Root Locus Analysis & Design

- A designer would like:
  - To know if the system is absolutely stable *and* the degree of stability.
  - To predict a system's performance by an analysis that does NOT require the actual solution of the differential equations.
  - The analysis to indicate readily the manner or method by which this system must be adjusted or compensated to produce the desired performance characteristics.

- Two Methods are available:
  – Root-Locus Approach
  – Frequency-Response Approach
- Root Locus Approach
  – Basic characteristic of the transient response of a closed-loop system is closely related to the location of the closed-loop poles.
  – If the system has a variable loop gain, then the location of the closed-loop poles depends on the value of the loop gain chosen.

- It is important to know how the closed-loop poles move in the *s* plane as the loop gain is varied.
- From a Design Viewpoint:
  - Simple gain adjustment may move the closed-loop poles to desired locations. The design problem then becomes the selection of an appropriate gain value.
  - If gain adjustment alone does not yield a desired result, addition of a compensator to the system is necessary.
- The closed-loop poles are the roots of the closed-loop system characteristic equation.

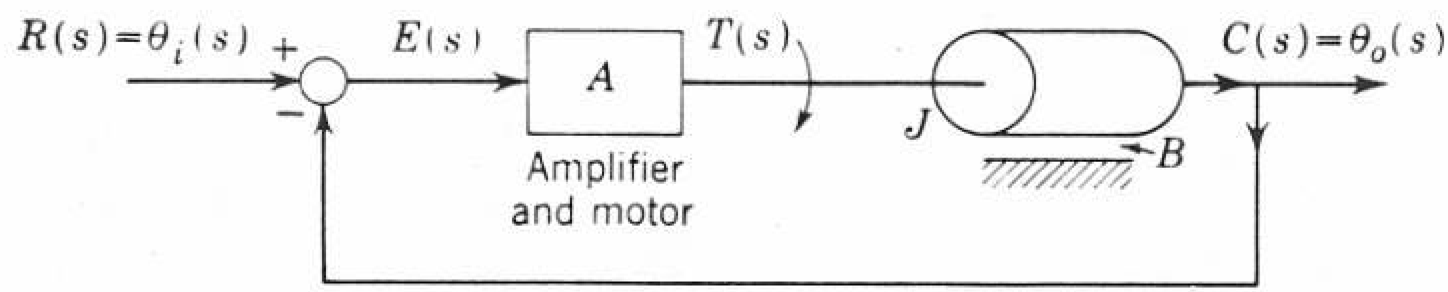
- *The Root Locus Plot* is a plot of the roots of the characteristic equation of the *closed-loop system* for all values of a system parameter, usually the gain; however, any other variable of the open-loop transfer function may be used.
- By using this method, the designer can predict the effects on the location of the closed-loop poles of varying the gain value OR adding open-loop poles and/or open-loop zeros.
- A designer MUST know how to generate the root loci of the closed-loop system BOTH by hand and with a computer (e.g., MatLab).

– Experience in sketching the root loci by hand is invaluable for interpreting computer-generated root loci, as well as for getting a rough idea of the root loci very quickly.

- Underlying Principle:
  - Poles of the closed-loop transfer function are related to the zeros and poles of the open-loop transfer function and also to the gain.
  - The values of *s* that make the open-loop transfer function equal to –1 must satisfy the characteristic equation of the closed-loop system.
  - The root-locus plot clearly shows the contributions of each open-loop pole and zero to the locations of the closed-loop poles.

– The root-locus plot also shows the manner in which the open-loop poles and zeros should be modified so that the response meets system performance specifications.

- Example: **Motor Position Control System**



$$G(s)=\frac{\theta_o(s)}{E(s)}=\frac{\frac{A}{J}}{s\left(s+\frac{B}{J}\right)}=\frac{K}{s(s+a)}$$

Assume a = 2.

$$\frac{C(s)}{R(s)}=\frac{K}{s(s+2)+K}=\frac{K}{s^2+2s+K}=\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$$

$$\omega_n=\sqrt{K} \qquad \zeta=\frac{1}{\sqrt{K}}$$

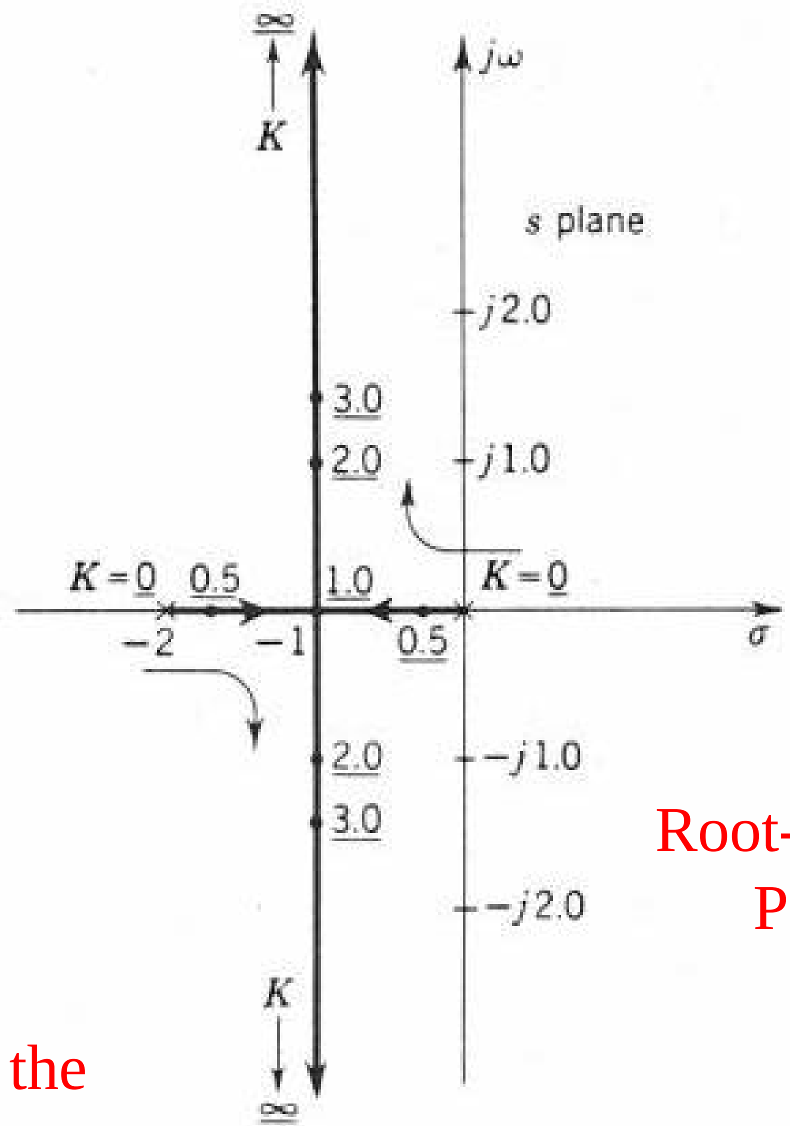
Problem:

Determine the roots of the characteristic equation for all values of K and plot these roots in the *s* plane.

Roots of the characteristic equation are given by:

$$s_{1,2} = -1 \pm \sqrt{1-K}$$

Note:
s = -2 and s = 0 are the open-loop poles



Root-Locus Plot

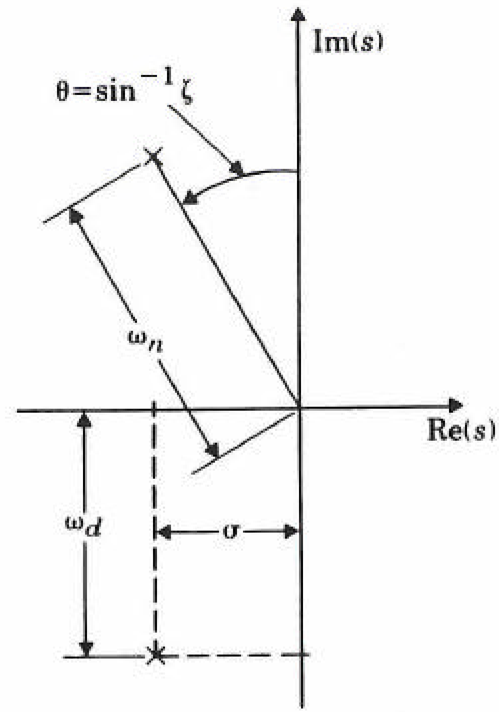
– For 0 < K < 1, the roots are real and lie on the real axis. For K > 1, the roots are complex.

– Once the root-locus plot has been obtained, it is possible to determine the variation in system performance with respect to a variation in K.

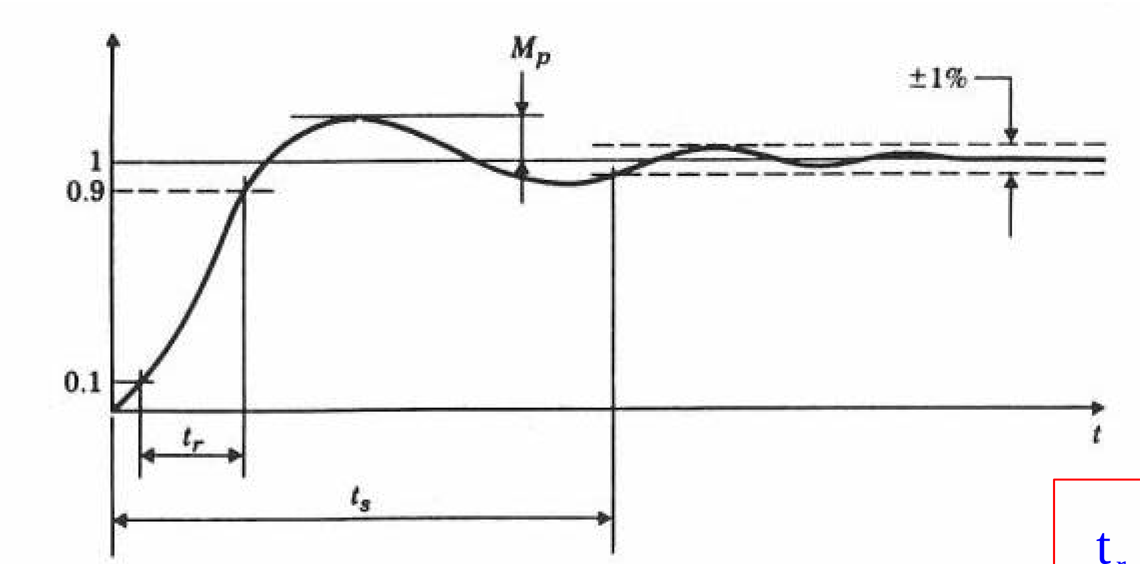
$$\frac{C(s)}{R(s)} = \frac{\omega_n^2}{s^2 + 2\varsigma\omega_n s + \omega_n^2}$$

$$s_{1,2} = -\varsigma\omega_n \pm i\omega_n\sqrt{1-\varsigma^2}$$

$$s_{1,2} = -\sigma \pm i\omega_d$$



- As K is increased from K = 1, we observe:
  - A decrease in the damping ratio ζ. This increases the overshoot of the time response.
  - An increase in the undamped natural frequency ω_n.
  - An increase in the damped natural frequency ω_d.
  - No effect on the rate of decay σ.
  - No matter how much the gain is increased in this simple linear second-order system, the system can never become unstable.



Second-Order System Unit Step Response

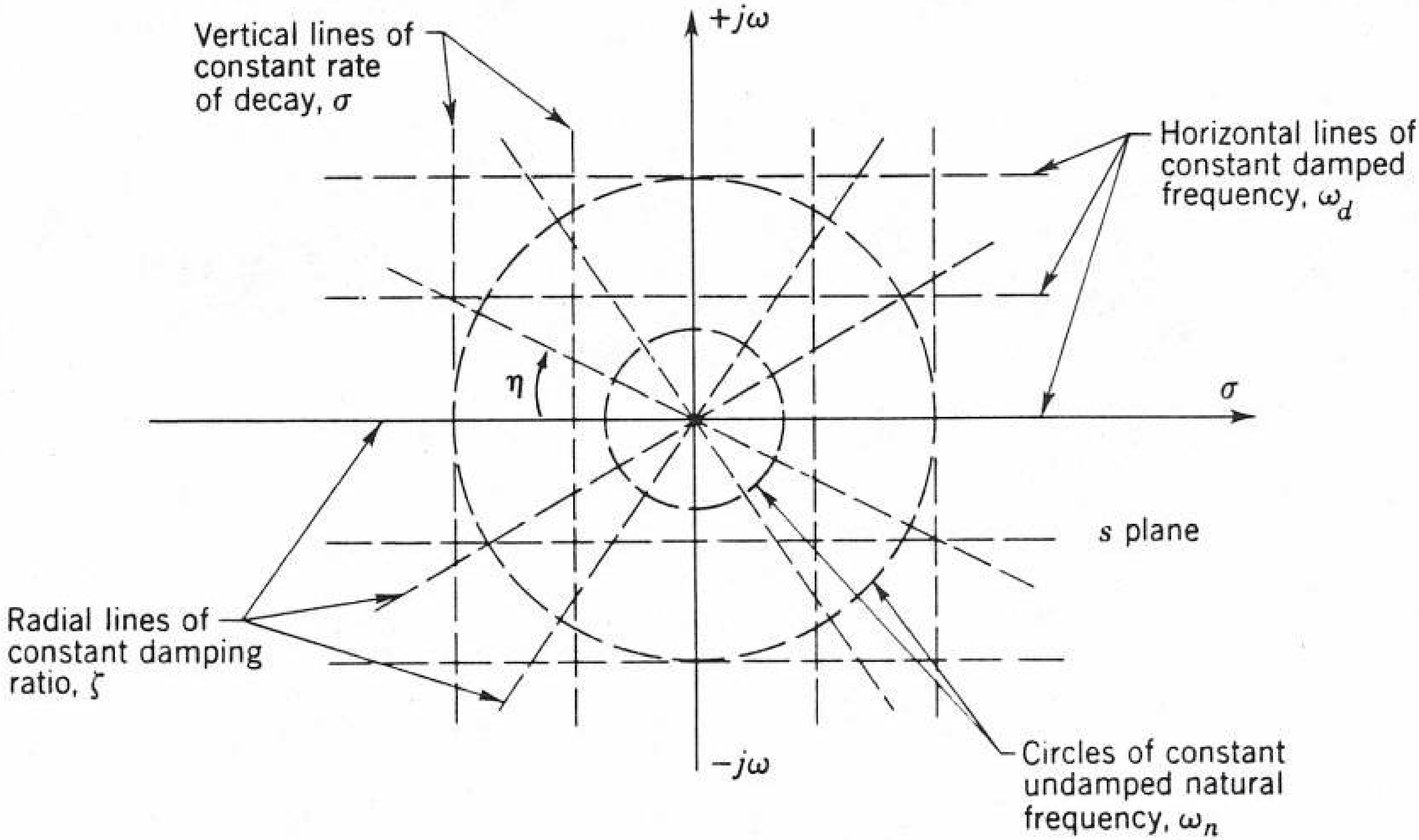
$$y(t) = 1 - e^{-\sigma t}\left(\cos\omega_d t + \frac{\sigma}{\omega_d}\sin\omega_d t\right)$$

