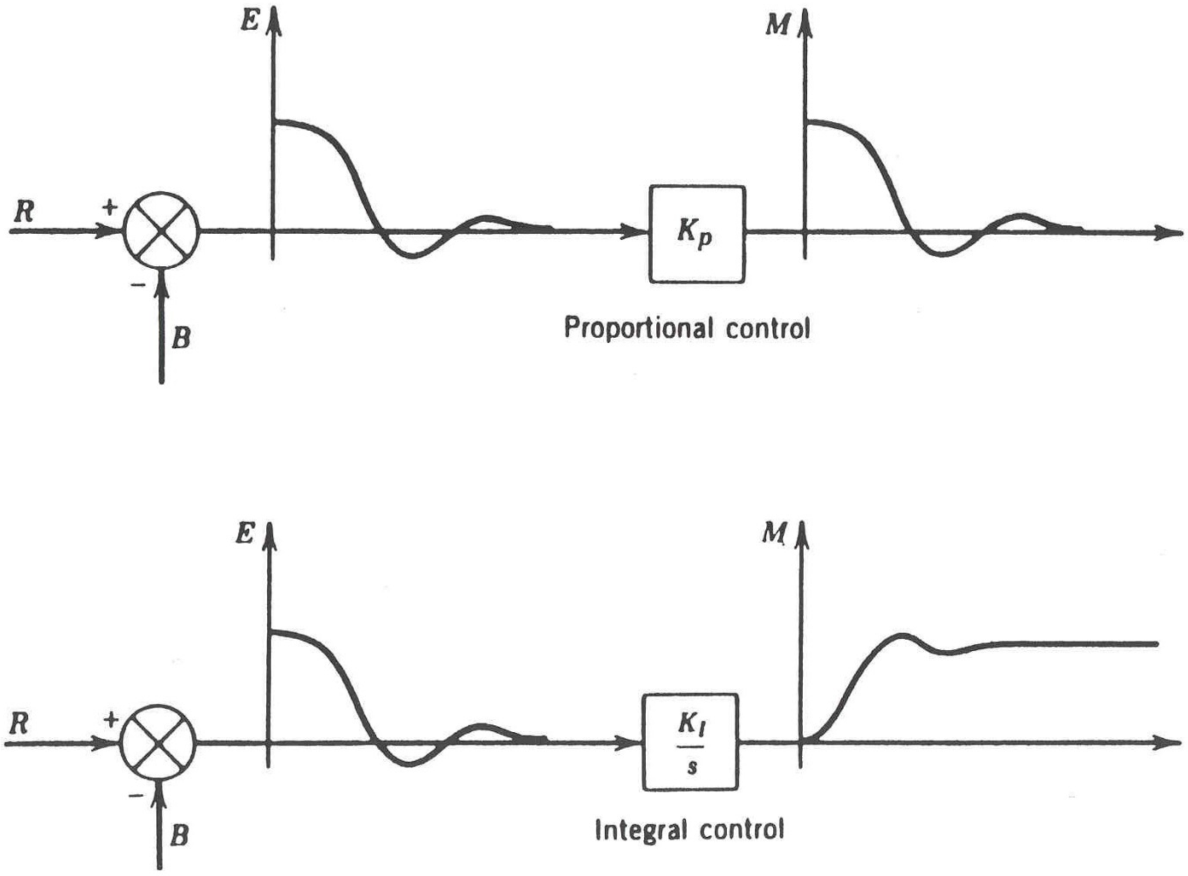
- Integral control is a means of removing steady-state errors without the need for manual reset. It is sometimes called *automatic reset*.

$$\frac{dm(t)}{dt} = K_i e(t) \qquad \frac{M(s)}{E(s)} = \frac{K_i}{s}$$

$$m(t) = K_i \int_0^t e(\tau)d\tau$$

- If the value of e(t) is doubled, then the value of m(t) varies twice as fast.
- For e(t) = 0, m(t) remains stationary.
- We have seen why proportional control suffers from steady-state errors. We need a control that can provide any needed steady output (within its design range, of course) when its input (system error) is zero.

Comparison: Proportional vs. Integral Control



Integral control has the undesirable side effects of reducing response speed and degrading stability

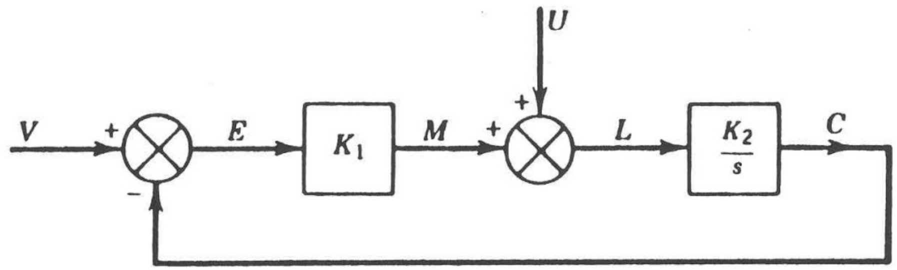
- Although integral control is very useful for removing or reducing steady-state errors, it has the undesirable side effect of reducing response speed and degrading stability.
- Why? Reduction in speed is most readily seen in the time domain, where a step input (a sudden change) to an integrator causes a ramp output, a much more gradual change.
- Stability degradation is most apparent in the frequency domain (Nyquist Criterion) where the integrator reduces the phase margin by giving an additional 90 degrees of phase lag at every frequency, rotating the (B/E)(iω) curve toward the unstable region near the -1 point.

- Occasionally an integrating effect will naturally appear in a system element (actuator, process, etc.) other than the controller.
- These gratuitous integrators can be effective in reducing steady-state errors. Although controllers with a single integrator are most common, double (and occasionally triple) integrators are useful for the more difficult steady-state error problems, although they require careful stability augmentation.
- Conventionally, the number of integrators between E and C in the forward path has been called the *system type number*.

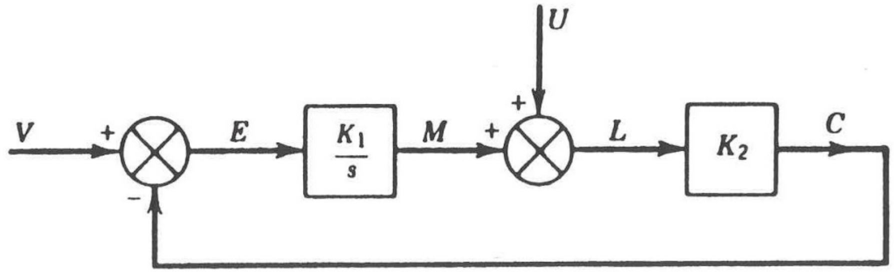
In addition to the number of integrators, their location (relative to disturbance injection points) determines their effectiveness in removing steady-state errors.

In Figure (a) the integrator gives zero steady-state error for a step command but not for a step disturbance.

By relocating the integrator as in Figure (b), either or both step inputs V_s and U_s can be "canceled" by M without requiring E to be nonzero.



(a)



(b)

Integrators must be located upstream from disturbance-injection points if they are to be effective in removing steady-state errors due to disturbances.

Location is not significant for steady-state errors caused by commands.

- Integral control can be used by itself or in combination with other control modes.
- Proportional + Integral (PI) Control is the most common mode.

$$m(t) = K_p e(t) + \frac{K_p}{T_I}\int_0^t e(\tau)d\tau$$

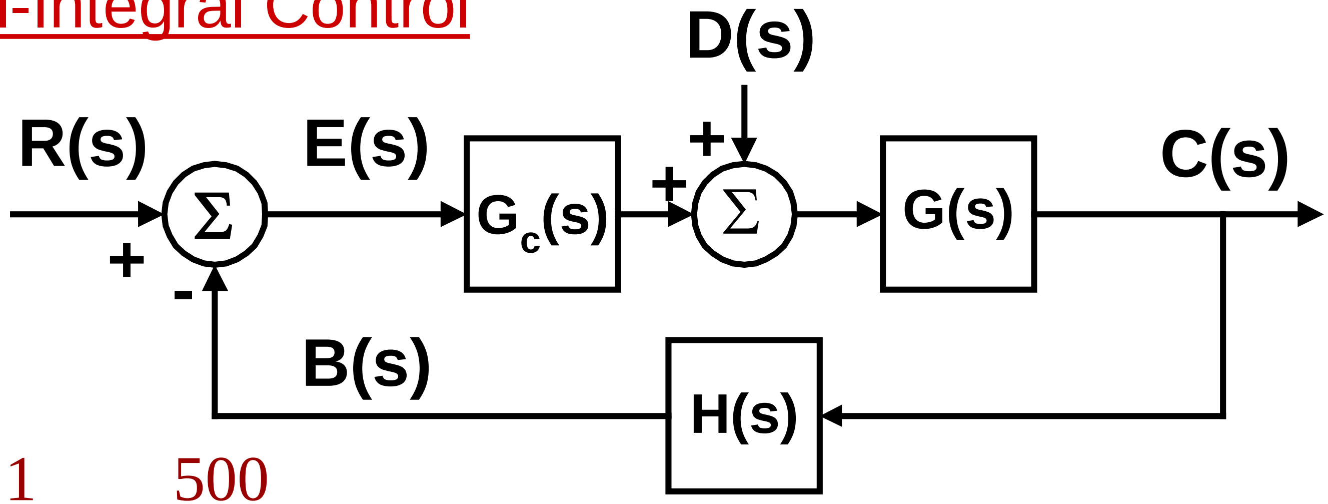
$$\frac{M(s)}{E(s)} = K_p\left(1 + \frac{1}{T_I s}\right)$$

- T_I = integral time
- $1/T_I$ = reset rate = number of times per minute that the proportional part of the control action is duplicated = repeats per minute

- Integral gain provides DC and low-frequency stiffness. When a DC error occurs, the integral gain will move to correct it. The higher the gain, the faster the correction. Fast correction implies a stiffer system.
- Don't confuse DC stiffness with dynamic stiffness. A system can be quite stiff at DC and not stiff at all at high frequencies! Higher integral gains will provide higher DC stiffness but will not substantially improve stiffness at or above the loop bandwidth.
- PI controllers are more complicated to implement than P controllers. Saturation becomes more complicated, as wind-up must be avoided. In analog controllers, clamping diodes must be added, and in digital controllers, saturation algorithms must be coded.

- Integral gain can cause instability. In the open loop, the integral, with its 90º phase lag, reduces PM. In the time domain, the common result of adding integral gain is overshoot and ringing. As a result, larger integral gains usually reduce bandwidth.
- PI controllers have two frequency ranges: high and low. The high range is served by K_P and the low by K_I.
- To tune a PI controller, set the P gain as it was in the P controller. After the higher frequency range is complete, K_I can be tuned. Raise K_I for 15% overshoot to a step input; a modest amount of overshoot to a step input is tolerable in most applications.