$$t_r \approx \frac{1.8}{\omega_n} \quad \text{rise time}$$

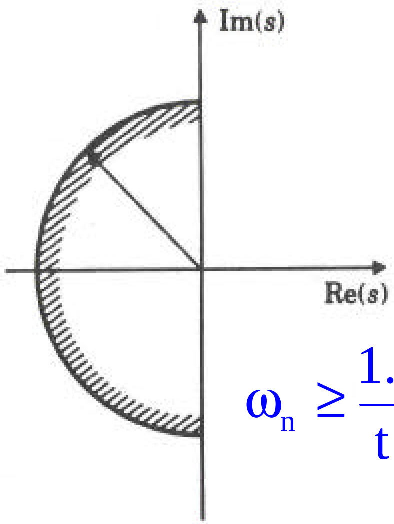
$$t_s \approx \frac{4.6}{\varsigma\omega_n} \quad \text{settling time}$$

$$M_p = e^{\frac{-\pi\varsigma}{\sqrt{1-\varsigma^2}}} \quad (0 \le \zeta < 1) \quad \text{overshoot}$$

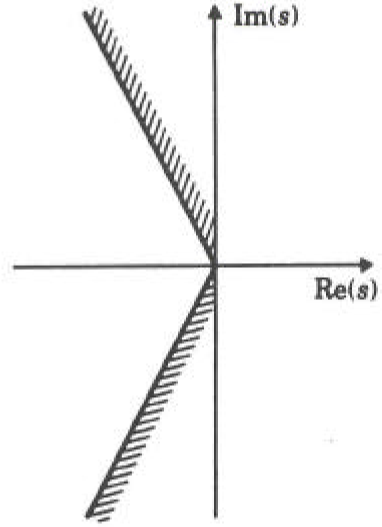
$$= \left(1 - \frac{\zeta}{0.6}\right) \quad (0 \le \zeta \le 0.6)$$

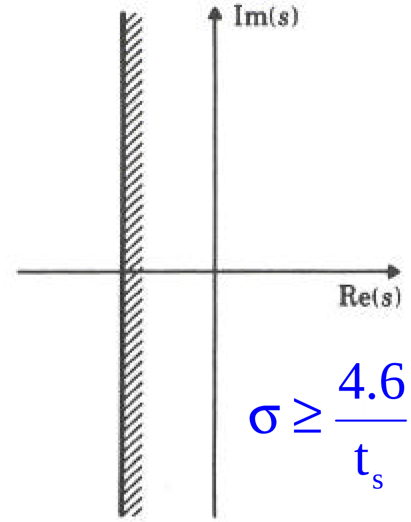


Constant Parameter Curves on the S Plane



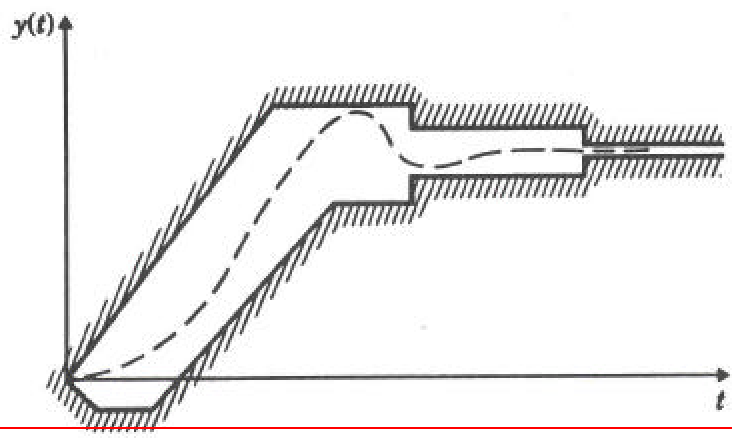
$$\omega_n \geq \frac{1.8}{t_r}$$

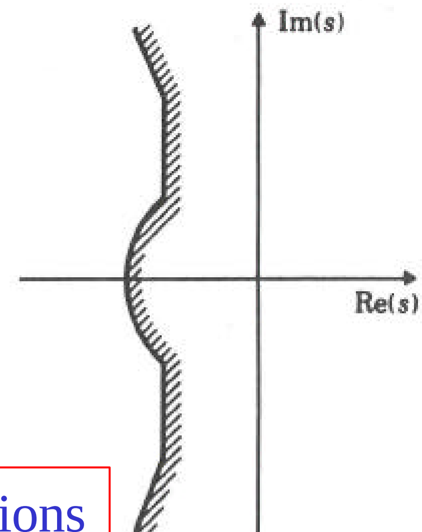




$$\sigma \geq \frac{4.6}{t_s}$$

$$\zeta \geq 0.6\left(1 - M_p\right) \qquad 0 \leq \zeta \leq 0.6$$





Time-Response Specifications vs. Pole-Location Specifications

- General effects of the addition of poles:
  - pull root locus to the right
  - lower system's relative stability
  - slow down the settling of the response

- Compare root locus plots of:

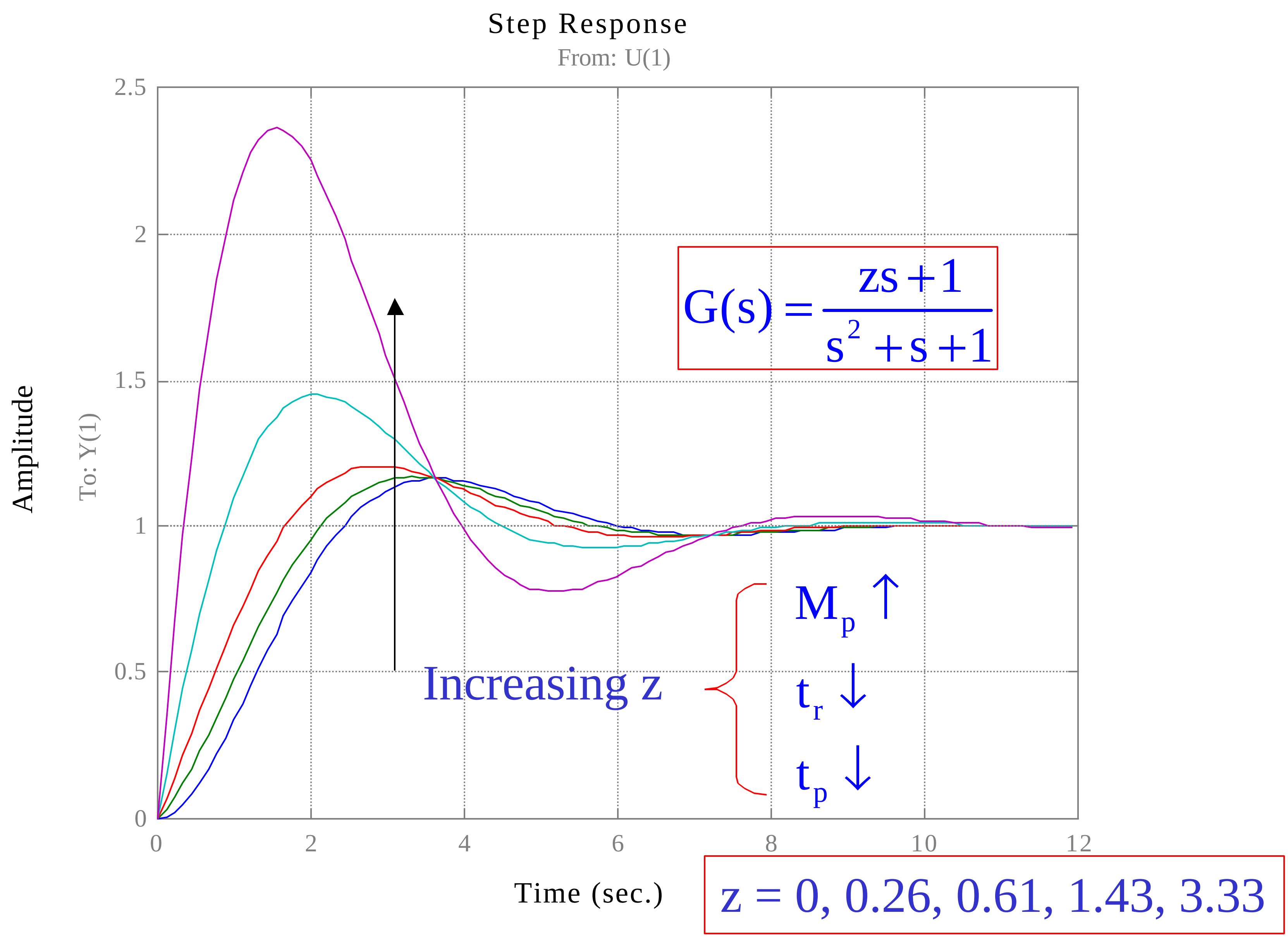
$$\frac{1}{(s+4)} \Rightarrow \frac{1}{(s+4)(s+2)} \Rightarrow \frac{1}{(s+4)(s+2)(s+1)}$$

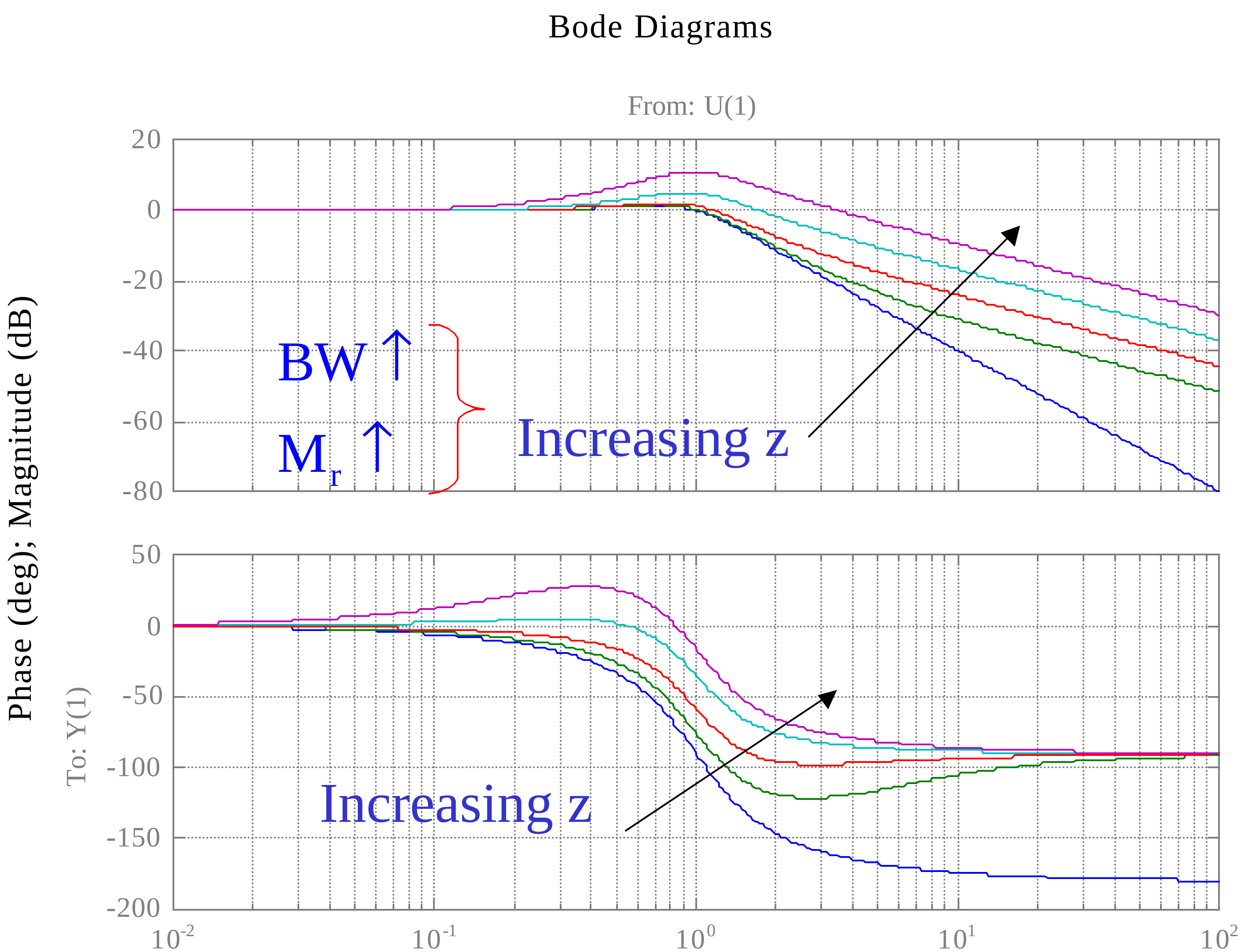
- General effects of the addition of zeros:
  - pull root locus to the left
  - makes system more stable
  - speed up the settling of the response

- Compare root locus plots of:

$$\frac{1}{(s+4)(s+2)(s+1)} \Rightarrow \frac{(s+6)}{(s+4)(s+2)(s+1)}$$

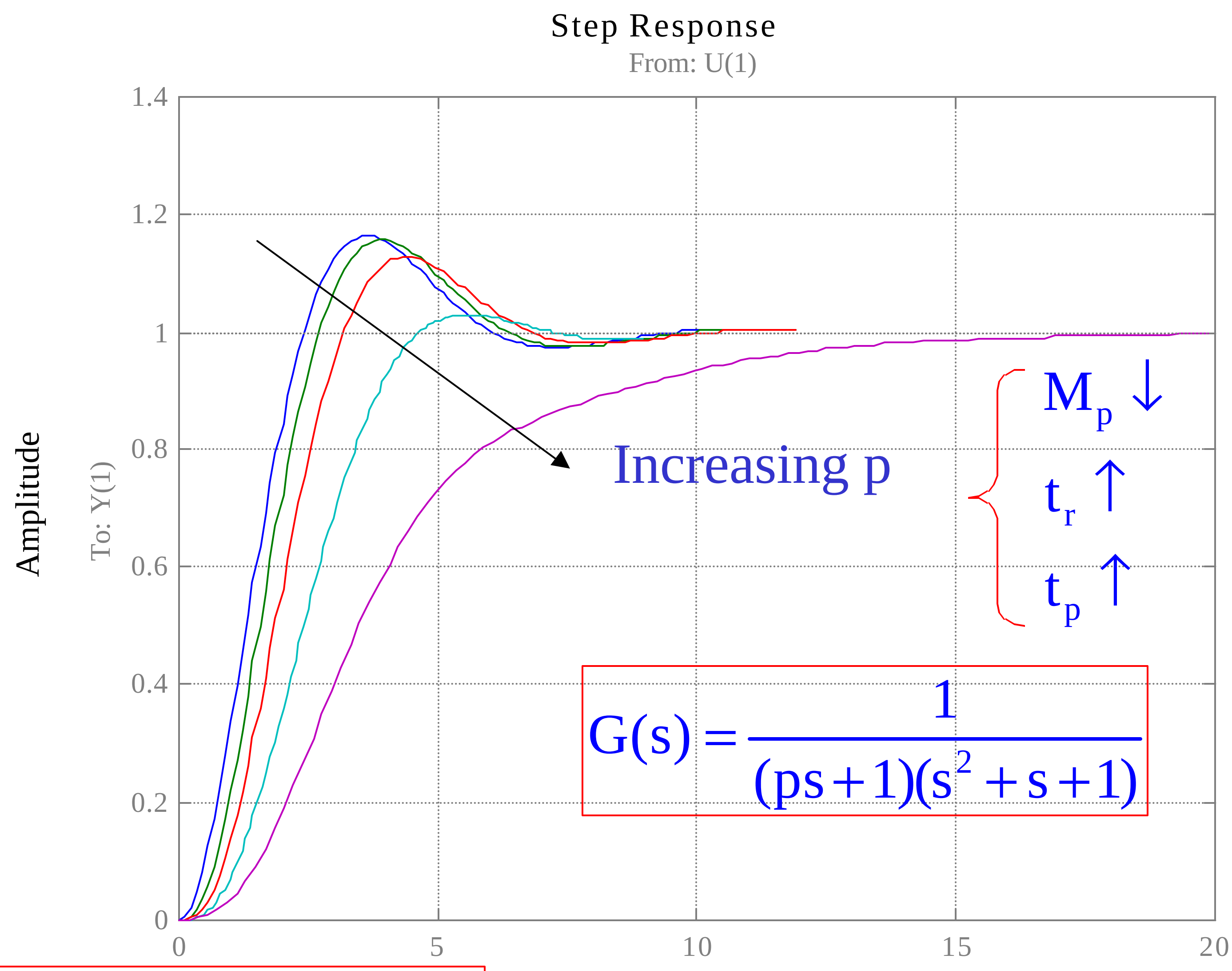
$$\Rightarrow \frac{(s+3)}{(s+4)(s+2)(s+1)} \Rightarrow \frac{(s+1.5)}{(s+4)(s+2)(s+1)}$$


Step Response
From: U(1)
Amplitude
To: Y(1)
Time (sec.)
0
0.5
1
1.5
2
2.5
0
2
4
6
8
10
12
$G(s)=\frac{zs+1}{s^2+s+1}$
Increasing z
$M_p \uparrow$
$t_r \downarrow$
$t_p \downarrow$
$z = 0, 0.26, 0.61, 1.43, 3.33$



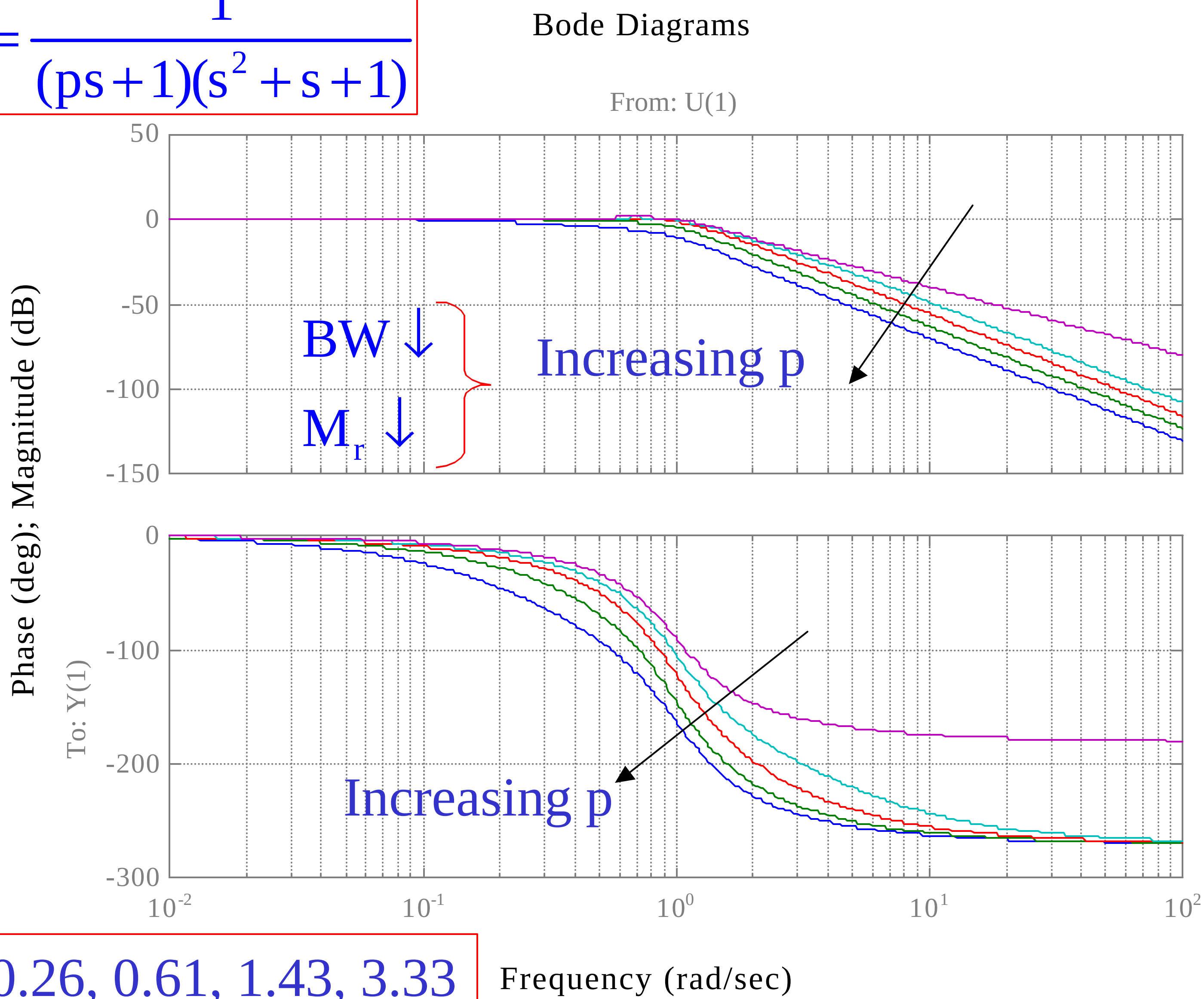
$z = 0, 0.26, 0.61, 1.43, 3.33$

$$G(s) = \frac{zs+1}{s^2+s+1}$$



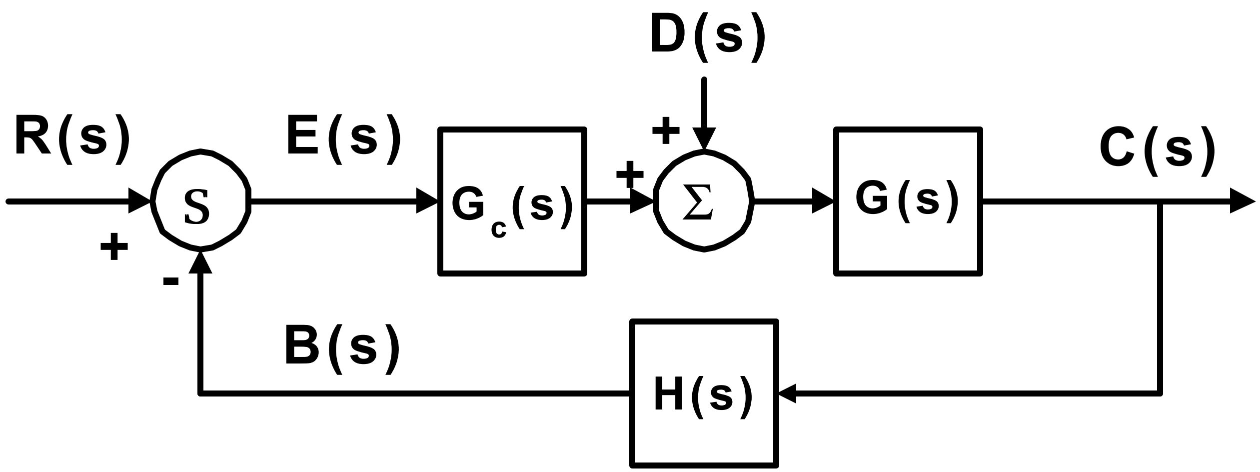
$p = 0, 0.26, 0.61, 1.43, 3.33$

$$G(s) = \frac{1}{(ps+1)(s^2+s+1)}$$



p = 0, 0.26, 0.61, 1.43, 3.33

- Root-Locus Plots: Angle & Magnitude Conditions



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

- The characteristic equation of the closed-loop system is:

$$1 + G_c(s)G(s)H(s) = 0$$

$$G_c(s)G(s)H(s) = -1$$

- Here we assume that $G_c(s)G(s)H(s)$ is a ratio of polynomials in *s*.
- $G_c(s)G(s)H(s)$ is a complex quantity:
  - Angle Condition

$$\angle G_c(s)G(s)H(s) = \pm 180^\circ (2k+1) \qquad (k = 0,\ 1,\ 2,\ \ldots)$$

  - Magnitude Condition

$$\left|G_c(s)G(s)H(s)\right| = 1$$

- The values of s that fulfill both the angle and magnitude conditions are the roots of the characteristic equation, or the closed-loop poles.
- A plot of the points of the complex plane satisfying the angle condition alone is the root locus.
- The roots of the characteristic equation (the closed-loop poles) corresponding to a given value of the gain can be determined from the magnitude condition.

– $G_c(s)G(s)H(s)$ often involves a gain parameter K and the characteristic equation may be written as:

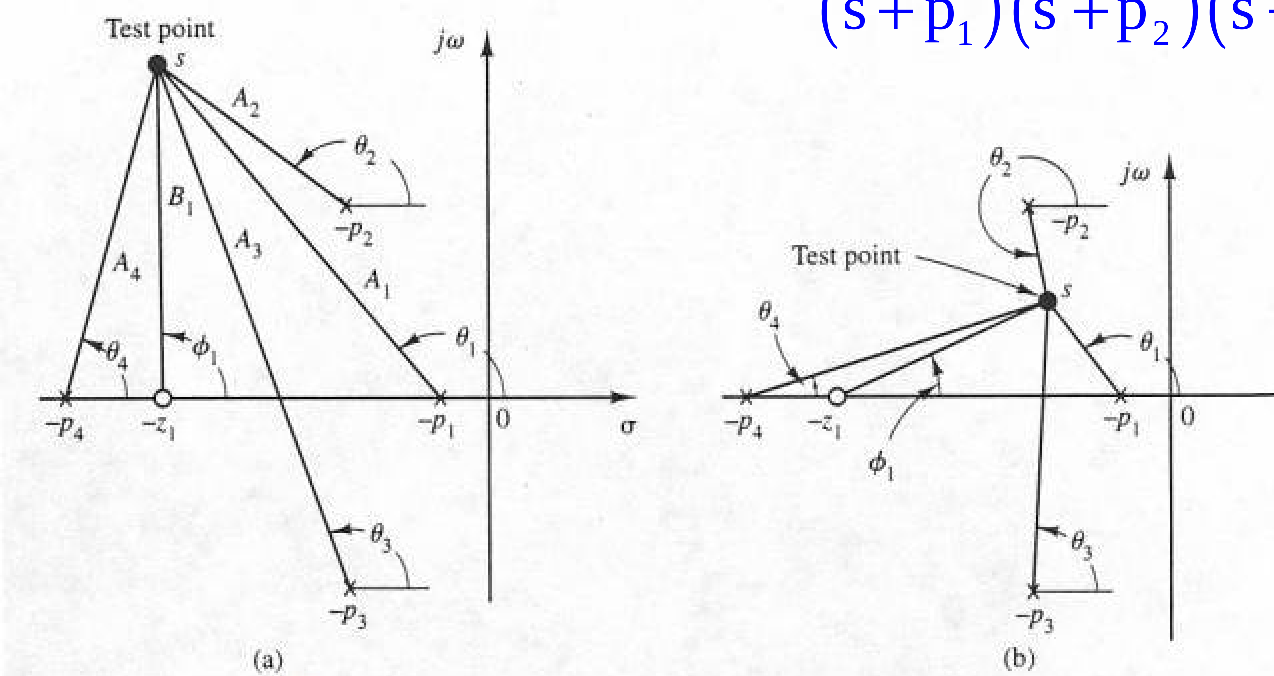
$$1+\frac{K(s+z_1)(s+z_2)\cdots(s+z_m)}{(s+p_1)(s+p_2)\cdots(s+p_n)}=0$$

– The root loci for the system are the loci of the closed-loop poles as the gain K is varied from zero to infinity.