Proportional-Integral Control



$$G(s) = \frac{1}{Js} = \frac{1}{0.002s} = \frac{500}{s}$$

$$G_{amp}(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

$$\omega = 5027 \text{ (rad/s)} = 800 \text{ (Hz)}$$

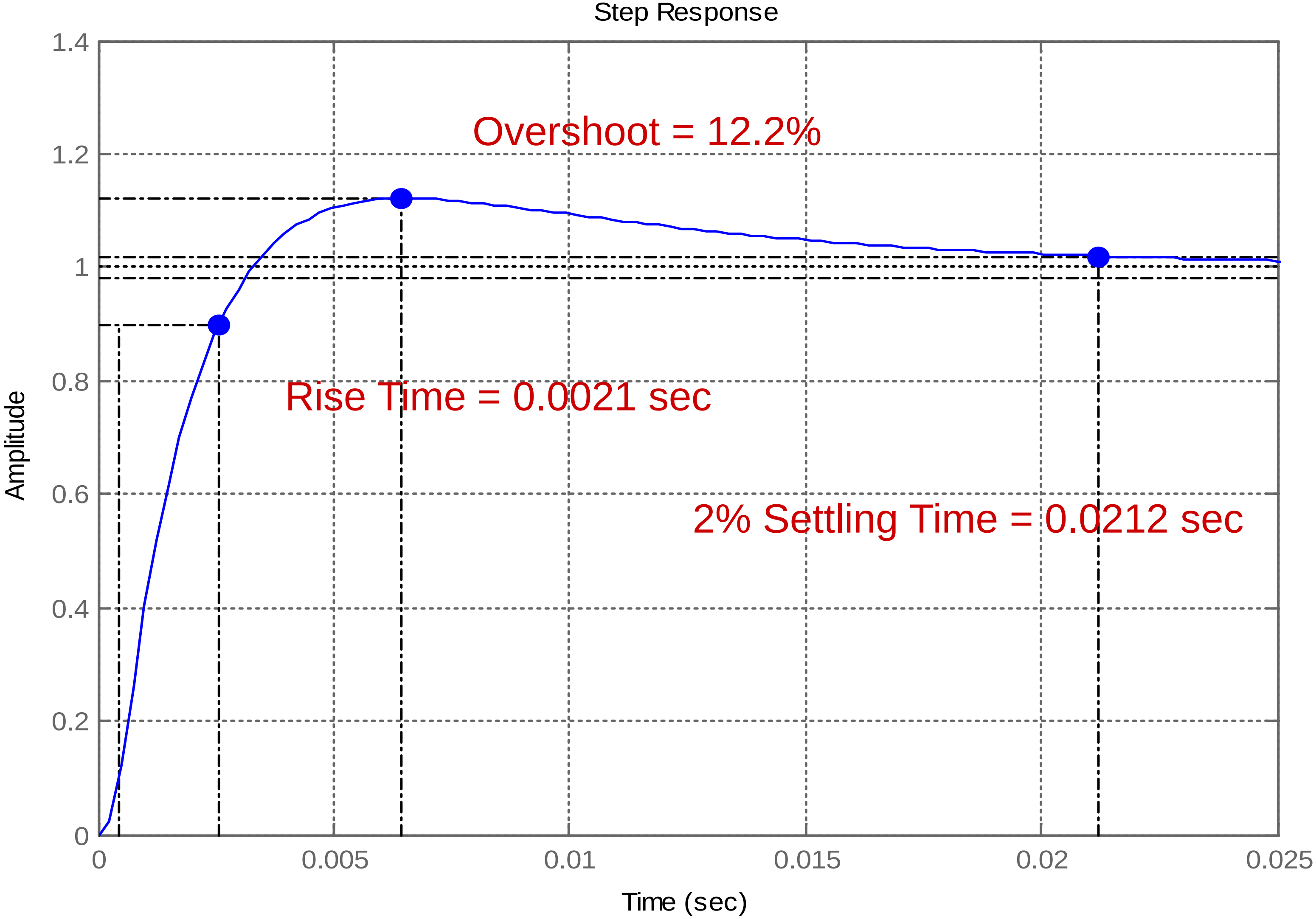
$$\zeta = 0.707$$

$$K_t = 1 \text{ (Nm/A)}$$

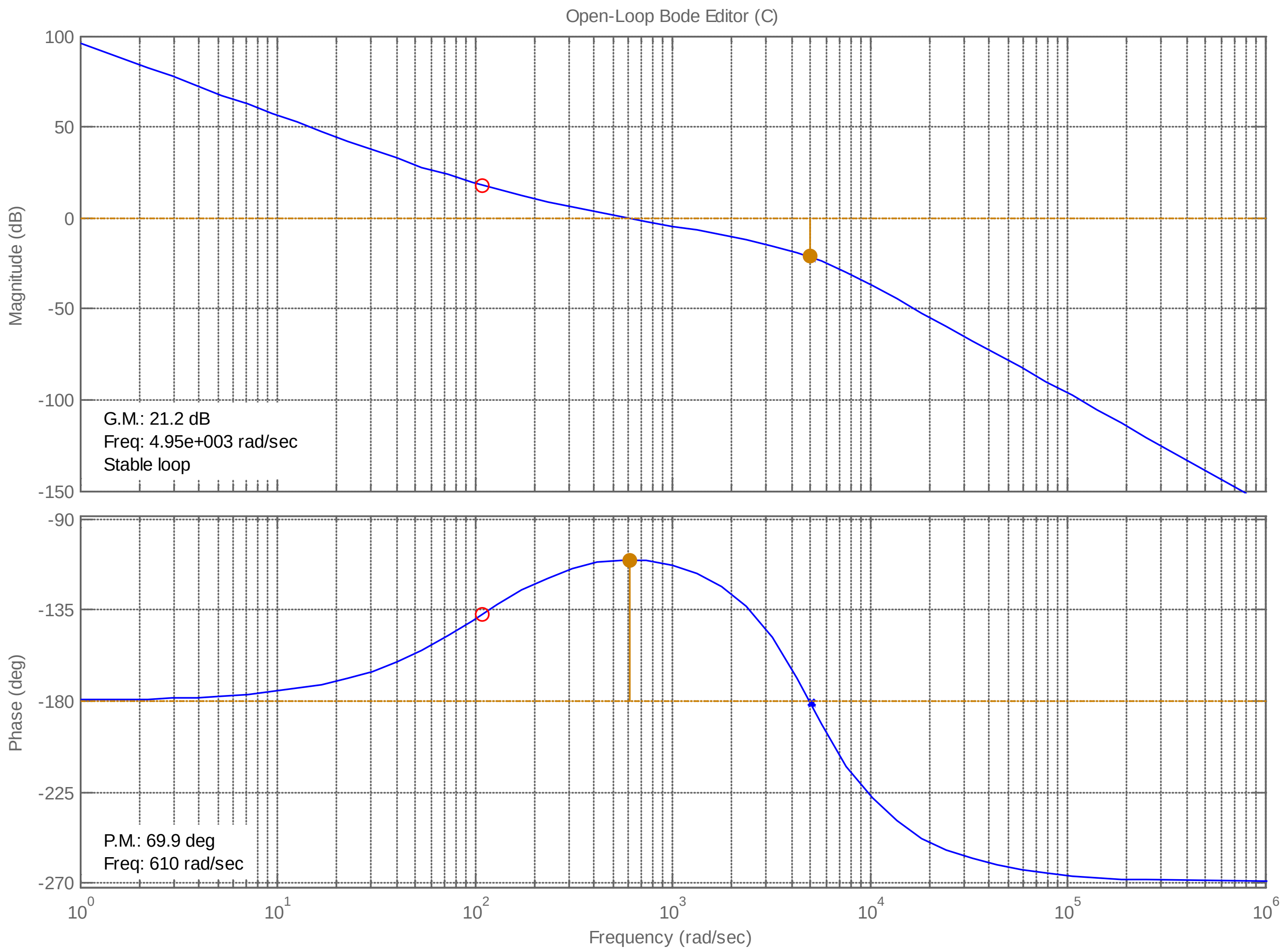
$$H(s) = 1$$

$$G_C(s) = K_P\left(1 + \frac{K_I}{s}\right) = K_P\left(\frac{s + K_I}{s}\right) = 1.2\left(\frac{s + 110}{s}\right)$$

Closed-Loop Step-Response Plot

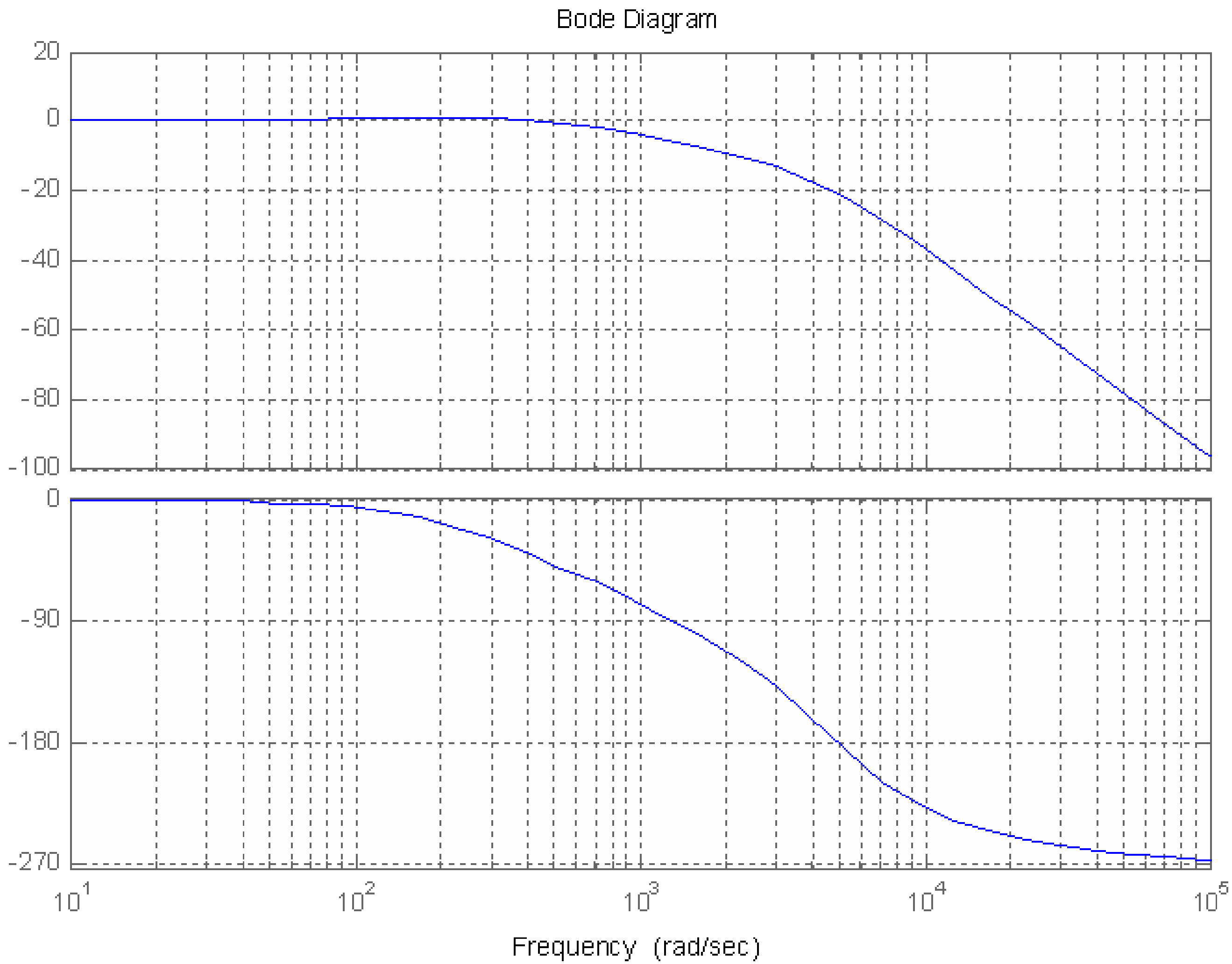


Open-Loop Bode Plots: PM = 69.9º and GM = 21.2 dB



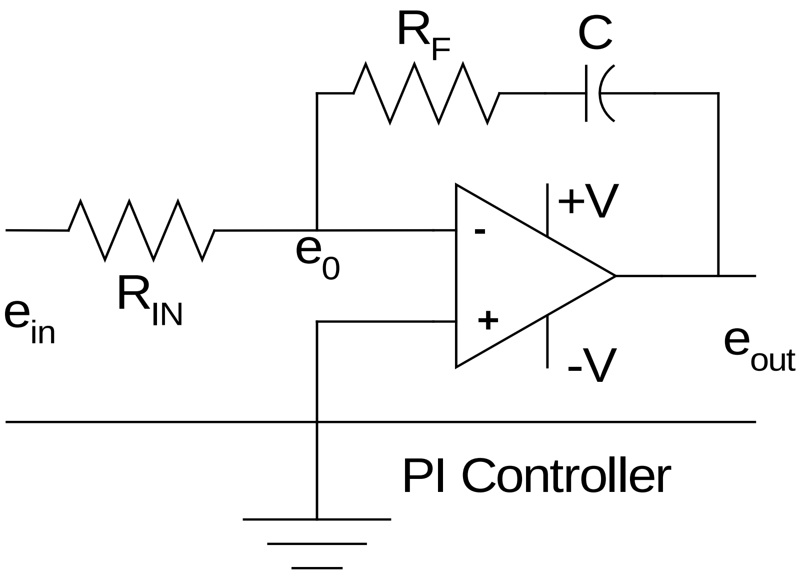
Closed-Loop Bode Plots

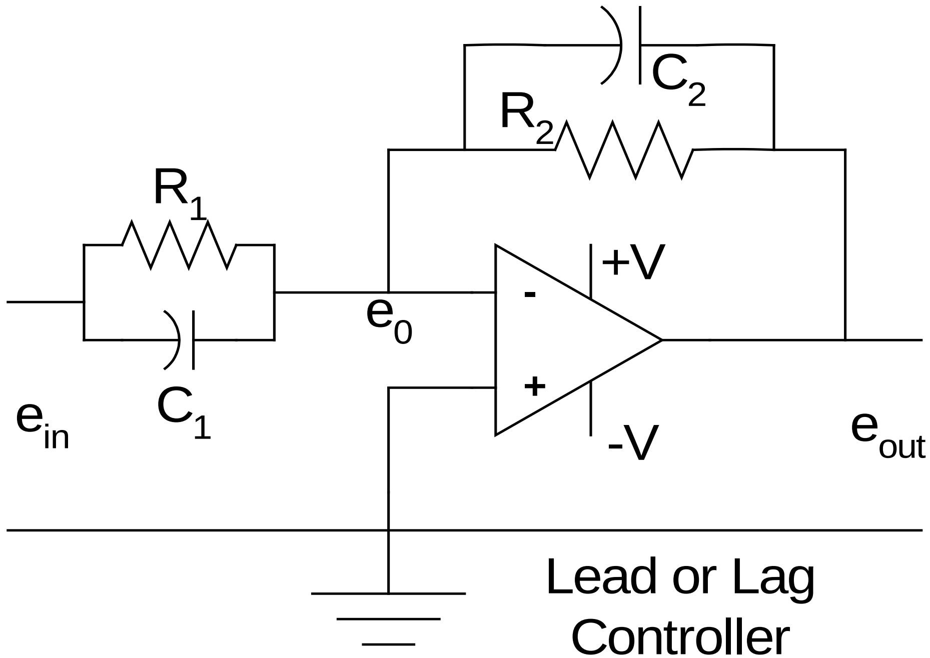
Closed-Loop Bandwidth = 851 rad/s
Peak Response = 1.08 dB @ 223 rad/s