– To begin sketching the root loci we must know the location of the poles and zeros of $G_c(s)G(s)H(s)$.

– Consider:

$$G(s)G_c(s)H(s) = \frac{K(s+z_1)}{(s+p_1)(s+p_2)(s+p_3)(s+p_4)}$$

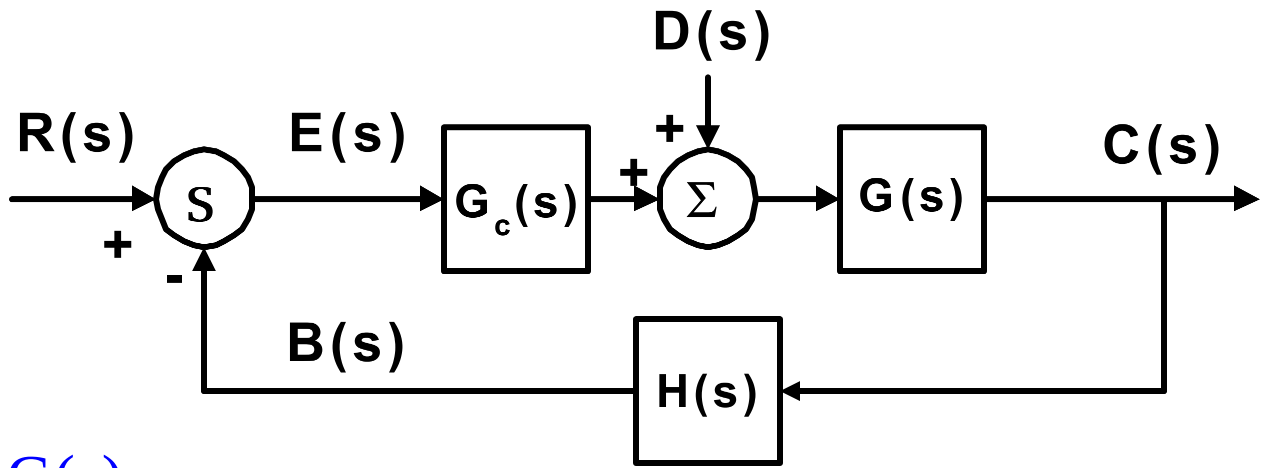


$$\angle G_c(s)G(s)H(s) = \phi_1 - \theta_1 - \theta_2 - \theta_3 - \theta_4$$

$$\left|G_c(s)G(s)H(s)\right| = \frac{K|s+z_1|}{|s+p_1||s+p_2||s+p_3||s+p_4|} = \frac{KB_1}{A_1A_2A_3A_4}$$

– Since the open-loop complex-conjugate poles and complex-conjugate zeros, if any, are always located symmetrically about the real axis, the root loci are always symmetrical with respect to the real axis.

Generalized
Block
Diagram



Assume: D(s) = 0

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

$$G_c(s) = \frac{N_c(s)}{D_c(s)} \qquad G(s) = \frac{N(s)}{D(s)} \qquad G_h(s) = \frac{N_h(s)}{D_h(s)}$$

$$\frac{C(s)}{R(s)} = \frac{N_c N D_h}{D_c D D_h + N_c N N_h}$$

Closed-Loop Zeros consist of the zeros of G_c and G and the poles of H

- Procedure for Applying the Root-Locus Method:
  - Derive the open-loop transfer function $G_c(s)G(s)H(s)$.
  - Factor the numerator and denominator of $G_c(s)G(s)H(s)$.
  - Plot zeros and poles of $G_c(s)G(s)H(s)$ in the *s* plane.
  - By use of computer or geometrical shortcuts, determine the locus that describes the roots of the closed-loop system characteristic equation:

$$1+G_c(s)G(s)H(s)=0$$

– If the gain K of the open-loop system is predetermined, the location of the exact roots of the closed-loop system characteristic equation are immediately known. If the location of the roots is specified, the required value of K can be determined.

– Determine the system's time response by taking the inverse Laplace Transform of C(s) or by computer simulation.

– If the response does not meet the desired specifications, determine the shape that the root locus must have to meet these specifications.

– Synthesize the compensator that must be inserted into the system, if other than gain adjustment is required, to make the required modification on the original locus. This process is called compensation.

– Note: When the open-loop transfer function is in the form shown, with the coefficients of *s* all equal to 1, then K is defined as the loop sensitivity.

$$G_c(s)G(s)H(s) = \frac{K(s+z_1)(s+z_2)\cdots(s+z_m)}{s^u(s+p_1)(s+p_2)\cdots(s+p_v)}$$

$$\text{where } u+v=n$$

- Useful Construction Rules for Negative Feedback
  - The number of branches of the root locus is equal to the number of poles of the open-loop transfer function.
  - For positive values of K, the root locus exists on those portions of the real axis for which the total number of real poles and zeros to the right is an odd number.
  - The root locus starts (K = 0) at the open-loop poles and terminates (K = ± ∞) at the open-loop zeros or at infinity.
  - The angles of the asymptotes of the root locus that end at infinity are determined by:

$$\gamma = \frac{(1+2k)180^\circ}{n-m}$$

$$k = 0,\ 1,\ 2,\ \ldots$$

n = number of finite open-loop poles

m = number of finite open-loop zeros

– The real-axis intercept of the asymptotes is give by:

$$\sigma_a = \frac{(\text{sum of poles}) - (\text{sum of zeros})}{n - m}$$

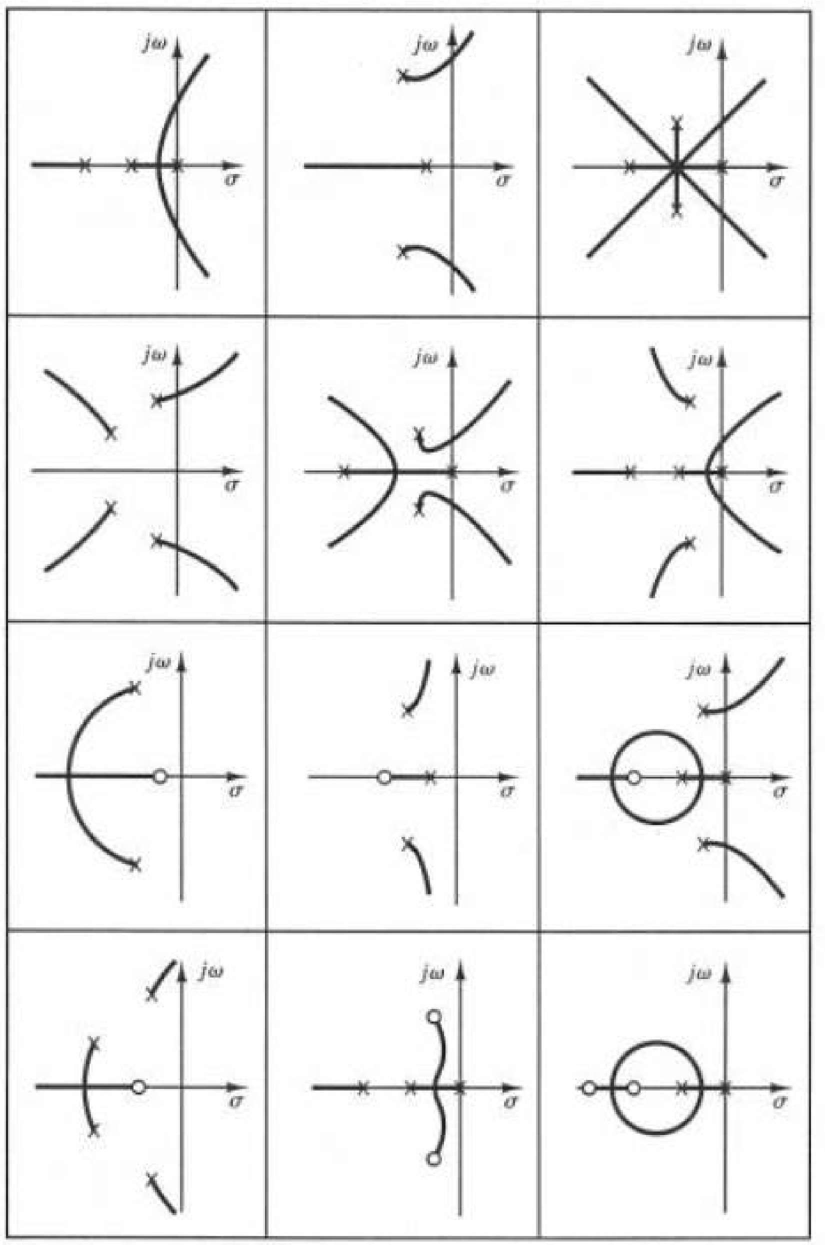
n = number of finite open-loop poles
m = number of finite open-loop zeros

– It is important to note that the asymptotes show the behavior of the root loci for $|s| >> 1$. A root locus branch may lie on one side of the corresponding asymptote or may cross the corresponding asymptote from one side to the other.

- The points where the root loci intersect the $i\omega$ axis can be found by letting $s = i\omega$ in the characteristic equation, equating both the real part and imaginary part to zero, and solving for ω and K. The values of ω thus found give the frequencies at which the root loci cross the imaginary axis. The K value corresponding to each crossing frequency gives the gain at the crossing point.
- The value of K corresponding to any point s on a root locus can be obtained using the magnitude condition or:

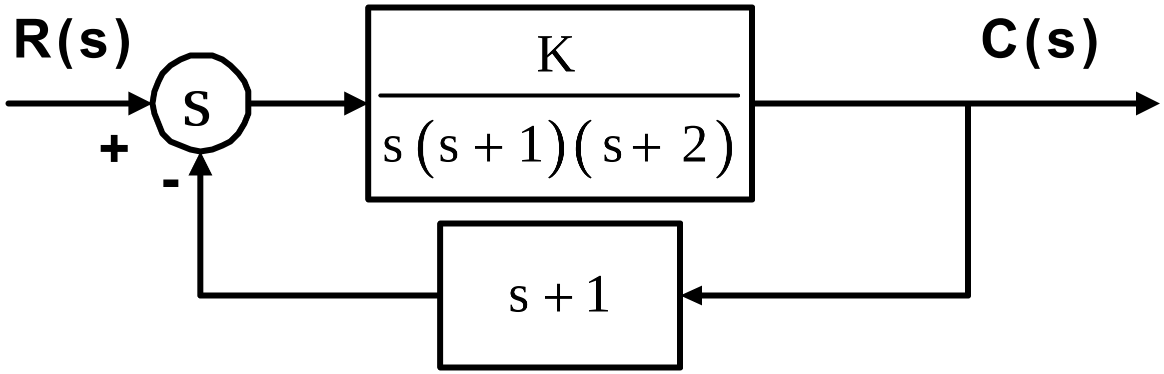
$$K = \frac{\text{product of lengths between point s to poles}}{\text{product of lengths between point s to zeros}}$$

Open-Loop Pole-Zero Configurations and the Corresponding Root Loci



- Comments on Root-Locus Plots
  - A slight change in the pole-zero configuration may cause significant changes in the root-locus configurations.
  - Cancellation of Poles of G(s) with zeros of H(s)
    - Note that if the denominator of G(s) and the numerator of H(s) involve common factors then the corresponding open-loop poles and zeros will cancel each other, reducing the degree of the characteristic equation by one or more.
    - The cancelled pole of G(s)H(s) is a closed-loop pole of the system and this must be added to the closed-loop poles obtained from the root-locus plot of G(s)H(s).

Consider:

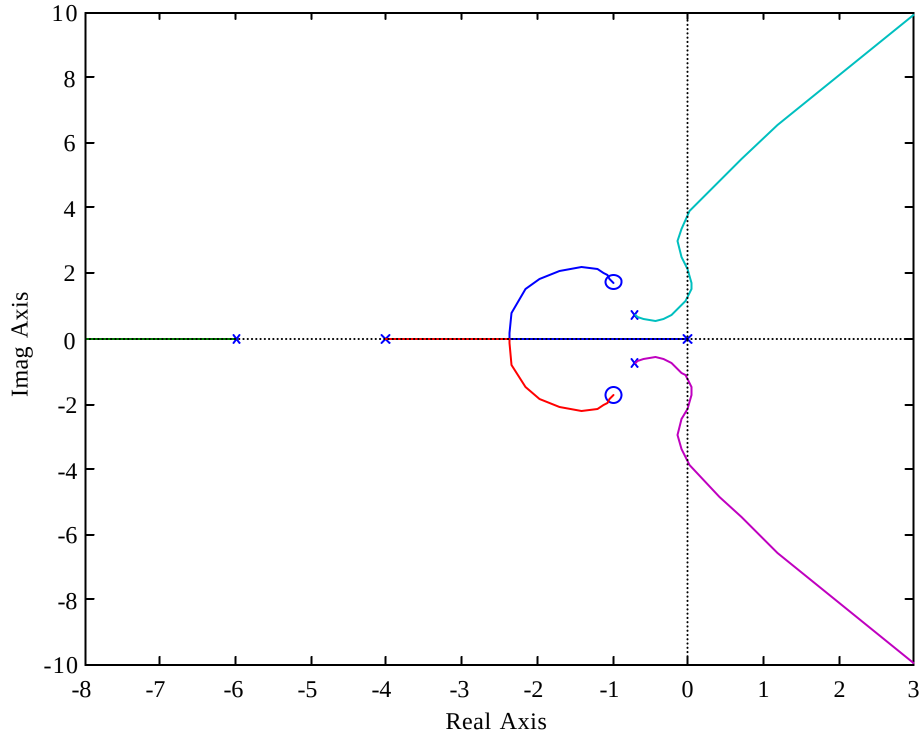


$$\frac{C(s)}{R(s)} = \frac{\dfrac{K}{s(s+1)(s+2)}}{1+\left[\dfrac{K}{s(s+1)(s+2)}\right](s+1)} = \frac{K}{(s+1)\left[s(s+2)+K\right]}$$

$$G(s)H(s) = \left[\frac{K}{s(s+1)(s+2)}\right](s+1) = \frac{K}{s(s+2)}$$

- Conditionally-Stable Systems
  - Consider the following system

Root-Locus Diagram



$$\frac{K\left(s^2+2s+4\right)}{s\left(s+4\right)\left(s+6\right)\left(s^2+1.4s+1\right)}$$

Open-Loop Transfer Function

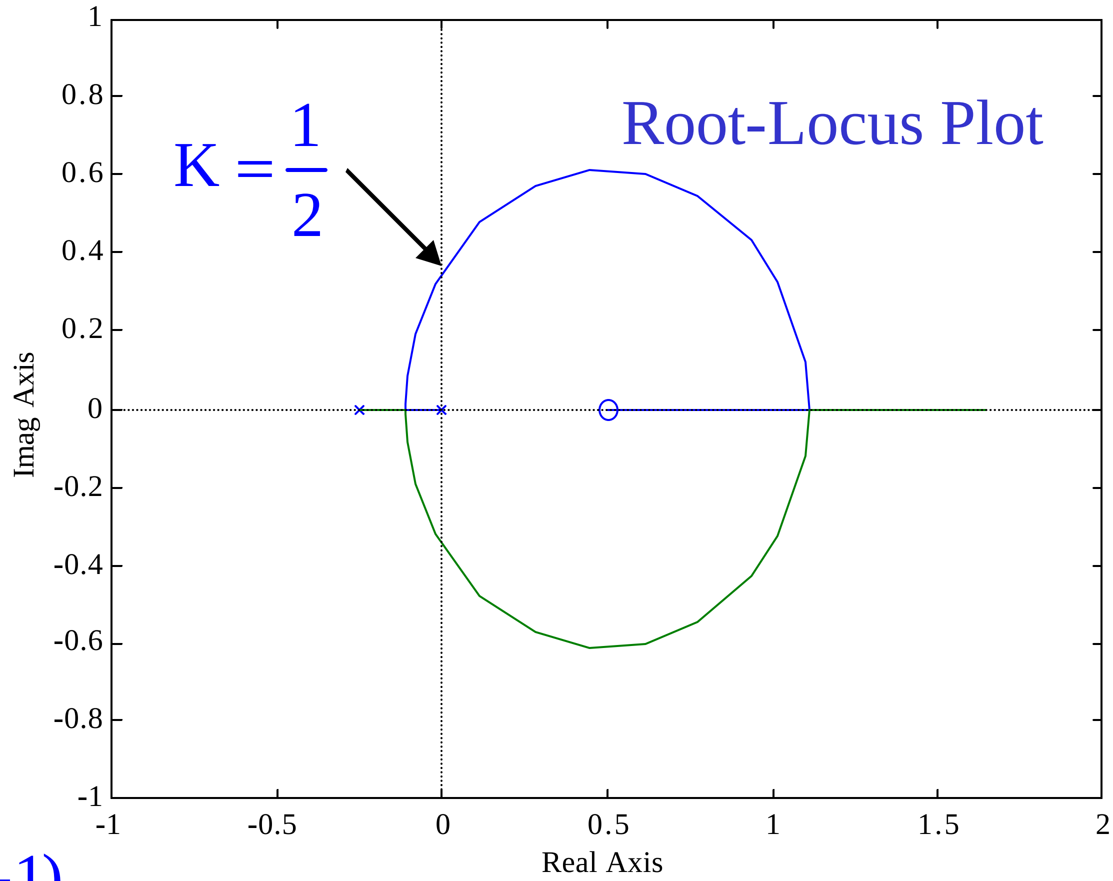
- This system is stable only for limited ranges of K: $0 < K < 12.5$ and $56 < K < 164$.
- The system becomes unstable for $12.5 < K < 56$ and $164 < K$.
- If K assumes a value corresponding to unstable operation, the system may break down or may become nonlinear due to a saturation nonlinearity that may exist.
- Such a system is called *conditionally stable* and it is not desirable since if the gain drops beyond the critical value for some reason, the system becomes unstable.

– Conditional stability may be eliminated by adding proper compensation.

- Nonminimum-Phase Systems
  - If all the poles and zeros of a system lie in the LHP, then the system is called *minimum phase*.
  - If at least one pole or zero lies in the RHP, then the system is called *nonminimum phase*.
  - The term nonminimum phase comes from the phase-shift characteristics of such a system when subjected to sinusoidal inputs.
  - Consider the open-loop transfer function:

$$G(s)H(s) = \frac{K\left(1-2s\right)}{s\left(4s+1\right)}$$

$$G(s)H(s) = \frac{K(1-2s)}{s(4s+1)}$$



Angle Condition:

$$\angle G(s)H(s) = \angle \frac{-K(2s-1)}{s(4s+1)}$$

$$= \angle \frac{K(2s-1)}{s(4s+1)} + 180° = \pm 180°(2k+1) \quad \text{or} \quad \angle \frac{K(2s-1)}{s(4s+1)} = 0°$$

- Systems with Time Delay
  - Any delay in measuring, in controller action, in actuator operation, in computer computation, and the like, is called *transport delay* or *dead time*, and it always reduces the stability of a system and limits the achievable response time of the system.
  - The input x(t) and the output y(t) of a dead time element are related by $y(t) = x(t - \tau_{dt})$ where τ_{dt} is dead time.
  - The Laplace transfer function of a dead time is given by:

$$\frac{Y(s)}{X(s)} = e^{-\tau_{dt} s}$$

- Dead-Time Approximations
  - The simplest dead-time approximation can be obtained by taking the first two terms of the Taylor series expansion of the Laplace transfer function of a dead-time element, τ_{dt}.

$$q_i(t) = \text{ input to dead-time element}$$

$$q_o(t) = \text{ output of dead-time element}$$

$$= q_i\left(t-\tau_{dt}\right)u\left(t-\tau_{dt}\right)$$

$$u\left(t-\tau_{dt}\right)=1 \quad \text{for } t \geq \tau_{dt}$$

$$u\left(t-\tau_{dt}\right)=0 \quad \text{for } t < \tau_{dt}$$

$$L\left[f\left(t-a\right)u\left(t-a\right)\right]=e^{-as}F\left(s\right)$$

$$\frac{Q_o}{Q_i}\left(s\right)=e^{-\tau_{dt}s}\approx 1-\tau_{dt}s$$

$$q_o\left(t\right)\approx q_i\left(t\right)-\tau_{dt}\frac{dq_i}{dt}$$

- The accuracy of this approximation depends on the dead time being sufficiently small relative to the rate of change of the slope of $q_i(t)$. If $q_i(t)$ were a ramp (constant slope), the approximation would be perfect for any value of τ_{dt}. When the slope of $q_i(t)$ varies rapidly, only small τ_{dt}'s will give a good approximation.

- A frequency-response viewpoint gives a more general accuracy criterion; if the amplitude ratio and the phase of the approximation are sufficiently close to the exact frequency response curves of $e^{-\tau_{dt}s}$ for the range of frequencies present in $q_i(t)$, then the approximation is valid.

– The Pade' approximants provide a family of approximations of increasing accuracy (and complexity), the simplest two being:

$$\frac{Q_o}{Q_i}(s) = \frac{2-\tau_{dt}s}{2+\tau_{dt}s} \qquad \frac{Q_o}{Q_i}(s) = \frac{2-\tau_{dt}s+\dfrac{(\tau_{dt}s)^2}{8}}{2+\tau_{dt}s+\dfrac{(\tau_{dt}s)^2}{8}}$$

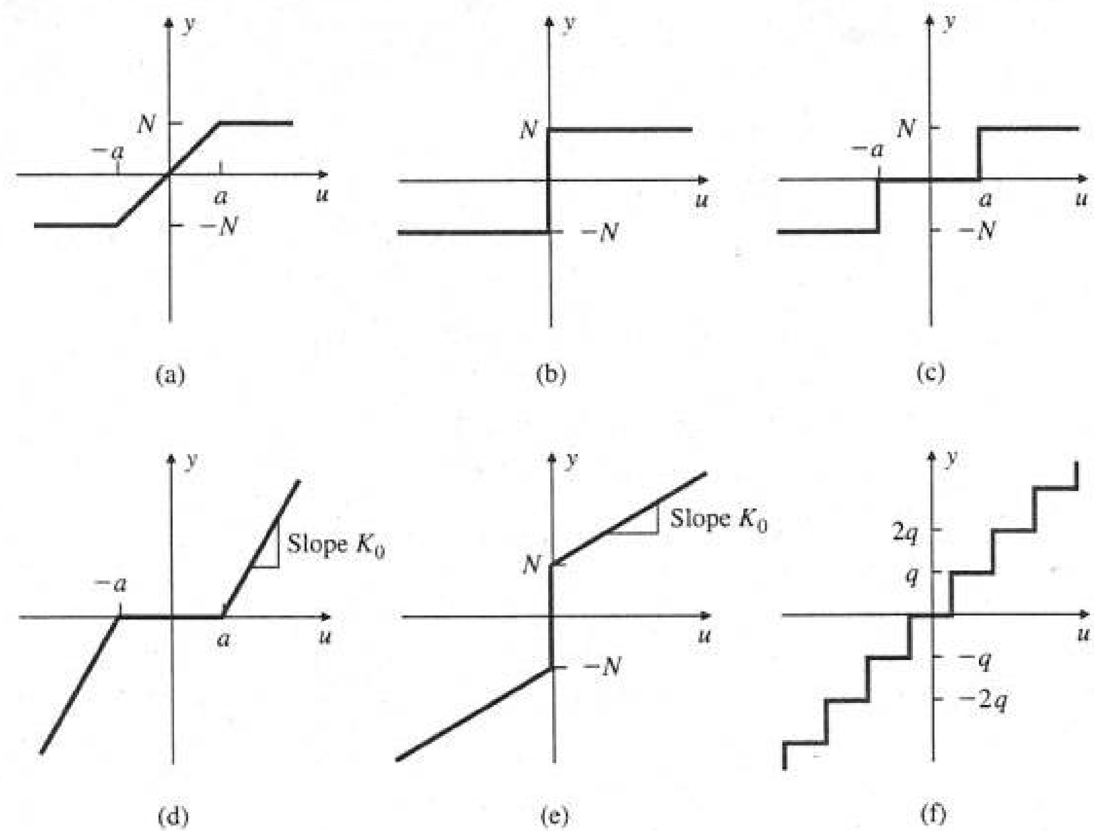
– In some cases, a very crude approximation given by a first-order lag is acceptable:

$$\frac{Q_o}{Q_i}(s) = e^{-\tau_{dt}s} \approx \frac{1}{\tau_{dt}s+1}$$

- Nonlinear Systems
  - Every real control system is nonlinear and we use linear approximations to the real models.
  - There is one important category of nonlinear systems for which some significant analysis can be done: systems in which the nonlinearity has no dynamics and is well approximated as a gain that varies as the size of its input signal varies.
  - The behavior of systems containing such a nonlinearity can be quantitatively described by considering the nonlinear element as a varying, signal-dependent gain.

Nonlinear Elements with No Dynamics

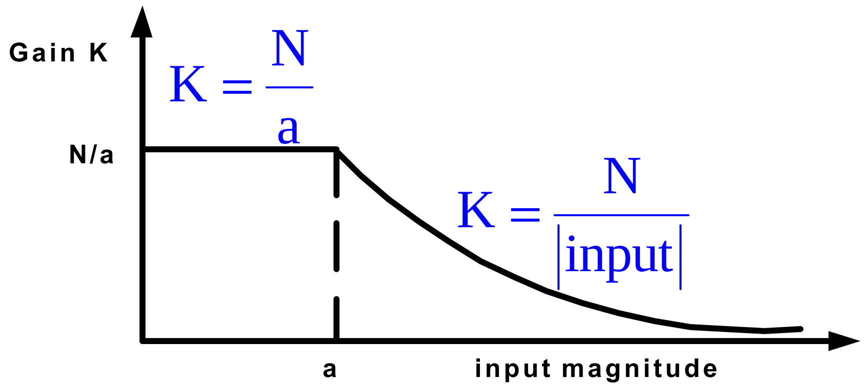
(a) Saturation
(b) Relay
(c) Relay with Dead Zone
(d) Gain with Dead Zone
(e) Pre-loaded Spring or Coulomb plus Viscous Friction
(f) Quantization



– As an example, consider the *saturation element*. All actuators saturate at some level; if they did not, their output would increase to infinity, which is physically impossible.

– For the saturation element, it is clear that for input signals with magnitudes < a, the nonlinearity is linear with the gain N/a. However, for signals > a, the output size is bounded by N, while the input size can get much larger than a, so once the input exceeds a, the ratio of output to input goes down.