Analog PI Controller and Lag Controller

$$\frac{e_{out}}{e_{in}} = -\frac{R_F C s + 1}{R_{IN} C s}$$

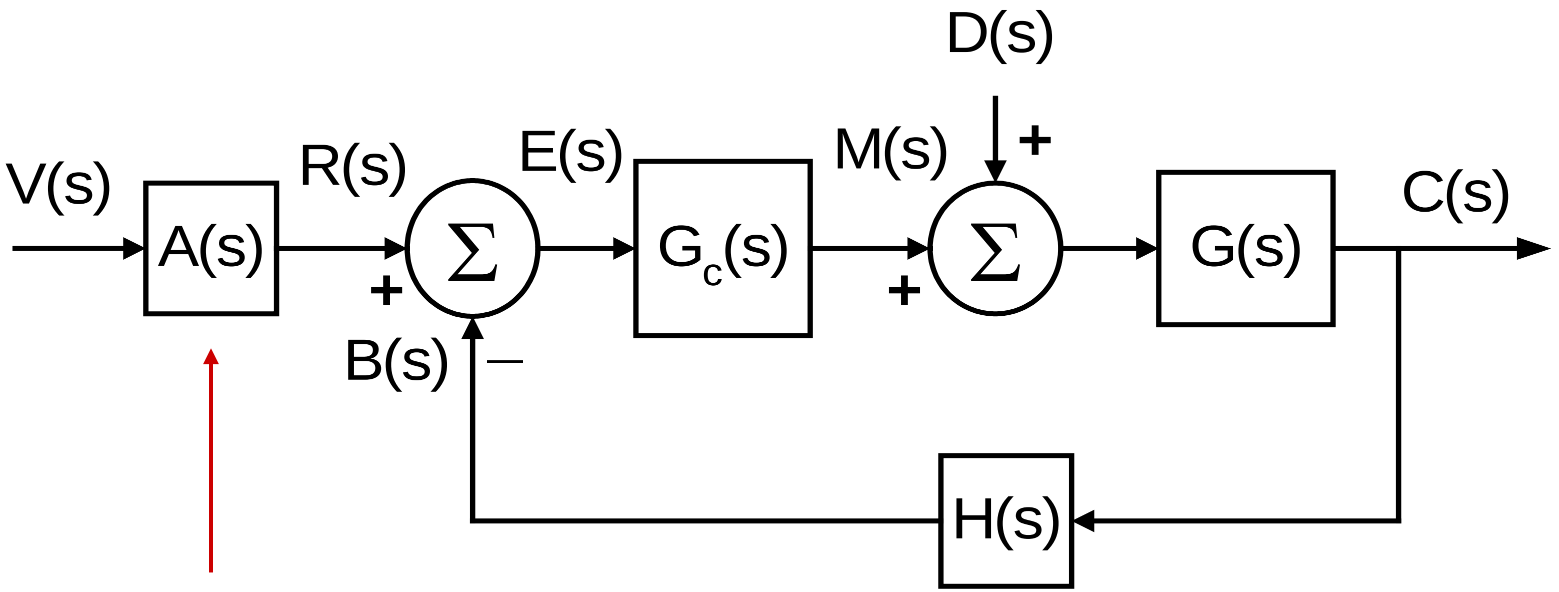




$$\frac{e_{out}}{e_{in}} = -\frac{R_2}{R_1}\frac{R_1 C_1 s + 1}{R_2 C_2 s + 1}$$

- PI+ Control
- This is an enhancement to PI control. Usually, because of the overshoot, the integral gain in PI control is limited to relatively small values.
- PI+ control uses a low-pass filter on the command signal to remove overshoot. In this way, the integral gain can be raised to higher values. PI+ control is useful in applications where the rejection of DC disturbances is paramount. The primary shortcoming of PI+ control is that the command filter also reduces the controller's command response.

PI+ Controller



Command Filter

$$A(s) = K_F + (1 - K_F)\frac{K_I}{s + K_I}$$

K_F = 1: all filtering removed

K_F = 0: filtering most severe

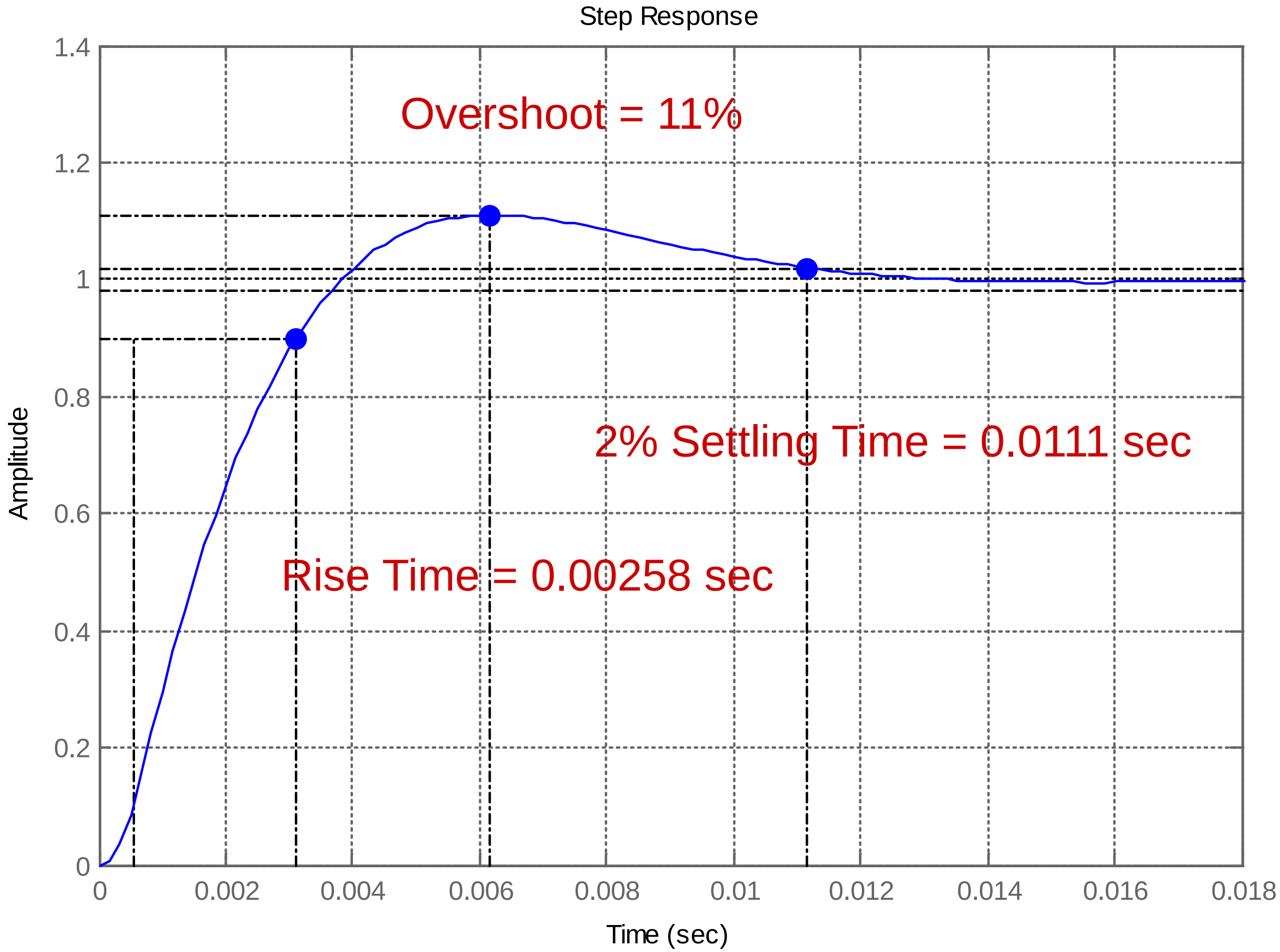
- When $K_F = 0$, the command is filtered by a single-pole, low-pass filter, with the pole at $-K_I$. This allows the highest integral gain but will also most severely limit the controller command response.
- Typically, $K_F = 0$ will allow an increase of almost three times in the integral gain but will reduce the bandwidth by ½ when compared with $K_F = 1$ (PI control).
- Finding the optimal value of K_F depends on the application, but a value of 0.65 has been found to work in many applications. This value typically allows the integral gain to more than double while reducing the bandwidth by only 15% - 20%.
- Why select the frequency of the command low-pass filter to be K_I? Why not set this higher or lower?

- The reason is that this frequency is excellent at canceling the peaking caused by the integral gain.
- One way to look at PI+ control is that it uses the command filter to attenuate the peaking caused by PI control. The peaking caused by K_I can be cancelled by the attenuation of a low-pass filter with a break of K_I.
- How do you tune a PI+ Controller?
- Tuning a PI+ controller is similar to tuning a PI controller except that you must choose the amount of filtering (K_F) before tuning the integral gain.

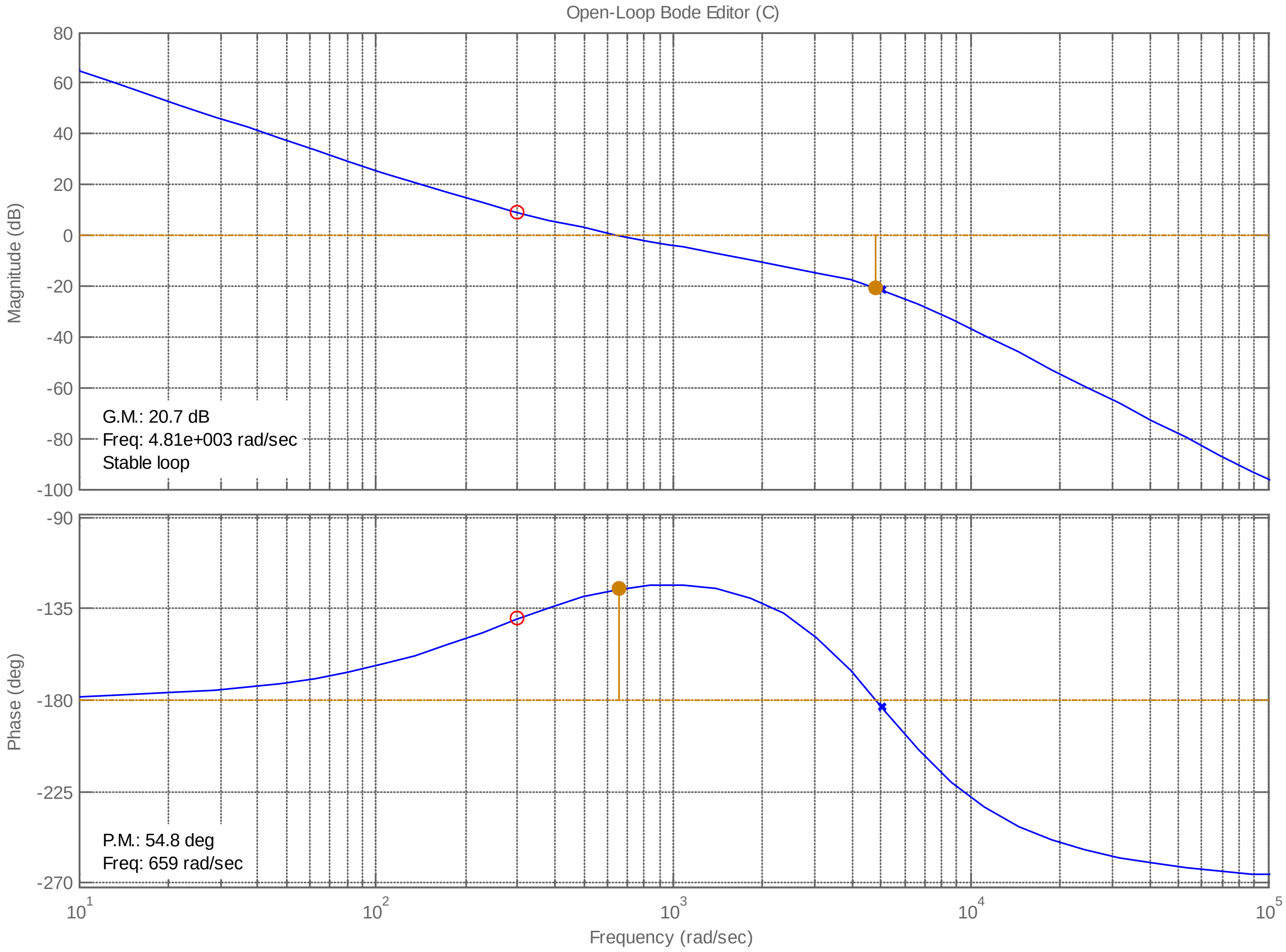
$K_F < 0.4$	high DC stiffness
$K_F = 0.65$	general
$K_F > 0.9$	fastest response