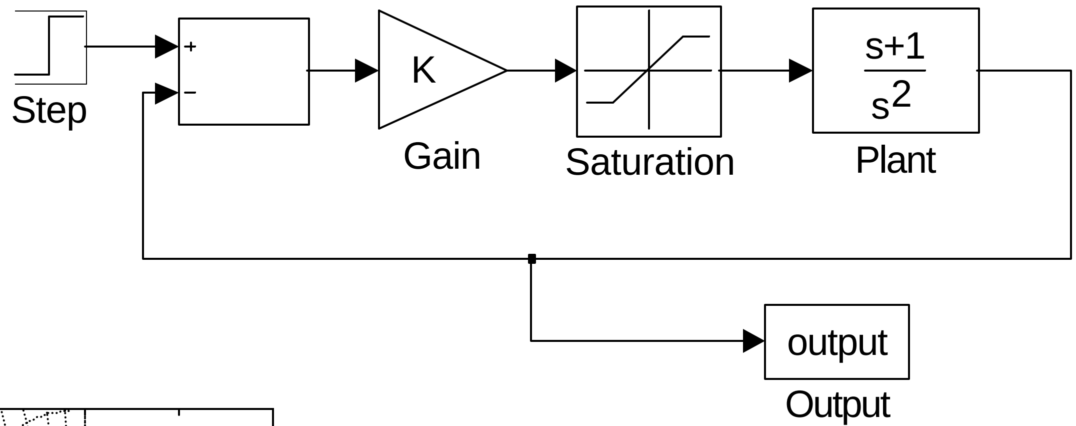
General Shape of the Effective Gain of Saturation

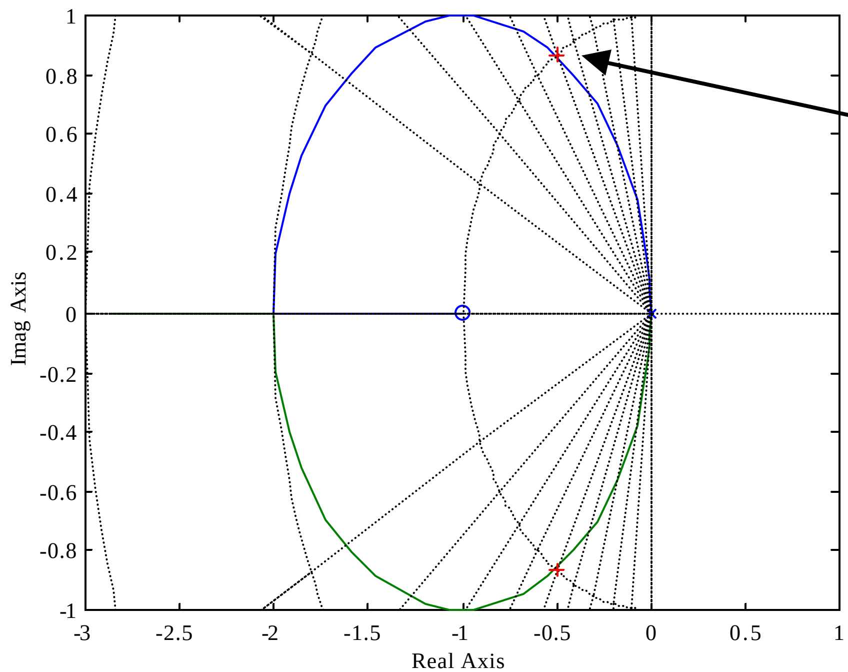


- An important aspect of control system design is *sizing the actuator*, which means picking the size, weight, power required, cost, and saturation level of the device.
- Generally, higher saturation levels require bigger, heavier, and more costly actuators.
- The key factor that enters into the sizing is the effect of the saturation on the control system's performance.

Saturation levels: ± 0.4

Dynamic System With Saturation





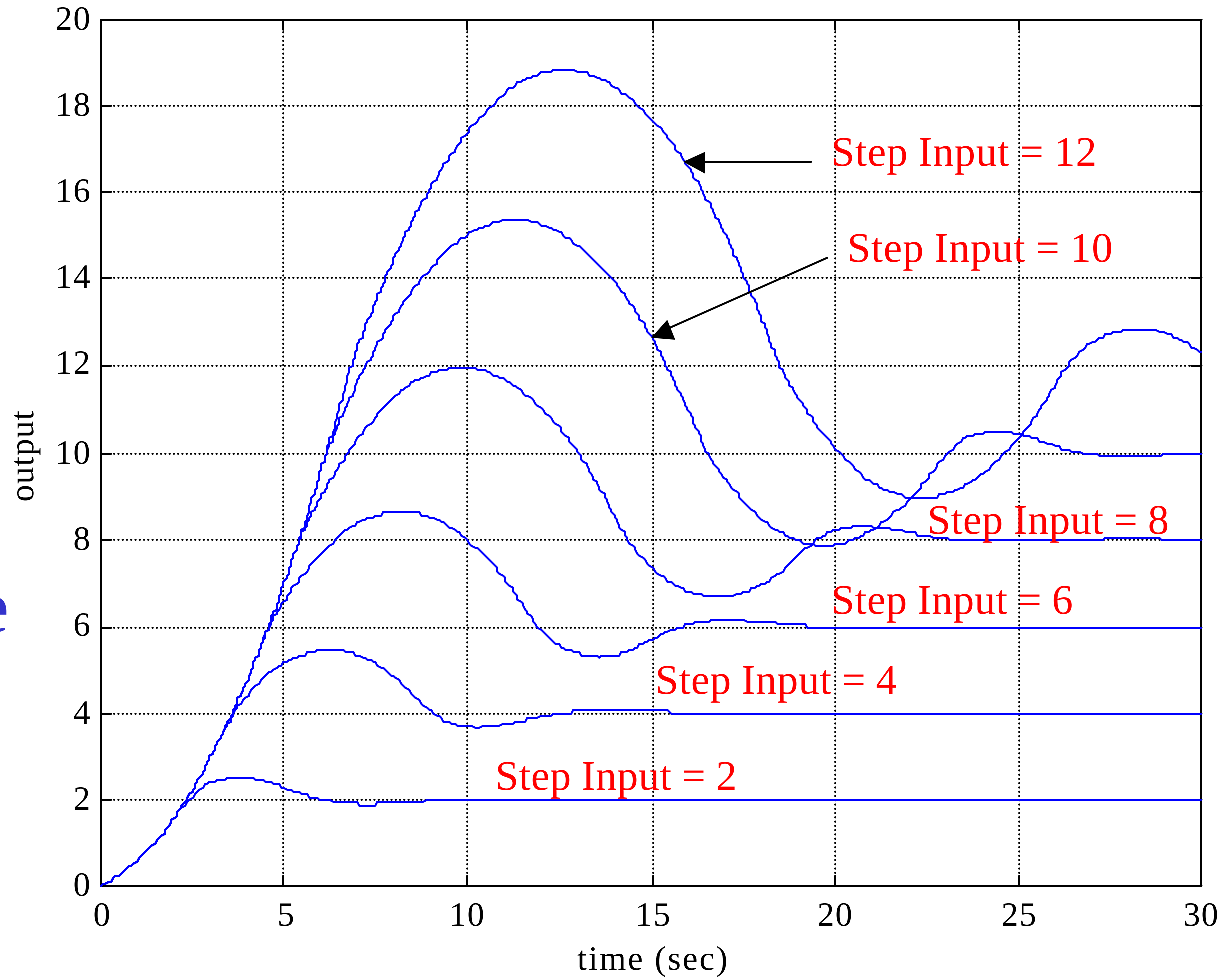
$\zeta = 0.5,\ K = 1$

Root-Locus Plot Without Saturation

As K is reduced, the roots move toward the origin of the s-plane with less and less damping.

Step-Response Results

$\zeta = 0.5, K = 1$

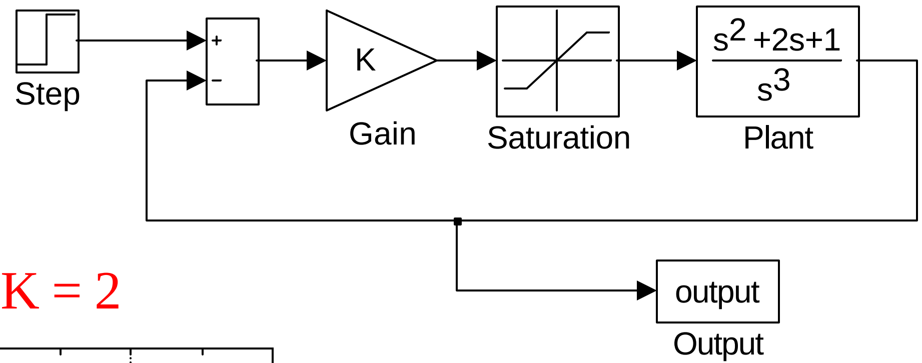


– Observations

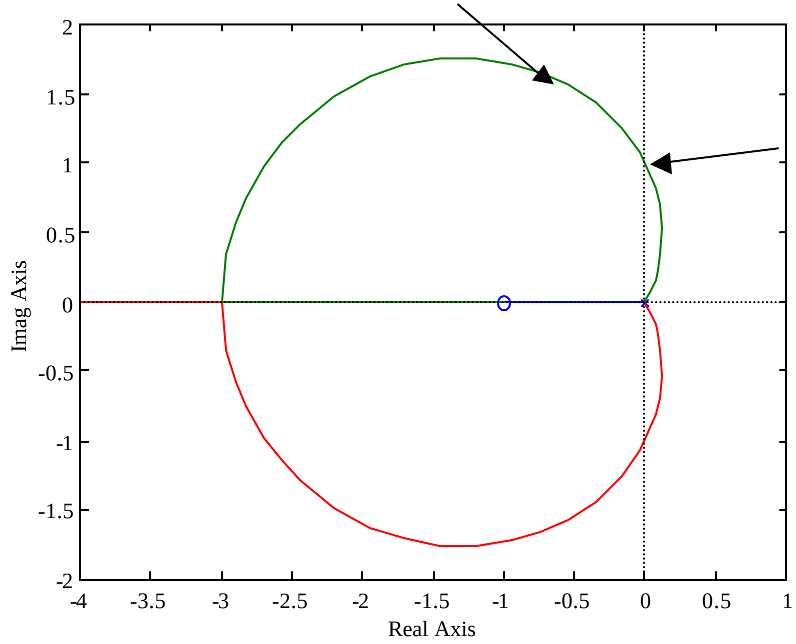
- As long as the signal entering the saturation remains less than 0.4, the system will be linear and should behave according to the roots at $\zeta = 0.5$.
- However, notice that as the input gets larger, the response has more and more overshoot and slower and slower recovery.
- This can be explained by noting that larger and larger input signals correspond to smaller and smaller effective gain K.
- From the root-locus plot, we see that as K decreases, the closed-loop poles move closer to the origin and have a smaller damping ζ.
- This results in the longer rise and settling times, increased overshoot, and greater oscillatory response.

– As another example, consider the block diagram below.