Closed-Loop Step-Response Plot

$K_I = 300$
$K_P = 1.2$
$K_F = 0.65$

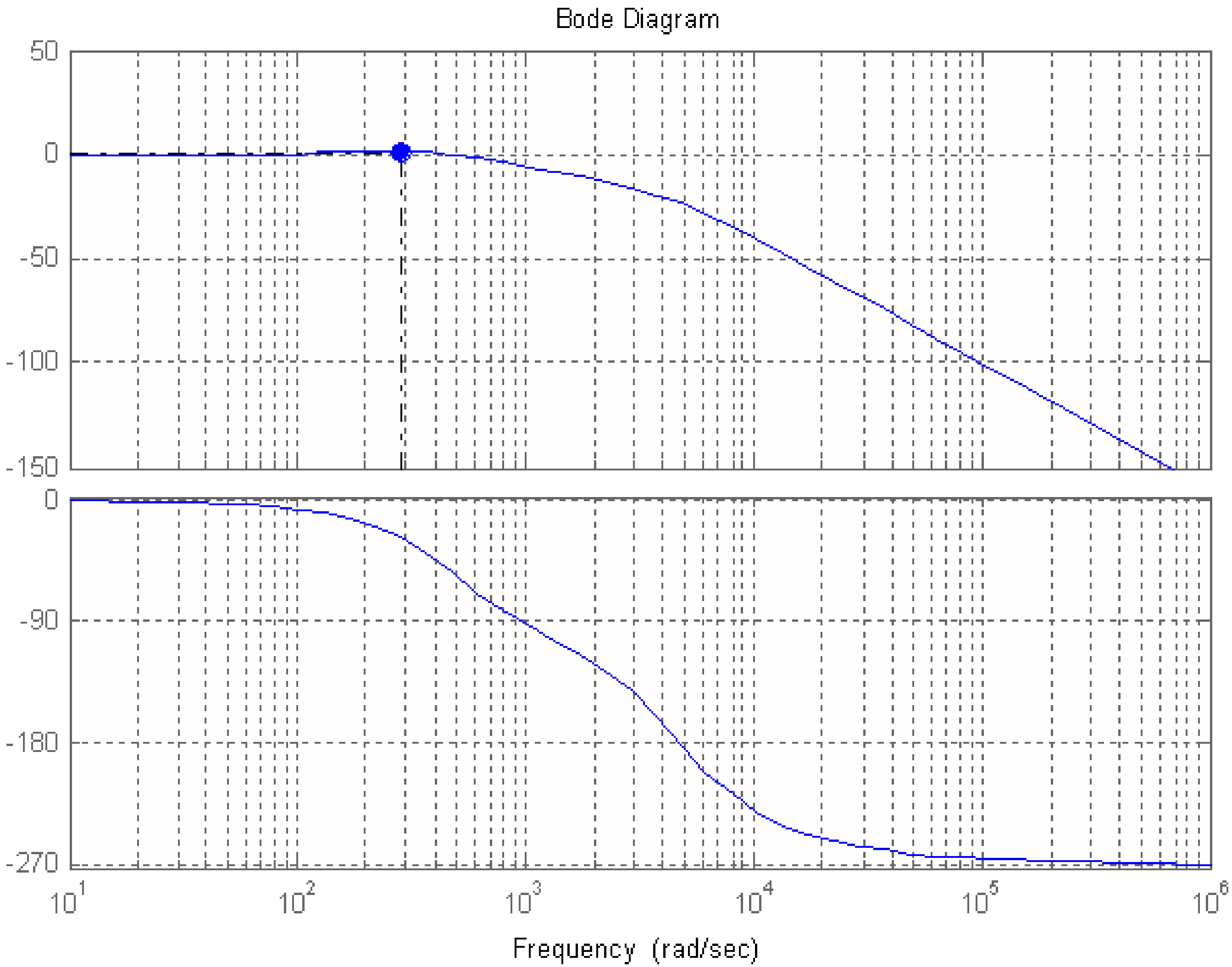


Open-Loop Bode Plots: PM = 54.8º and GM = 20.7 dB



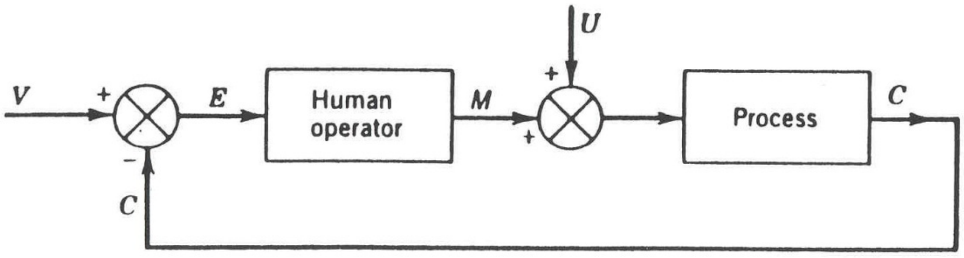
Closed-Loop Bode Plots

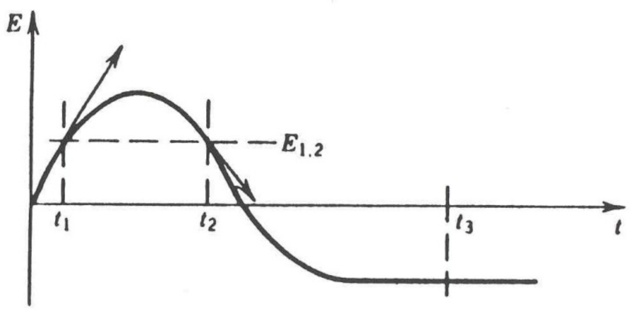
Closed-Loop Bandwidth = 731 rad/s
Peak Response = 0.841 dB @ 284 rad/s



- <u>Derivative Control</u>
- On-off, proportional, and integral control actions can be used as the sole effect in a practical controller.
- But the various derivative control modes are always used in combination with some more basic control law. This is because the derivative mode produces no corrective effect for any constant error, no matter how large, and therefore would allow uncontrolled steady-state errors.
- One of the most important contributions of derivative control is in system stability augmentation. If absolute or relative stability is the problem, a suitable derivative control mode is often the answer.
- The stabilization or "damping" aspect can easily be understood qualitatively from the following discussion.

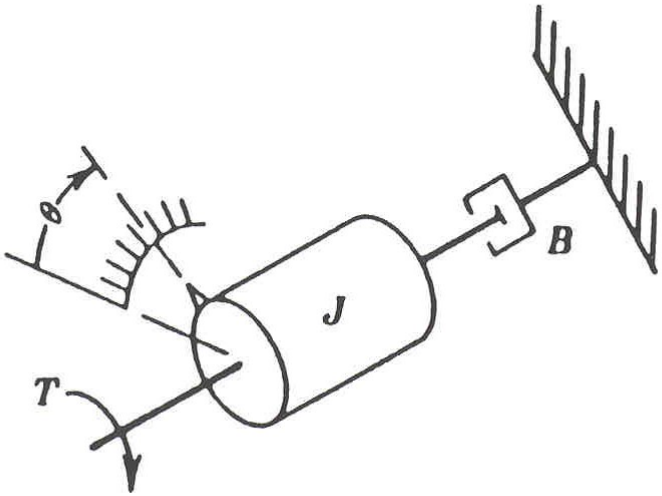
- Invention of integral control may have been stimulated by the human process operators' desire to automate their task of manual reset. Derivative control hardware may first have been devised as a mimicking of human response to changing error signals. Suppose a human process operator is given a display of system error *E* and has the task of changing manipulated variable *M* (say with a control dial) so as to keep *E* close to zero.





- If you were the operator, would you produce the same value of *M* at t_1 as at t_2? A proportional controller would do exactly that.
- A stronger corrective effect seems appropriate at t_1 and a lesser one at t_2 since at t_1 the error E is $E_{1,2}$ and increasing, whereas at t_2 it is also $E_{1,2}$ but decreasing.
- The human eye and brain senses not only the ordinate of the curve but also its trend or slope. Slope is clearly dE/dt, so to mechanize this desirable human response we need a controller sensitive to error derivative.
- Such a control can, however, not be used alone since it does not oppose steady errors of any size, as at t_3, thus a combination of proportional + derivative control, for example, makes sense.

- The relation of the general concept of derivative control to the specific effect of viscous damping in mechanical systems can be appreciated from the figure below.
- Here an applied torque T tries to control position θ of an inertia J. The damper torque on J behaves exactly like a derivative control mode in that it always opposes velocity $d\theta/dt$ with a strength proportional to $d\theta/dt$ making motion less oscillatory.



- Derivatives of *E*, *C*, and almost any available signal in the system are candidates for a useful derivative control mode.
- First derivatives are most common and easiest to implement.
- The noise-accentuating characteristics of derivative operations may often require use of approximate (low-pass filtered) derivative signals.
- Derivative signals can sometimes be realized better with sensors directly responsive to the desired value, rather than trying to differentiate an available signal.
- In addition to stability augmentation, derivative modes may also offer improvements in speed of response and steady-state errors.

- The derivative gain advances the phase of the loop by virtue of the 90° phase lead of a derivative. Using derivative gain will usually allow the system responsiveness to increase, allowing the bandwidth to nearly double in some cases.
- Derivative gain has high gain at high frequencies. So while some derivative gain does help the phase margin, too much hurts the gain margin by adding gain at the phase crossover frequency, typically a high frequency. This makes the derivative gain difficult to tune. The designer sees overshoot improve because of increased PM, but a high-frequency oscillation, which comes from reduced GM, becomes apparent.

- Derivatives are also very sensitive to noise. The derivative gain needs to followed by a low-pass filter to reduce noise content. However, the lower break frequency of the filter, the less benefit can be gained from the derivative gain.
- Proportional + Derivative Control

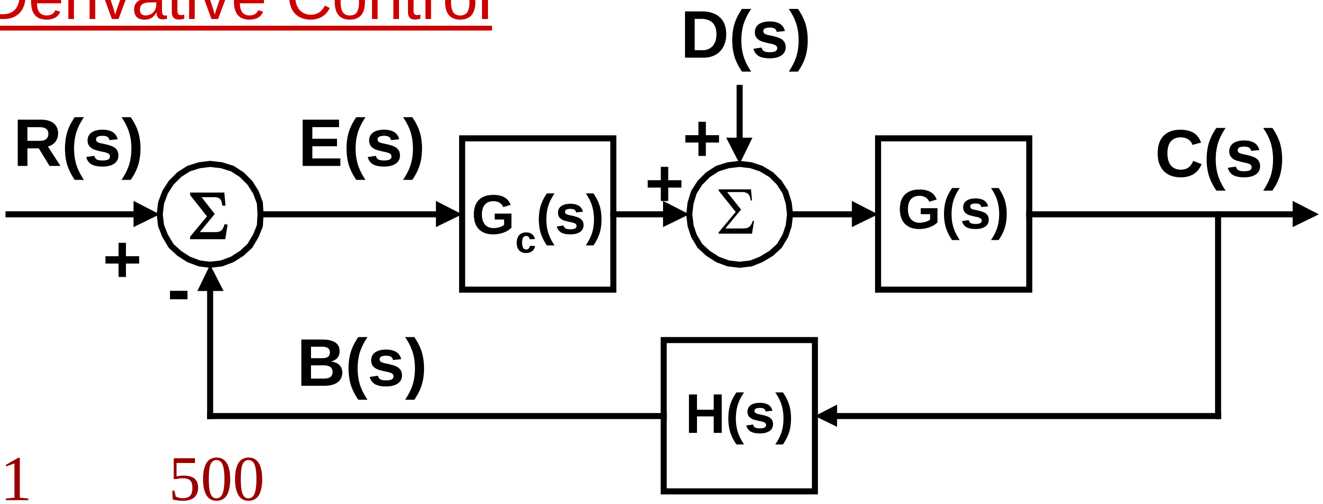
$$m(t) = K_p e(t) + K_p T_d \frac{de(t)}{dt}$$

$$\frac{M(s)}{E(s)} = K_p \left(1 + T_d s\right)$$

  – T_d = derivative time = time interval by which the rate action advances the effect of the proportional control action
- Derivative control has an anticipatory character, however, it can never anticipate any action that has not yet taken place.
- Derivative control amplifies noise signals and may cause a saturation effect in the actuator.

- Tuning a PD Controller
  - Zero K_D and set K_P low.
  - Apply a square wave command at about 10% of the desired loop bandwidth. Use large amplitude, but avoid saturation.
  - Raise K_P for some overshoot, but no ringing.
  - Raise K_D to eliminate most overshoot.
  - If it is too noisy, reduce noise at the source or lower K_D or lower K_P.

Proportional-Derivative Control



$$G(s) = \frac{1}{Js} = \frac{1}{0.002s} = \frac{500}{s}$$

$$G_{amp}(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

$$\omega = 5027\ (\text{rad/s}) = 800\ (\text{Hz})$$

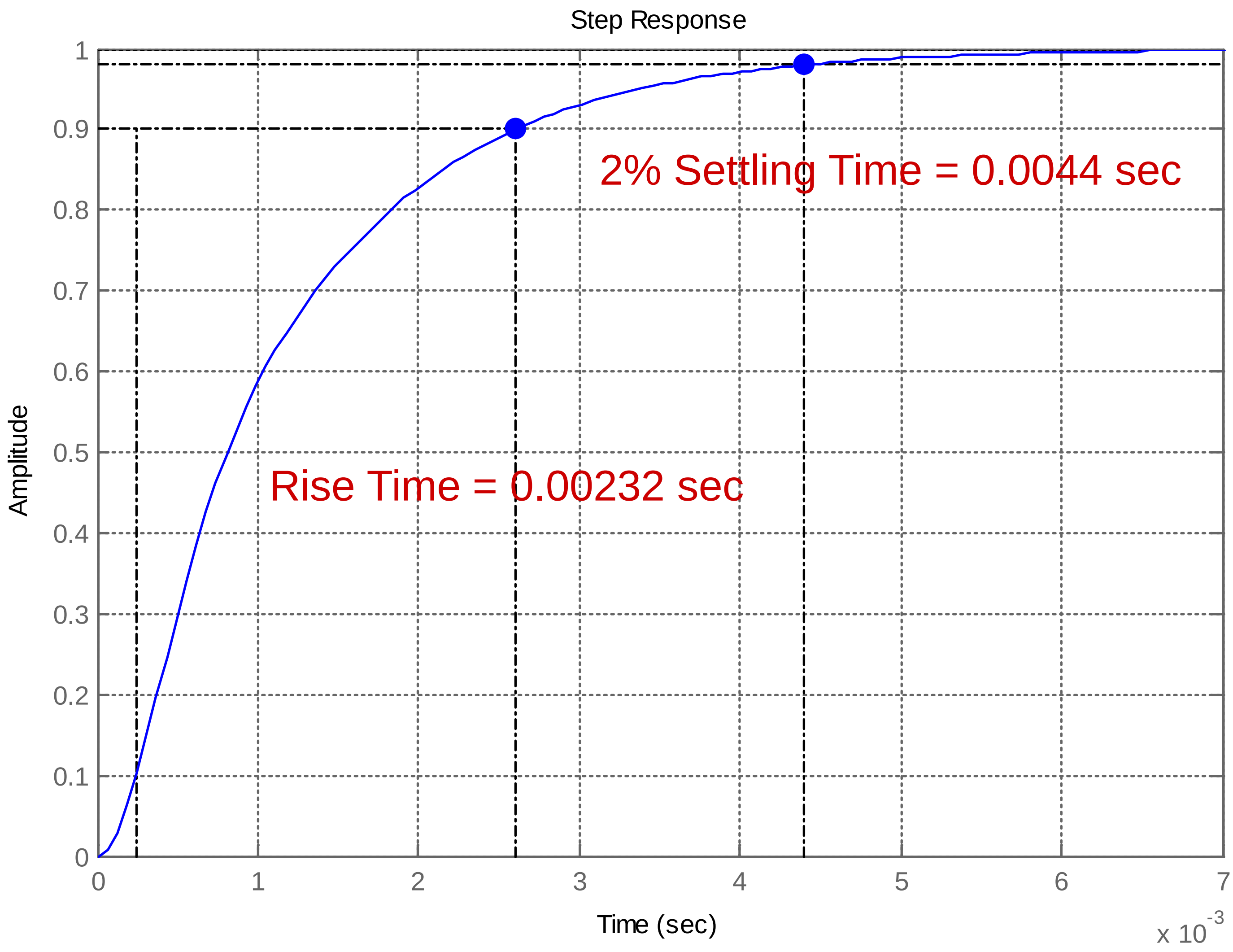
$$\zeta = 0.707$$

$$K_t = 1\ (\text{Nm/A})$$

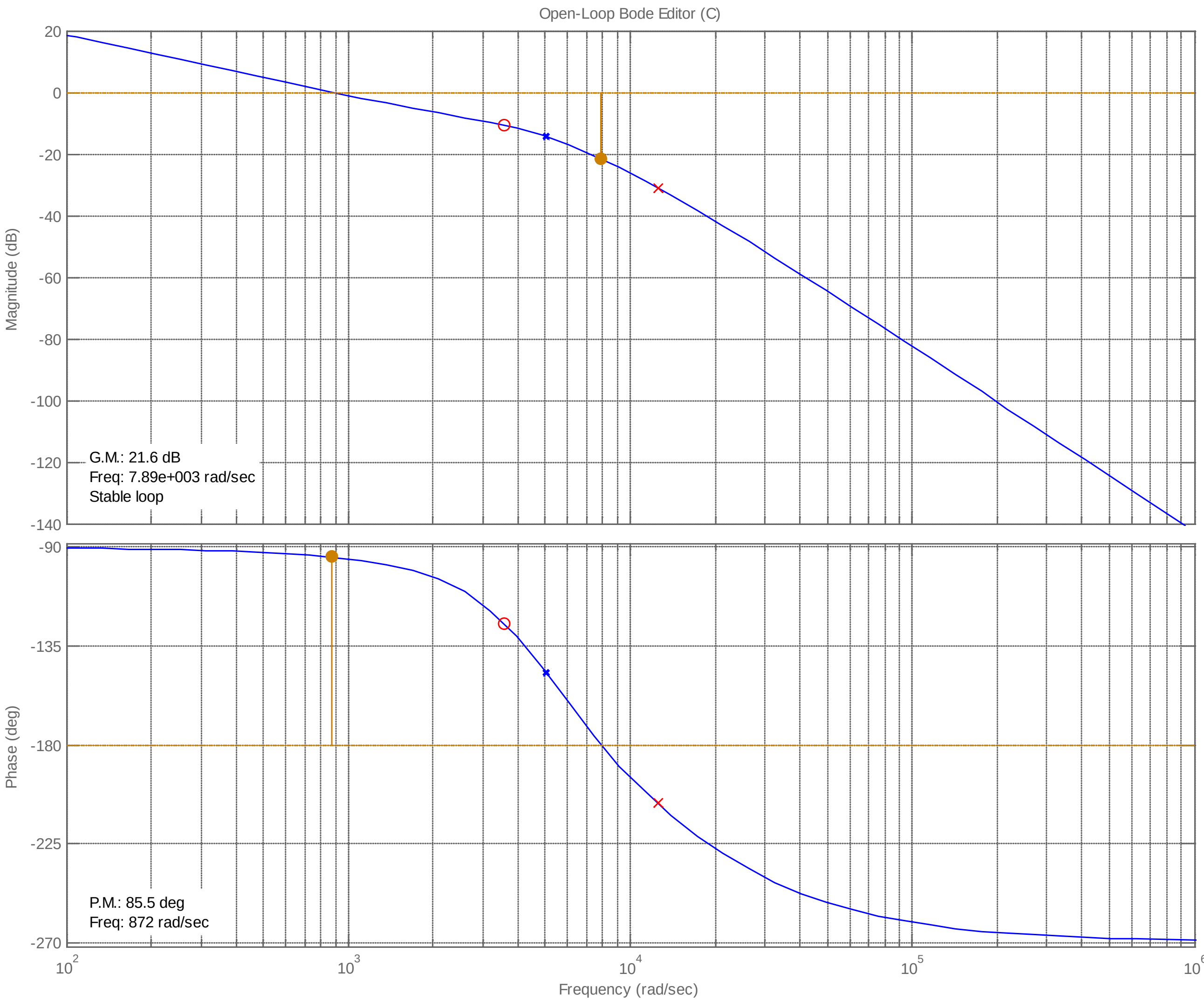
$$H(s) = 1$$

$$G_C(s) = K_P\left(1 + \frac{K_D s}{\tau s + 1}\right) = 1.7\left(1 + \frac{0.0002s}{8.0\left(10^{-5}\right)s + 1}\right)$$

Closed-Loop Step-Response Plot

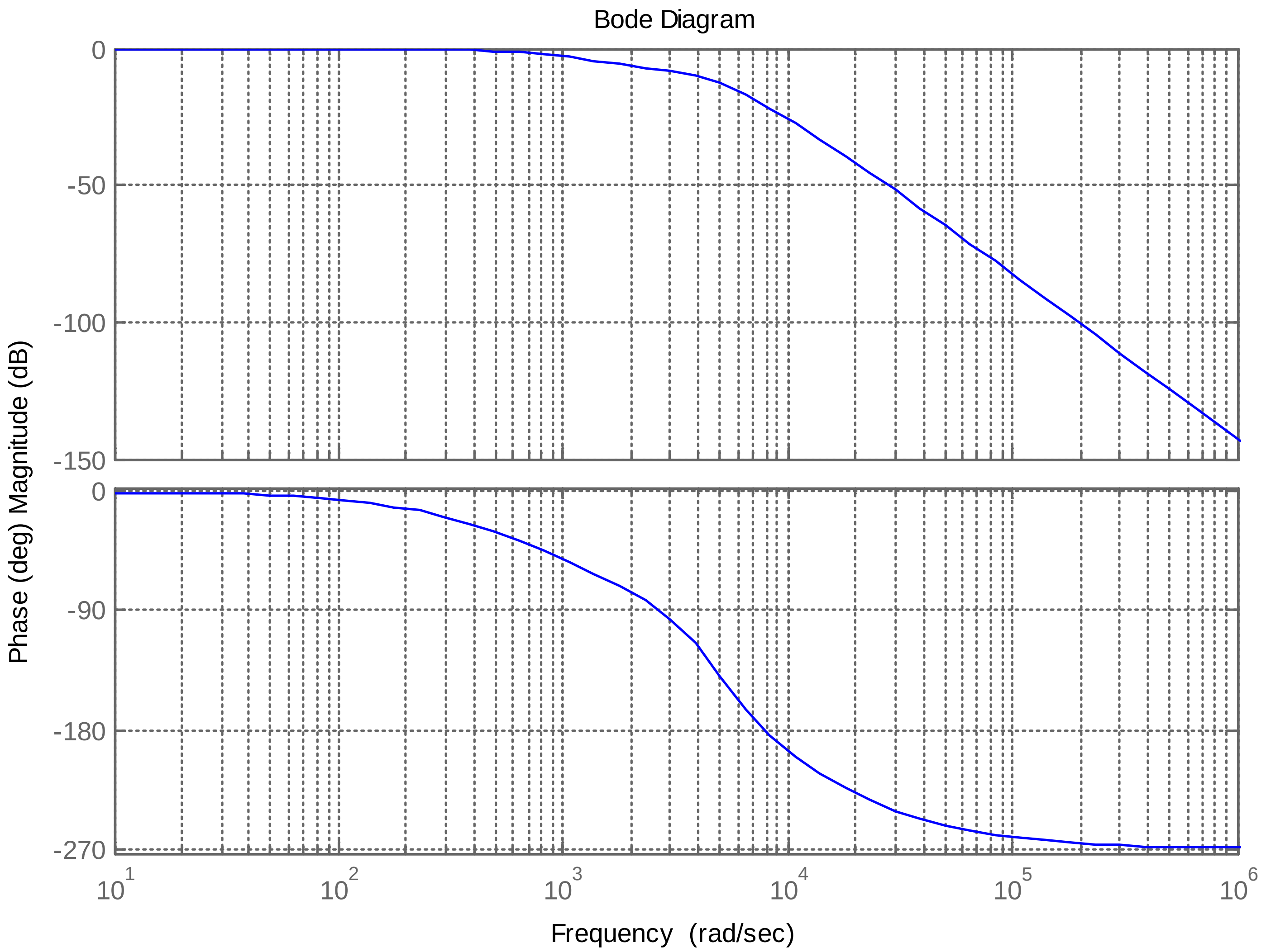


Open-Loop Bode Plots: PM = 85.8º and GM = 21.6 dB



Closed-Loop Bode Plots

Closed-Loop Bandwidth = 940 rad/s



- Proportional + Integral + Derivative (PID) Control

$$m(t) = K_p e(t) + \frac{K_p}{T_i}\int_0^t e(\tau)d\tau + K_p T_d \frac{de(t)}{dt}$$

$$\frac{M(s)}{E(s)} = K_p\left(1 + \frac{1}{T_i s} + T_d s\right)$$

  – About 90% of the industrial controllers is use today utilize PID or modified PID control schemes.
  – Analog PID controllers are mostly hydraulic, pneumatic, electric, and electronic types or their combinations.

$$G_c(s) = K_P\left(1 + K_D s + \frac{K_I}{s}\right) = K_P\left(\frac{K_D s^2 + s + K_I}{s}\right)$$

- Digital PID control through the use of microprocessors is becoming more common.
- Most PID controllers are adjusted on site and many different tuning rules have been proposed.
- PID controllers with automatic tuning, automatic generation of gain schedules, and continuous adaptation are now possible.
- Modified forms of PID control (e.g., I-PD and two-degrees-of-freedom PID control) are currently in use in industry.
- Many practical methods for bumpless switching (from manual operation to automatic operation) and gain scheduling are commercially available.
- Usefulness of PID control lies in its general applicability to most control systems.