Saturation levels: ± 1



ζ =0.5, K = 2



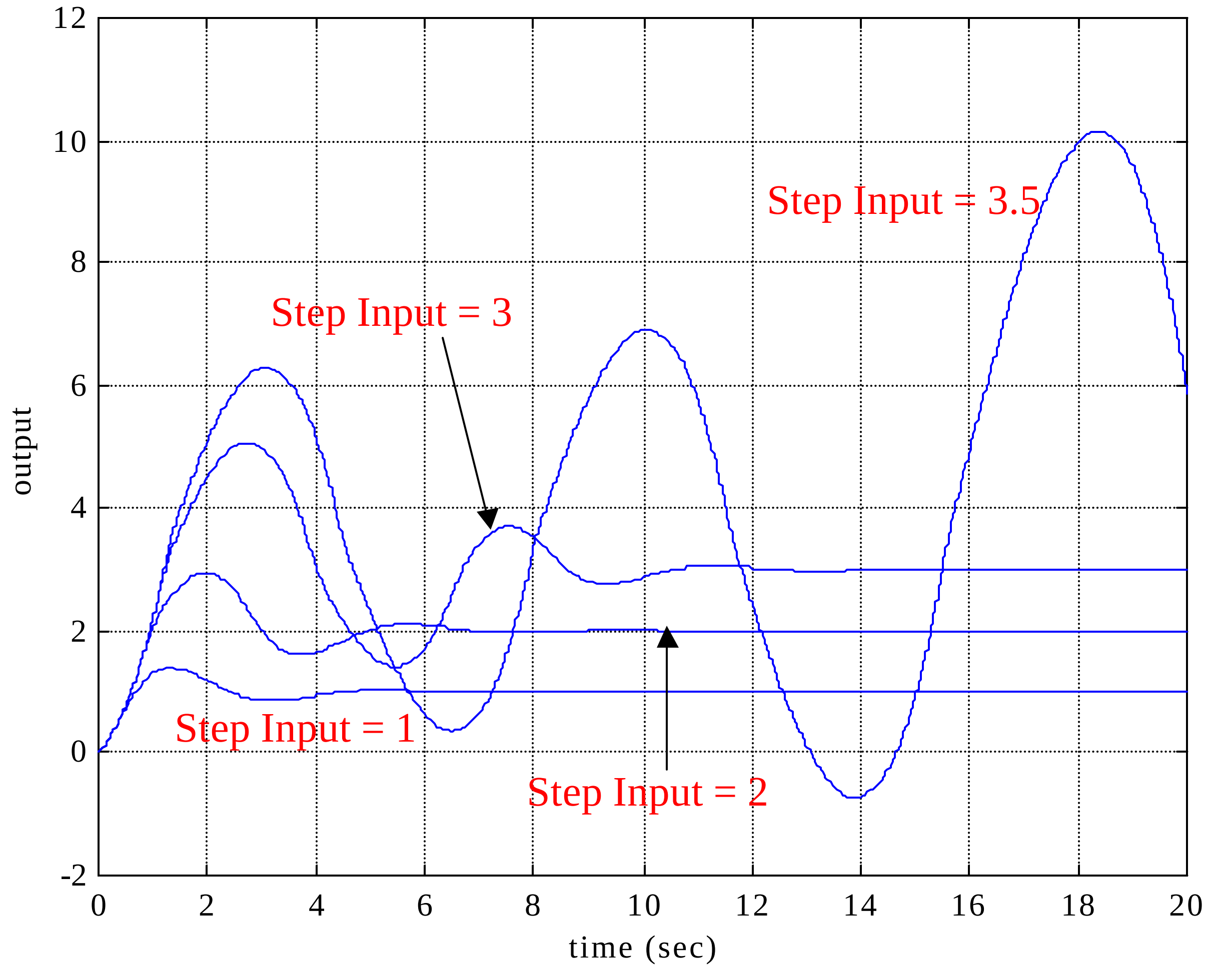
K = 1/2

Root-Locus Plot Without Saturation

System is stable for large gains but unstable for smaller gains

Step-Response Results

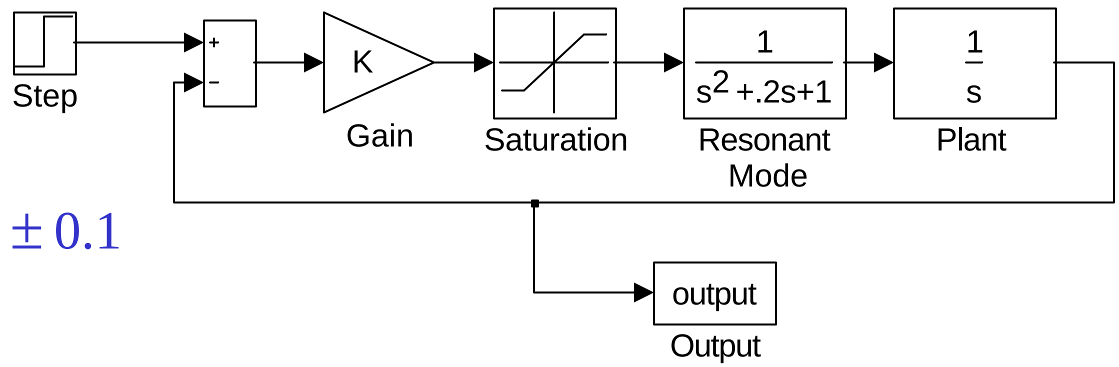
ζ =0.5, K = 2



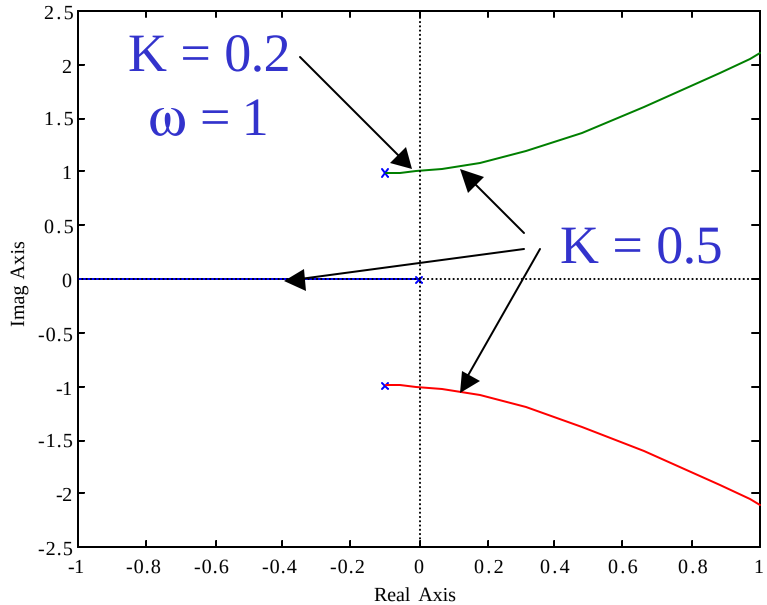
– Observations

- For K = 2, which corresponds to $\zeta = 0.5$ on the root locus, the system shows responses consistent with $\zeta = 0.5$ for small signals.
- As the signal strength is increased, the response becomes less well damped.
- As the signal strength is increased even more, the response becomes unstable.

– As a final example, consider the following block diagram.

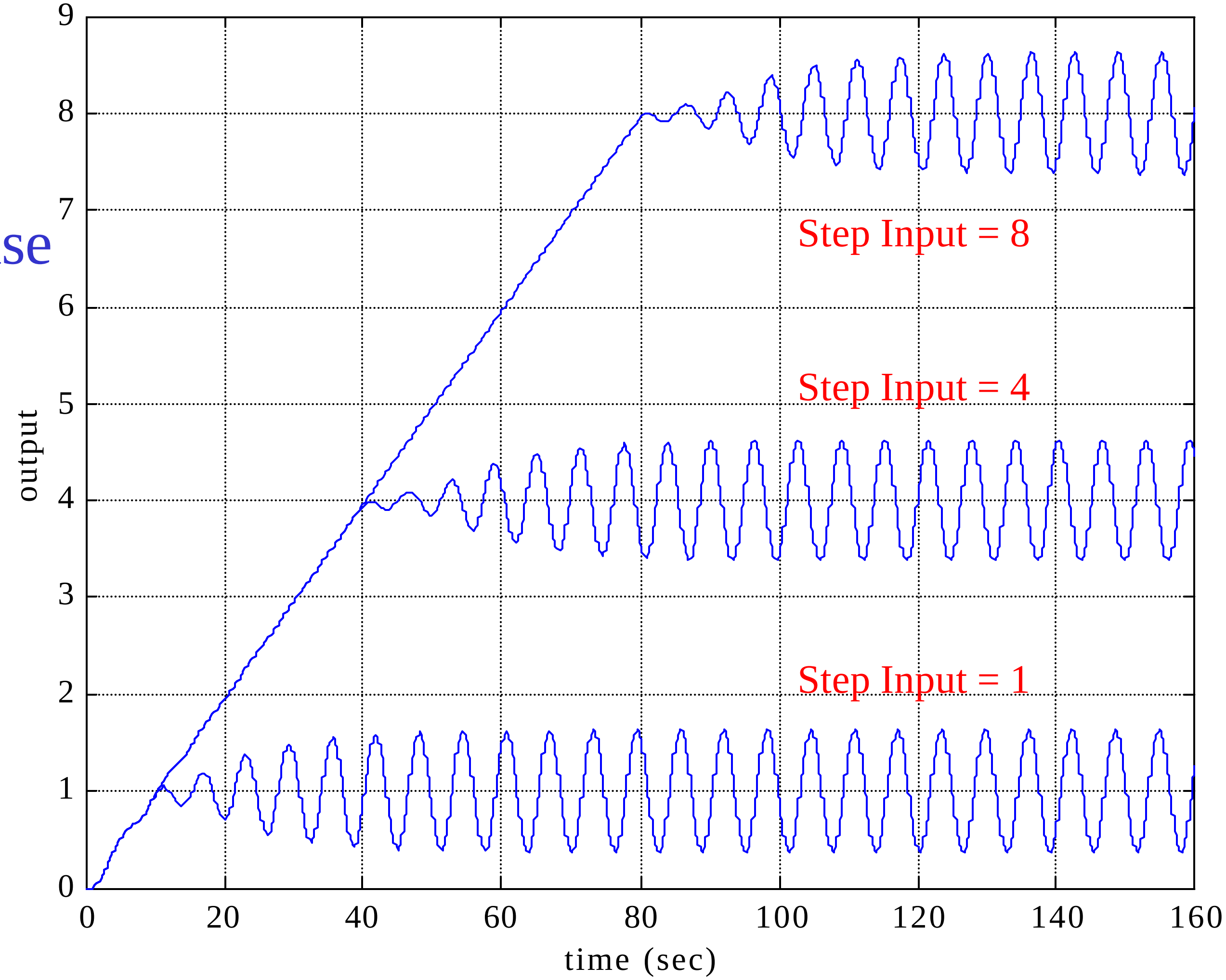


Saturation levels: ± 0.1



Root-Locus Plot Without Saturation

Step-Response Results



K = 0.5

– Observations

- This system is typical of electromechanical control problems where the designer perhaps at first is not aware of the resonant mode corresponding to the denominator term $s^2 + 0.2s + 1$ ($\omega = 1$, $\zeta = 0.1$).
- A gain of $K = 0.5$ is enough to force the roots of the resonant mode into the RHP. At this gain our analysis predicts a system that is initially unstable, but becomes stable as the gain decreases.
- Thus we see that the response of the system with saturation builds up due to the instability until the magnitude is sufficiently large that the effective gain is lowered to $K = 0.2$ and then stops growing!

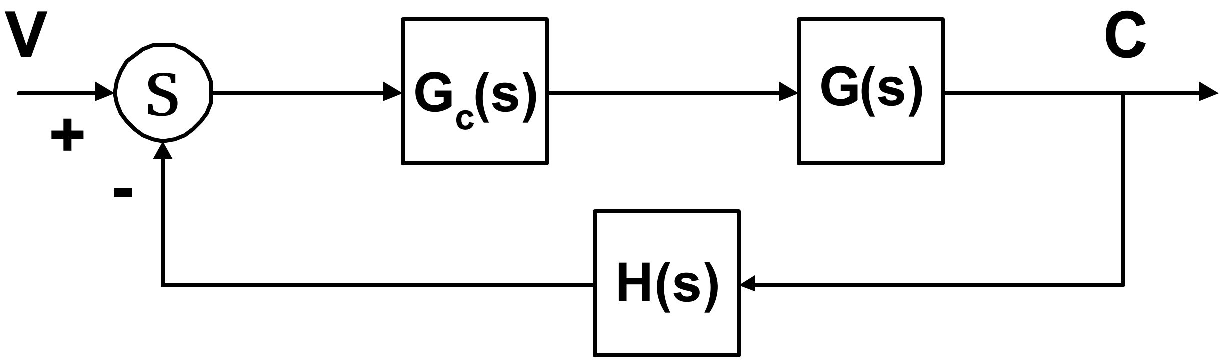
- The error builds up to a fixed amplitude and then starts to oscillate. The oscillations have a frequency of 1 rad/sec and hold constant amplitude at any DC equilibrium value (for the three different step inputs).
- The response always approaches a periodic solution of fixed amplitude known as a *limit cycle*, so-called because the response is cyclic and is approached in the limit as time grows large.
- In order to prevent the limit cycle, the root locus has to be modified by compensation so that no branches cross into the RHP. One common method to do this for a lightly-damped oscillatory mode is to place compensation zeros near the poles, but at a slightly lower frequency.

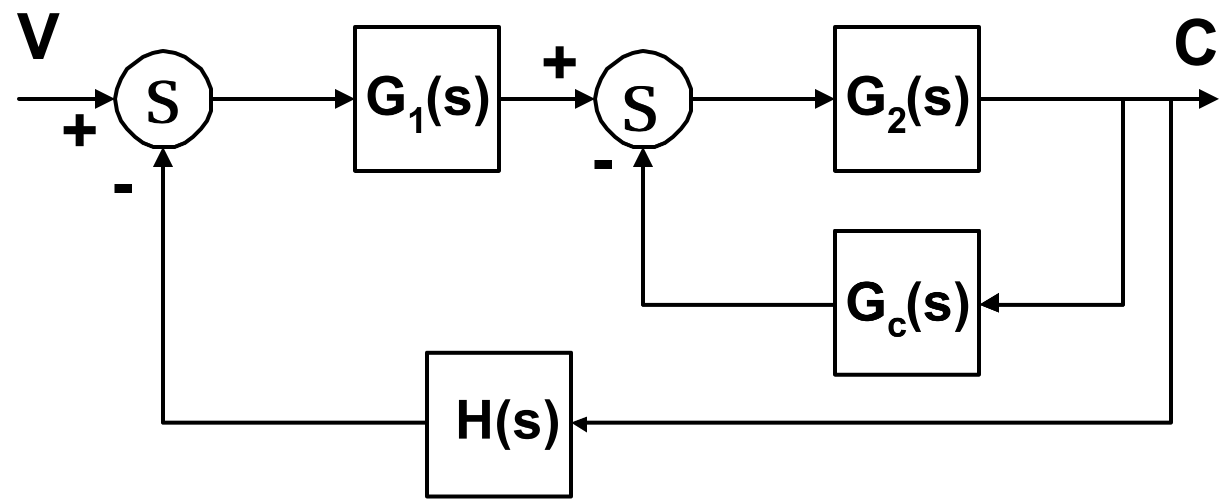
Compensation: General

- SISO Linear Time-Invariant Control Systems
- Performance Specifications:
  - speed of response
  - relative stability
  - steady-state accuracy
- State the performance specifications precisely to yield an optimal control system design for the given purpose.

- System Compensation:
  - Set the gain
    - This is the first step in adjusting the system for satisfactory performance. In many practical cases, however, the adjustment of gain alone may not provide sufficient alteration of the system behavior to meet the given specifications. Increasing the gain value will improve the steady-state behavior but will result in poor stability or even instability.
  - Redesign the system
    - Modify the structure or incorporate additional devices or components to alter the overall behavior so that the system will behave as desired. Such a redesign or addition of a suitable device is called *compensation* and the device inserted into the system for the purpose of satisfying the specifications is called a *compensator.*

Series vs. Feedback (or Parallel) Compensation





- Series vs. Feedback Compensation
  - Whether the compensator is in the feedforward path or the feedback path, the open-loop transfer function and closed-loop poles are identical. Therefore, they both have the same root-locus and Bode plots, so the stability properties are similar.
  - The closed-loop zeros are different, however, so the steady-state errors are different.
  - Because feedback reduces the effects of parameter variations with respect to elements on the forward path, the series configuration has better sensitivity properties and has traditionally been more popular.

- The location of a feedback compensator depends on the complexity of the basic system, the accessibility of the insertion points, the form of the feedback signal, the signal with which it is being compared, and the desired improvement.
- The problem boils down to a suitable design of a series or feedback compensator. The choice between series and feedback compensation depends on the nature of the signals in the system, the power levels at various points, available components, the designer’s experience, economic considerations, etc.

- Series compensation, in general, may be simpler, however it frequently requires additional amplifiers to increase the gain and/or provide isolation. To avoid power dissipation, the series compensator is inserted at the lowest energy point in the feedforward path.
- In general, the number of components required in feedback compensation is less than in series compensation, provided a suitable signal is available, because energy transfer is from a high power level to a low power level.

- Compensators:
  - physical devices: electronic, pneumatic, hydraulic, mechanical
  - lead, lag, lead-lag
- Assure absolute stability of the closed-loop system.
- Model for control system design: Design Model
- Model with nonlinearities, loading effects, other parasitic effects, to evaluate control system design: Truth Model
- By trial and error, the final system must meet performance specifications and be reliable and economical.

- Assume here the plant is given and unalterable.
- Insert compensator to compensate for the undesirable and unalterable characteristics of the plant.
- Continuous-time compensators are considered.
- Series compensators are considered.
- *Root-Locus Method*: graphical method for determining the locations of all closed-loop poles from knowledge of the locations of the open-loop poles and zeros as some parameter is varied.

- If simple gain adjustment doesn't meet the performance specifications, reshape the root loci of the system by the insertion of a suitable compensator so that a pair of dominant closed-loop poles can be placed at the desired location.
- Dominant Complex-Conjugate Poles
  - The other poles must be far to the left of the dominant poles, so that the transients due to these other poles are small in amplitude and die out rapidly.
  - Any other pole which is not far to the left of the dominant complex-conjugate poles must be near a zero so that the magnitude of the transient term due to that pole is small.