- How to Tune a PID Controller
  - A PID controller is a two-zone controller. The P and D gains jointly form the higher frequency zone. The I gain forms the low-frequency zone.
  - The benefit of the D gain is that it allows the P gain to be set higher than it could be otherwise.
  - The first step is to tune the controller as if it were a P controller, but to allow more overshoot than normal, understanding that the D gain will cure the problem. Typically, the P gain can be raised 25% to 50% over the value from the P and PI controllers.
  - The next step is to add a little D gain to cure the overshoot induced by the higher than normal P gain.
  - The P & D gains together form the high-frequency zone.

- Next, the I gain is tuned, much as it was in the PI controller. The expectation is that the P and I gains will be about 30% higher than they were in the PI controller.
- The phase margin compared to that of a PI controller will be about the same, however, the gain margin will be less because the high-frequency zone of the PID controller is so much higher than that of the PI controller as evidenced by the higher bandwidth of the PID controller. Reduced gain margin is a concern because the gains of plants often change during normal operation and is of particular concern in systems where the gain can increase, e.g., saturation of an inductor in a current controller, declining inertia in a motion system, or declining thermal mass in a temperature controller. These all raise the gain of the plant and reduce GM.

- Given the same plant and amplifier, a PID controller will provide faster response than a PI controller but will often be harder to control and more sensitive to changes in the plant.
- The problems with noise in the PI controller are exacerbated by the use of a differential gain, as the gain of a true derivative increases without bound as the frequency increases. In most working systems, a low-pass filter is placed in series with the derivative to limit gain at the highest frequencies. If the noise content of the feedback or command signals is high, the best cure is to reduce the noise at its source. Beyond that, lowering the frequency of the derivative's low-pass filter will help, but it will also limit the effectiveness of the D gain.

- Noise can also be reduced by reducing the D gain directly, but this is usually a poorer alternative than lowering the low-pass filter frequency.
- If the signal is too noisy, the D gain may need to be abandoned altogether.

• Ziegler-Nichols Tuning Method
  - A popular method for tuning P, PI, and PID controllers is the Ziegler-Nichols method.
  - This method starts by zeroing the I and D gains and then raising the P gain until the system is unstable.
  - The value of K_p at the point of instability is called K_{MAX}; the frequency of oscillation is called f_0 (rad/s).

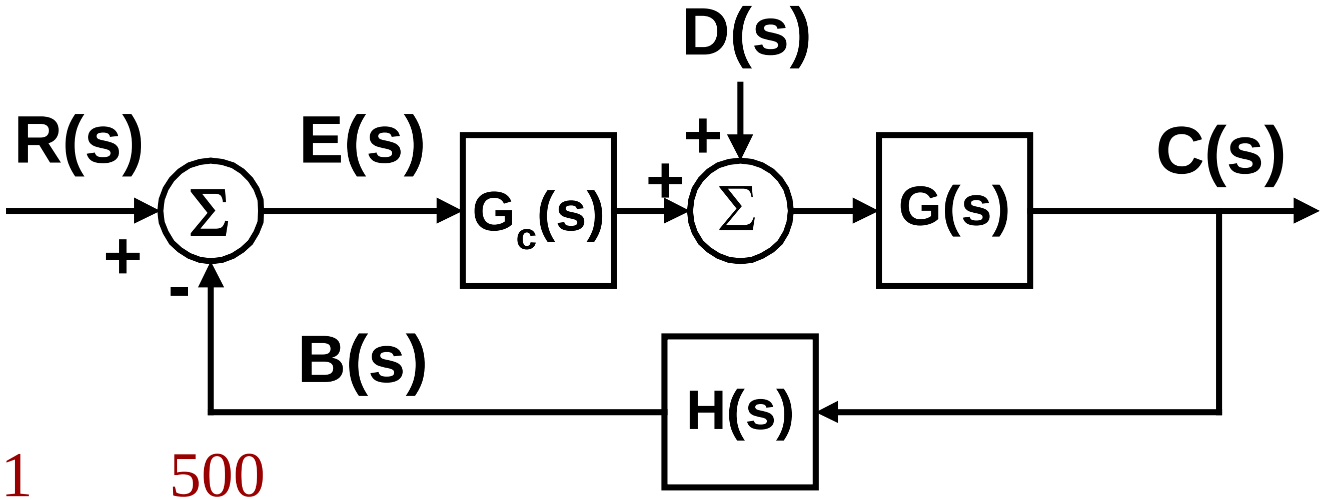
– The method then backs off the P gain a predetermined amount and sets the I and D gains as a function of f_0.

	K_P	K_I	K_D
P	$0.5\ K_{MAX}$	0	0
PI	$0.45\ K_{MAX}$	$1.2\ f_0$	0
PID	$0.6\ K_{MAX}$	$2.0\ f_0$	$0.125\ /\ f_0$

– It is assumed that K_P is in series with K_I and K_D.

– The Ziegler-Nichols method is too aggressive (small stability margins) for many industrial control systems. Also, in general, the gains from the Ziegler-Nichols method will be higher than from then other methods discussed here.

Proportional-Integral-Derivative Control



$$G(s) = \frac{1}{Js} = \frac{1}{0.002s} = \frac{500}{s}$$

$$G_{amp}(s) = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

$$\omega = 5027 \text{ (rad/s)} = 800 \text{ (Hz)}$$

$$\zeta = 0.707$$

$K_P = 1.7$

$K_I = 160$

$K_D = 0.0002$

$$\frac{1}{\tau} = 2000 \text{ Hz}$$

$$= 12566 \text{ rad/s}$$

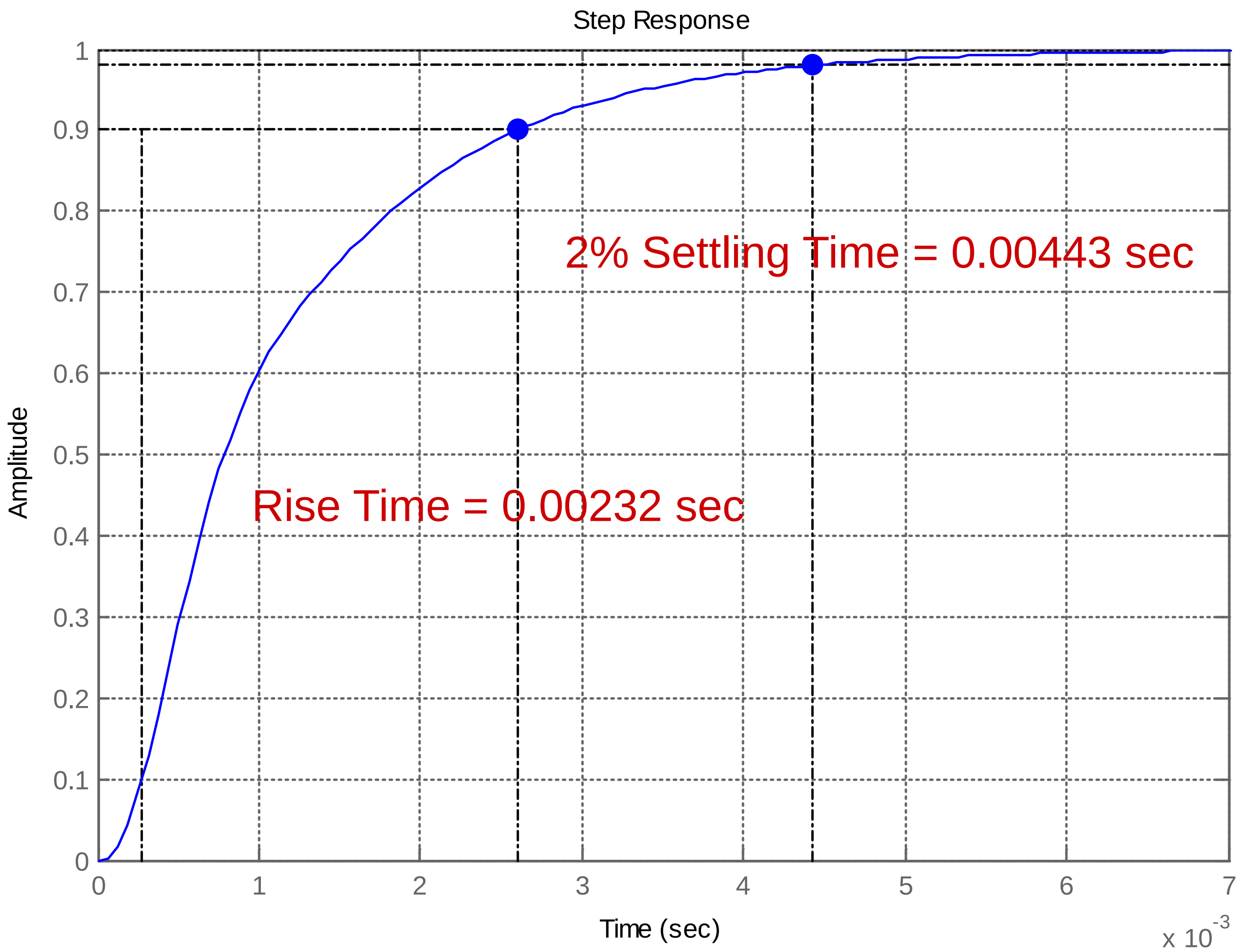
$$K_t = 1 \text{ (Nm/A)}$$

$$G_C(s) = K_P\left(1 + \frac{K_D s}{\tau s + 1} + \frac{K_I}{s}\right)$$

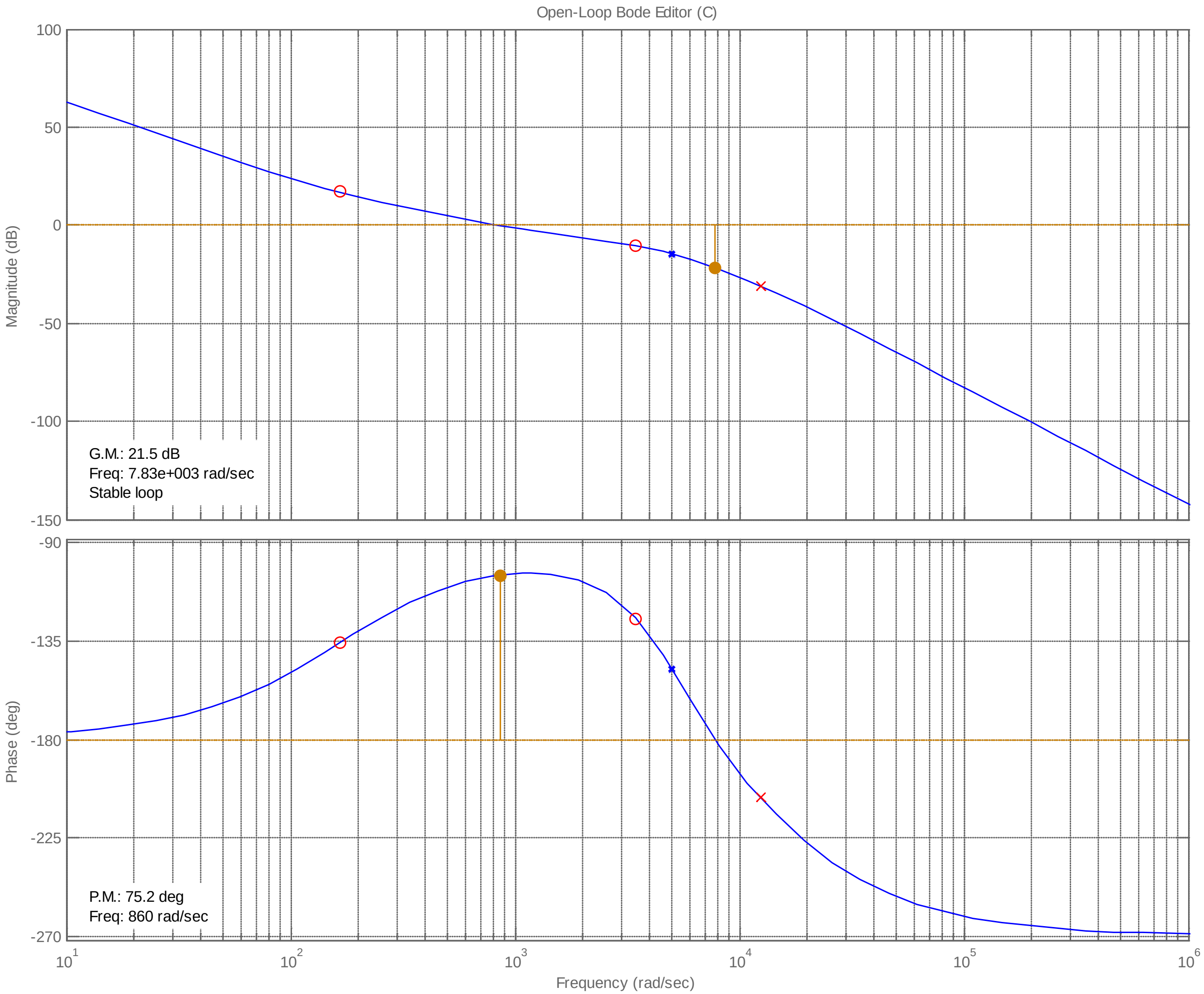
$$H(s) = 1$$

$$= K_P\left(\frac{(\tau + K_D)s^2 + (K_I\tau + 1)s + K_I}{s(\tau s + 1)}\right)$$

Closed-Loop Step-Response Plot

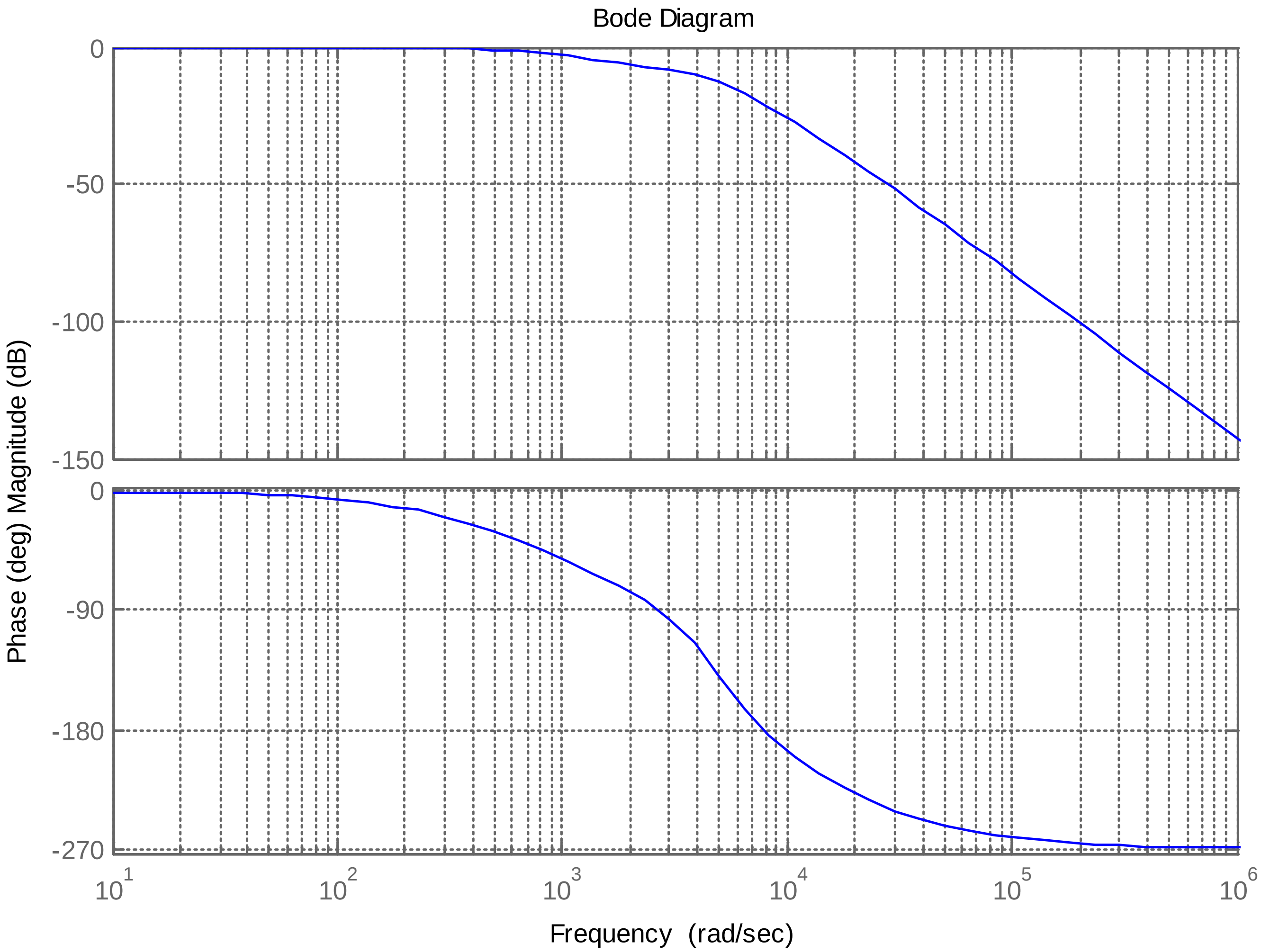


Open-Loop Bode Plots: PM = 75.2º and GM = 21.5 dB



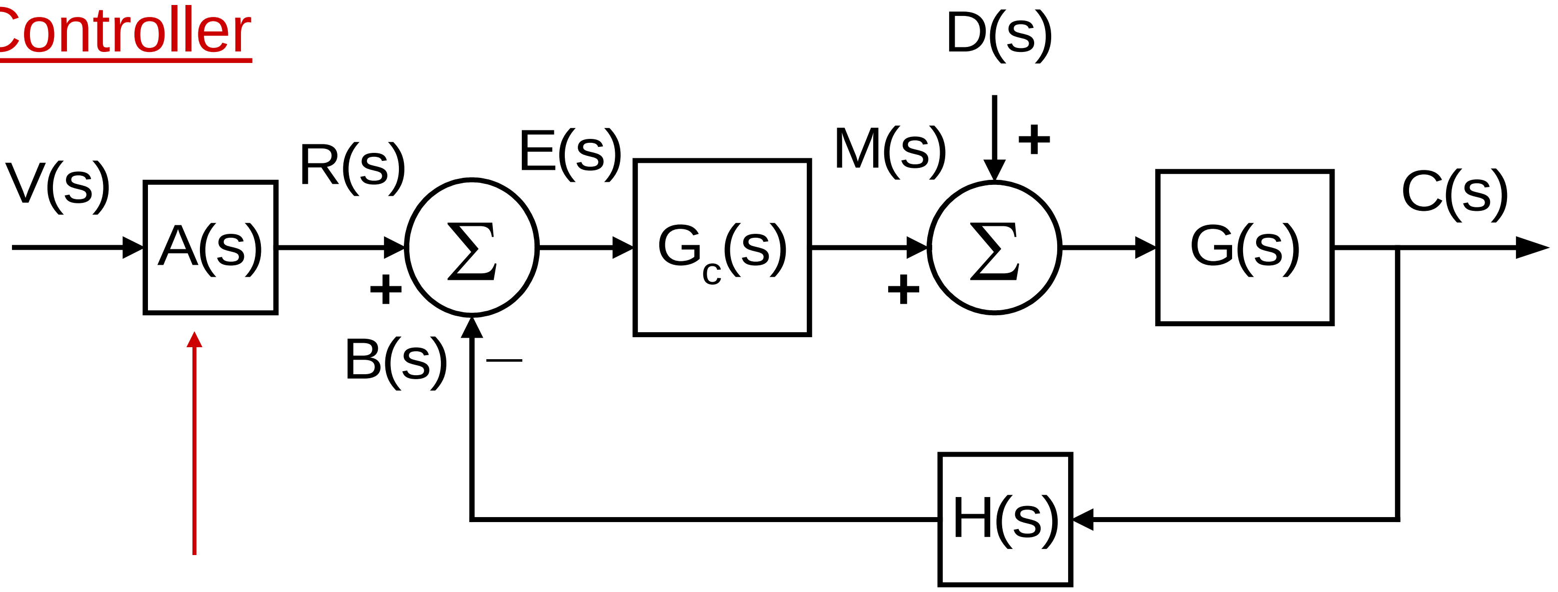
Closed-Loop Bode Plots

Closed-Loop Bandwidth = 942 rad/s



- PID+ Control
  - A PID+ controller is a PID controller modified with the command filter.
  - As with the PI+ controller, the goal for the PID+ controller is to allow higher integral gains for improved DC stiffness.
  - Again, the input filter cancels peaking caused by high integral gains.
  - As with PI+ control, the command response suffers as the stiffness improves.
  - Tuning a PID+ controller is the same as tuning a PID controller except that the value of K_F must be selected before tuning the I gain (similar to the PI+ controller).

PID+ Controller



Command Filter

$$A(s) = K_F + (1 - K_F)\frac{K_I}{s + K_I}$$

$K_F = 1$: all filtering removed

$K_F = 0$: filtering most severe

$K_F < 0.4$ high DC stiffness

$K_F = 0.65$ general

$K_F > 0.9$ fastest response

$$G_C(s) = K_P\left(1 + \frac{K_D s}{\tau s + 1} + \frac{K_I}{s}\right)$$

$$= K_P\left(\frac{(\tau + K_D)s^2 + (K_I\tau + 1)s + K_I}{s(\tau s + 1)}\right)$$

$K_I = 400$

$K_I = 1.7$

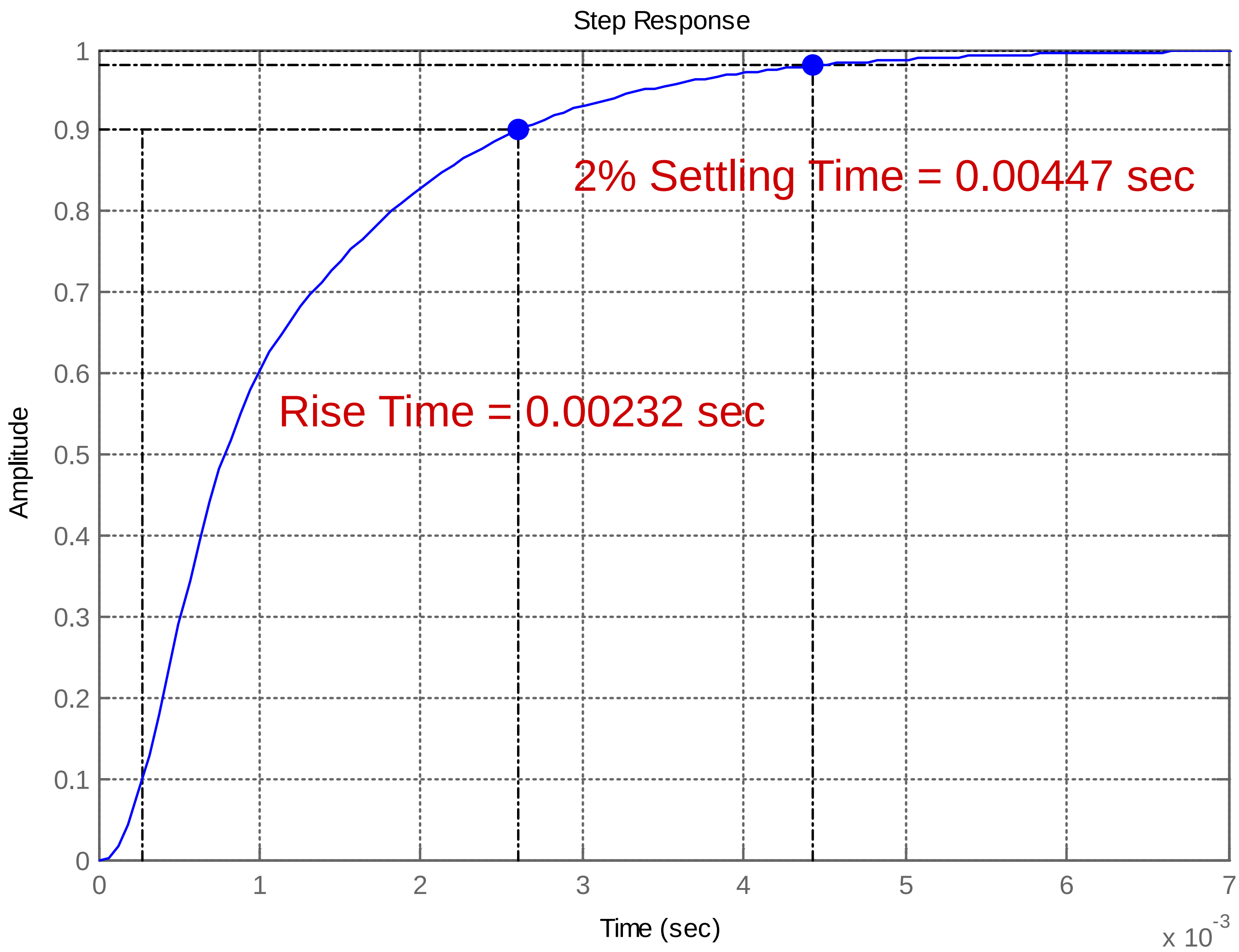
$K_D = 0.0002$

$K_F = 0.65$

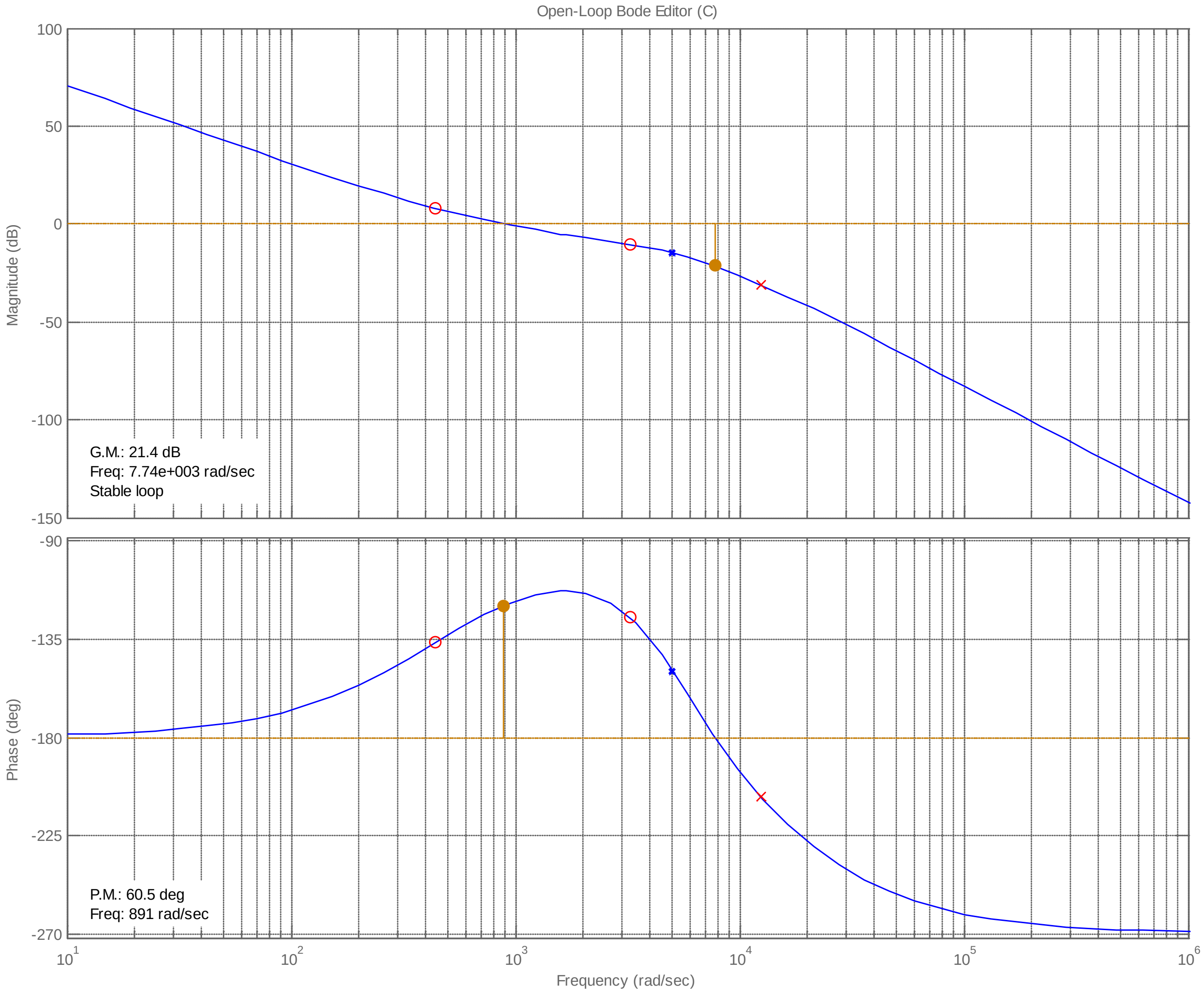
$$\frac{1}{\tau} = 2000 \text{ Hz}$$

$$= 12566 \text{ rad/s}$$

Closed-Loop Step-Response Plot

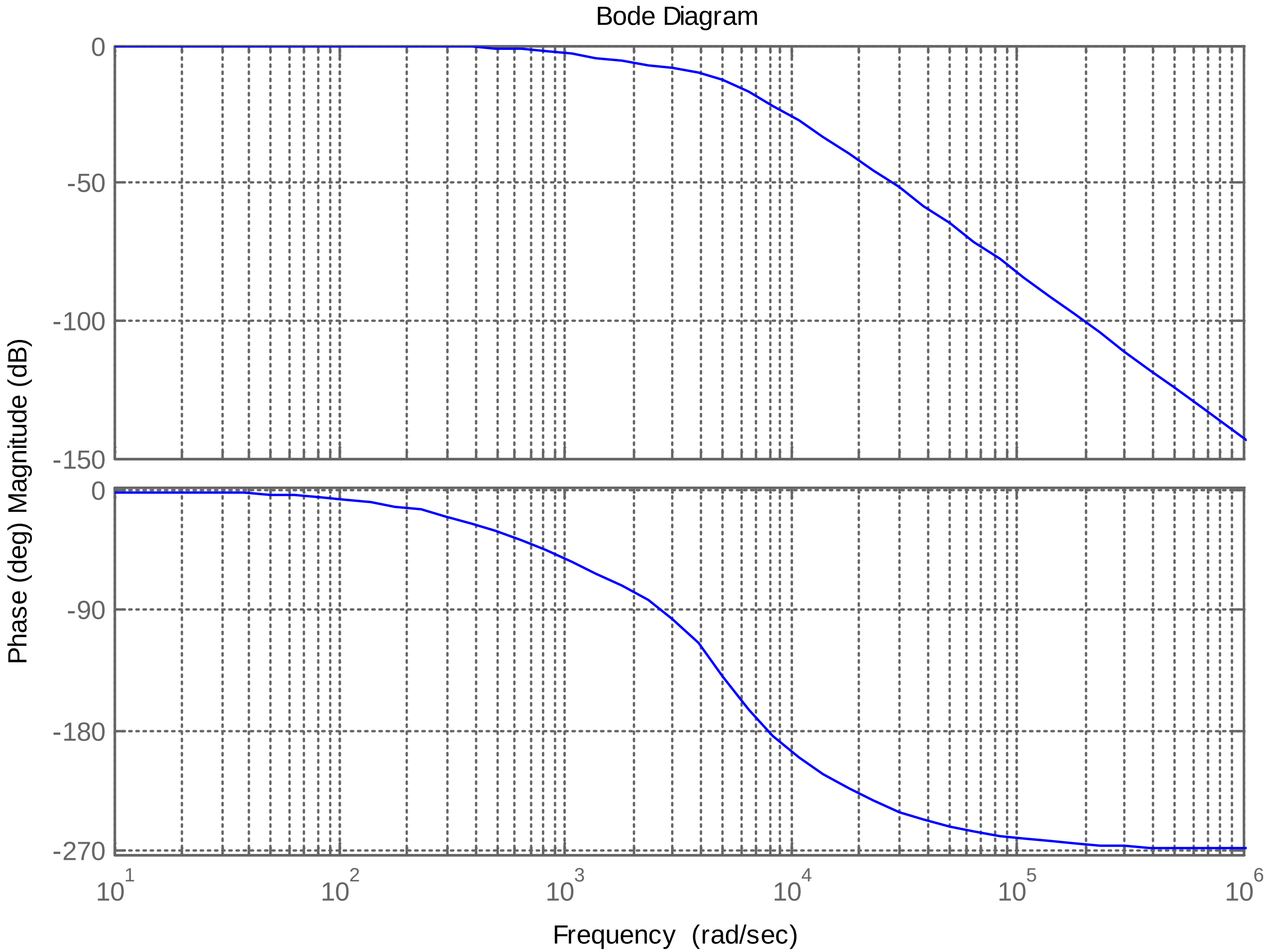


Open-Loop Bode Plots: PM = 60.5º and GM = 21.4 dB



Closed-Loop Bode Plots

Closed-Loop Bandwidth = 954 rad/s



- Choosing A Controller
  - The simple P controller provides performance suitable for many applications.
  - The introduction of the I term provides DC stiffness but reduces PM.
  - The command filter in PI+ and PID+ control allows even higher DC stiffness but reduces bandwidth.
  - The D term provides higher responsiveness but erodes gain margin and adds phase shift, which is a disadvantage if this loop is to be enclosed in an outer loop.
  - In selecting a controller, first determine whether the application needs the D gain; if not, avoid it as it adds complexity, increases noise susceptibility, and steals GM.

- Next, make sure the application can support D gains; systems that are noisy configured as PI controllers may not work well with a D gain.
- After that, examine the application for the needed DC stiffness. If none is required, avoid the I gain. If some is needed, use the standard form PI; if maximum DC stiffness is required, add the input filter by using PI+ control.

Approximate Control Modes

- Proportional + Integral (PI)
  - Phase-Lag Compensation
- Proportional + Derivative (PD)
  - Phase-Lead Compensation
- Proportional + Integral + Derivative (PID)
  - Lead / Lag Compensation

- We have introduced the basic control modes: on-off, proportional, integral, and derivative. Each of these has its own advantages and drawbacks, and thus it is not surprising that many practical applications are best served by some combination of basic modes.
- We have also considered the most basic or idealized versions of the modes so that their essential features could be brought out most clearly without confusing side issues. Practical versions of some controllers are not able to realize completely the ideal behavior and also may require a modified design technique. Sometimes a non-ideal controller can meet specifications with simpler hardware or software. For these reasons, approximate forms of control modes should be considered.

- Phase-Lag Compensation
  - PI control provides the steady-state-error benefits of pure integral control with faster response and improved stability.
  - Phase-Lag Compensation is the approximate version of PI Control realized in many practical controllers. It cannot attain the zero steady-state errors possible with perfect integral control but this is not a fatal defect because realistic error specifications always must allow some steady-state error.

- Phase-Lead Compensation
  - Since derivative control is never used alone and we have already briefly discussed PD control, let's concentrate on the approximate version, phase-lead compensation.
  - If a basic system has had its gain set for desired relative stability and we then find that its response speed is too slow, phase-lead compensation may be helpful. Also, if a basic system is structurally unstable (gain setting does not provide stability), phase-lead compensation may stabilize the system. Usually, phase-lead compensation also provides a modest gain increase, so steady-state errors are reduced whether this was a problem or not.

- Proportional + Integral + Derivative (PID) Control
  - This combination of basic control modes can improve all aspects (stability, speed, steady-state errors) of system performance and is the most complex method available as an off-the-shelf general-purpose controller. If we look at analog pneumatic and electronic controllers, their microprocessor-based digital versions, or the individual control loops implemented in a large general-purpose digital process computer, over and over again we see successful applications of P, PI, PD, and PID controls. The basis of the strength of the PID modes is their simplicity; they "make sense."

- Lag/Lead Compensation
  - The approximate version of PID control implemented in many practical controllers is called lag/lead compensation. Mathematically it is exactly a cascading of the phase-lag and phase-lead controllers already discussed.
  - The effects on system performance are also a superposition of the two separate effects, thus a lag/lead controller can improve all aspects of performance (as can a PID): stability, speed, and steady-state errors. Selection of the parameters is performed by essentially designing the two compensators separately.