- *Reshaping the Root Locus*: the purpose of reshaping the root locus generally falls into one of the following categories.
  - A)
    - System is stable; Transient response is satisfactory; Steady-state error is too large.
    - Here the gain must be increased to reduce the steady-state error without appreciably reducing system stability.
  - B)
    - System is stable; Transient response is unsatisfactory; Steady-state response is satisfactory.
    - Root locus must be reshaped so that it is moved farther to the left, away from the imaginary axis.

- C)
  - System is stable; Transient response is unsatisfactory; Steady-state error is unsatisfactory.
  - Root locus must be moved to the left and gain must be increased.
- D)
  - System is unstable for all values of gain.
  - Root locus must be reshaped so that part of each branch falls in the LHP, thereby making the system stable.

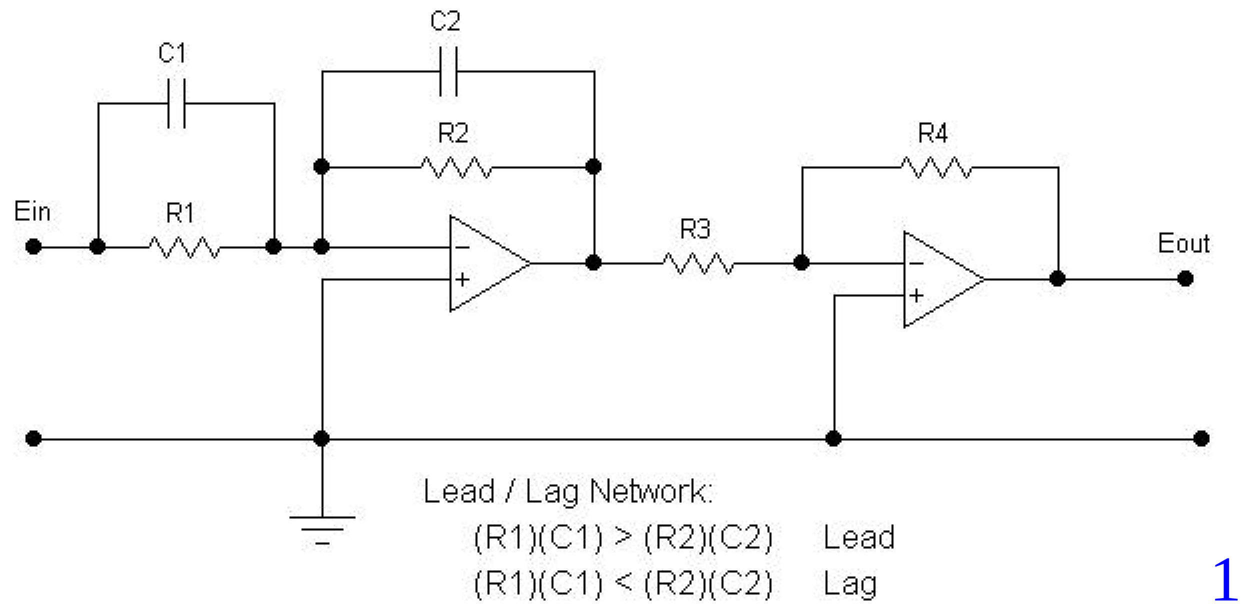
- Compensation of a system by the introduction of poles and zeros is used to improve the operating performance.

Lead Compensation

- Root-locus approach to design is very powerful when the specifications are given in terms of time-domain quantities.
- If the original system is either unstable for all values of gain or is stable but has undesirable transient-response characteristics, reshaping the root locus is necessary in the broad region of the imaginary axis and the origin in order that the dominant closed-loop poles be at desired locations in the complex plane.

- Lead Compensation is an approximate version of PD (proportional + derivative) control. PD control has the following characteristics: $M(s) = (K_P + K_D s)E(s)$
  – Introduction of an additional zero in the forward transfer function reshapes the root locus so that it is moved farther to the left of the imaginary axis.
  – The signal M(s) is now proportional to both the magnitude and rate of change of E(s).
  – The system reacts not only to the magnitude of E(s) but also to its rate of change. If E(s) is changing rapidly, then M(s) is large and the system responds faster. The net result is to speed up the response of the system.

- Derivative action amplifies any spurious signal or noise that may be present and so severely decreases signal-to-noise ratio. The noise amplification may saturate electronic amplifiers so that the system does not operate properly.
- However, PD control is often achieved with a sensor that can directly measure the derivative of the output, e.g., tachometer, which measures velocity. In this case, the derivative term is usually placed in a minor feedback loop around the plant.
- In the feedforward path, the derivative term in the PD controller usually contains a pole to filter out high-frequency noise, which is exactly what a lead compensator is.



$$\frac{E_o(s)}{E_i(s)} = \frac{R_2R_4}{R_1R_3}\frac{R_1C_1s+1}{R_2C_2s+1} = \frac{R_4C_1}{R_3C_2}\frac{s+\dfrac{1}{R_1C_1}}{s+\dfrac{1}{R_2C_2}}$$

Lead or Lag Electronic Compensator

$$= K_c\frac{s+\dfrac{1}{T}}{s+\dfrac{1}{\alpha T}} \quad \left\{ \begin{aligned} T &= R_1C_1 \\ \alpha T &= R_2C_2 \\ K_c &= \frac{R_4C_1}{R_3C_2} \end{aligned} \right.$$

- Digital Implementation of Continuous Controllers
  - The nature of the computer in a computer-controlled system is such that it can work with only one controlled variable at a time. Because of this, having sampled a particular value (obtained its current value), the computer usually must initiate an appropriate action or correction quickly and then move on to the next controlled variable. Thus, although a given variable is continuous in time, the computer has knowledge of its value only at discrete points of time.
  - In general, the control computer performs the following tasks:
    - Obtains a sample value of the process output c_n

- Calculates the error e_n from the relationship $e_n = r_n - c_n$ where r_n is the reference or desired value stored in the computer
- Computes the proper value for the manipulated process input m_n
- Outputs m_n to the appropriate control element
- Continues with the next controlled variable

– A control algorithm must be provided so that the computer can calculate values for the manipulated variable. Because the current error value e_n and stored values of previous errors and control outputs are available, a useful form for control algorithms is:

$$m_n - m_{n-1} = K_0 e_n + K_1 e_{n-1} + K_2 e_{n-2}$$

– This is a linear, constant-coefficient, difference equation where the subscript *n* indicates the present value, subscript *n-1* indicates a value one sample period earlier, subscript *n-2* indicates a value two sample periods earlier, and so on.

– The proportional-integral-derivative (PID) controller is the most widely used controller in use today. It can stabilize a system, increase the speed of response of a system, and reduce steady-state errors of a system. Its continuous transfer function in terms of the differential operator *D* is:

$$\frac{m}{e}(D) = K_P + \frac{K_I}{D} + K_D D = \frac{1}{D}\left(K_D D^2 + K_P D + K_I\right)$$

– The differential equation corresponding to this transfer function is:

$$\frac{dm}{dt} = K_D \frac{d^2 e}{dt^2} + K_P \frac{de}{dt} + K_I e$$

– To convert this differential equation to an equivalent difference equation we approximate the derivatives by finite differences. Of the various schemes possible, the simplest uses the following expressions, where *T* represents the time between samples (the sample period):

$$\frac{de_n}{dt} \approx \frac{e_n - e_{n-1}}{T}$$

$$\frac{d^2 e_n}{dt^2} \approx \frac{\dfrac{de_n}{dt} - \dfrac{de_{n-1}}{dt}}{T} \approx \frac{e_n - 2e_{n-1} + e_{n-2}}{T^2}$$

– Applying this to the differential equation for the PID controller results in the following difference equation:

$$m_n = m_{n-1} + \left(\frac{K_D}{T} + K_P + K_I T\right)e_n - \left(\frac{2K_D}{T} + K_P\right)e_{n-1} + \left(\frac{K_D}{T}\right)e_{n-2}$$

$$m_n - m_{n-1} = \left(\frac{K_D}{T} + K_P + K_I T\right)e_n - \left(\frac{2K_D}{T} + K_P\right)e_{n-1} + \left(\frac{K_D}{T}\right)e_{n-2}$$

– This is of the form:

$$m_n - m_{n-1} = K_0 e_n + K_1 e_{n-1} + K_2 e_{n-2}$$

– This difference equation is programmed into the computer once the values of K_P, K_I, and K_D are determined from control system analysis and design techniques.

- Sample rate T should be faster than 20 times the closed-loop system bandwidth in order to assure that the digital controller will match the performance of the continuous controller.

- Digital Implementation of the Lead or Lag Controller

$$\frac{m(s)}{e(s)} = K\frac{s+a}{s+b}$$

$$\left[s+b\right]m(s) = K\left[s+a\right]e(s)$$

$$\frac{dm}{dt} + bm = K\left[\frac{de}{dt} + ae\right]$$

$$\left[\frac{m_n - m_{n-1}}{T}\right] + bm_n = K\left[\frac{e_n - e_{n-1}}{T} + ae_n\right]$$

$$m_n = \frac{1}{1+bT}m_{n-1} + \frac{K\left(1+aT\right)}{1+bT}e_n - \frac{K}{1+bT}e_{n-1}$$

- Lead Compensation:
  - increases relative stability
  - may increase steady-state error
  - appreciably improves transient response
  - may accentuate high-frequency noise effects
- Lead Compensator:

$$G_c(s) = K_c \alpha \frac{Ts+1}{\alpha Ts+1} = K_c \frac{s+\frac{1}{T}}{s+\frac{1}{\alpha T}} \qquad (0 < \alpha < 1)$$

- Minimum value of α (usually 0.05) is limited by the physical construction of the compensator.

Lead Compensator Design Procedure

- From performance specifications determine the desired location for the dominant closed-loop complex-conjugate poles (e.g., $s = s_{1,2}$).
- Can gain adjustment alone satisfy specifications? If not, calculate the angle deficiency:

$$\phi_c = \pm 180^\circ (2k+1) - \angle G(s_1)$$

Is ϕ_c + or - ? too large?

- Assume lead compensator $G_c(s)$ is:

$$G_c(s) = K_c \alpha \frac{Ts+1}{\alpha Ts+1} = K_c \frac{s+\frac{1}{T}}{s+\frac{1}{\alpha T}} \qquad (0 < \alpha < 1)$$

– α and T are determined from the angle deficiency. K_c is determined from the requirement of the open-loop gain.

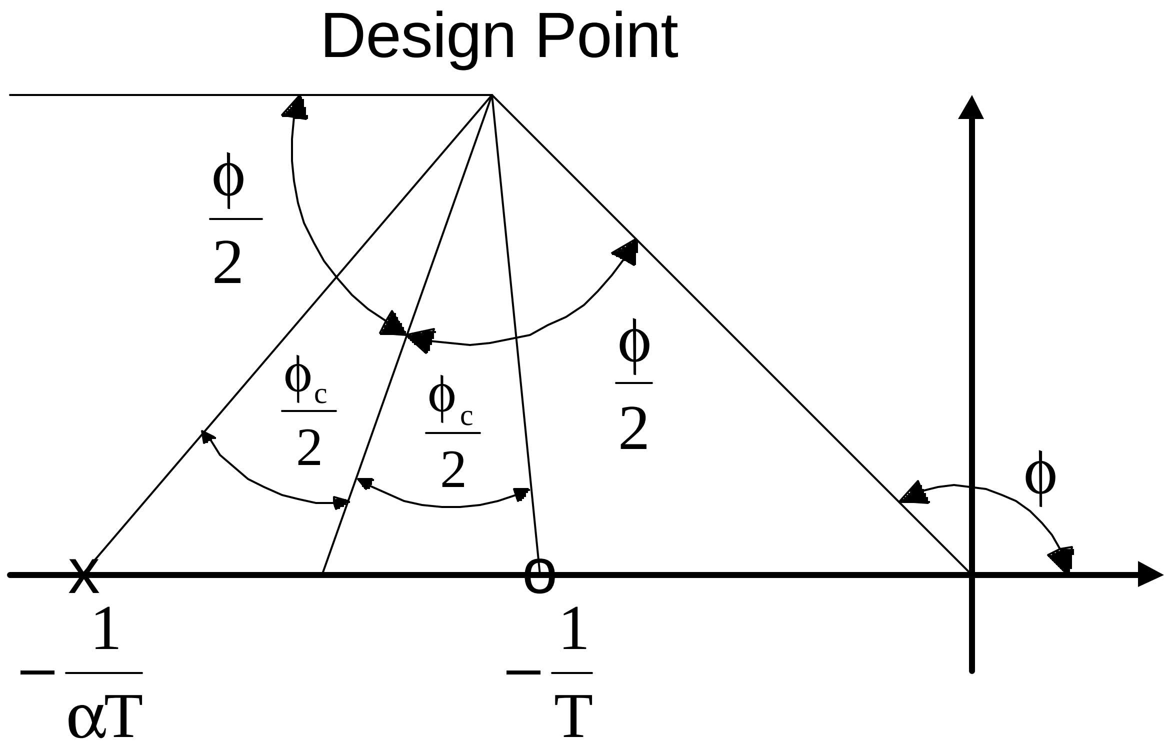
- Locate the pole and zero of the compensator so that the lead compensator will contribute the necessary angle ϕ_c. If no other requirements are imposed on the system, try to make the value for α as large as possible, as it leads to a smaller steady-state error.

$$\phi = \angle(s_1) \qquad \theta_p = \frac{\phi - \phi_c}{2} \qquad \theta_z = \frac{\phi + \phi_c}{2}$$

$$p_c = -\frac{1}{\alpha T} = \mathrm{Re}(s_1) - \frac{\mathrm{Im}(s_1)}{\tan(\theta_p)} \qquad z_c = -\frac{1}{T} = \mathrm{Re}(s_1) - \frac{\mathrm{Im}(s_1)}{\tan(\theta_z)}$$

- Determine the open-loop gain of the compensated system from the magnitude criterion:

$$\left|G_c(s_1)G(s_1)H(s_1)\right| = 1.0$$



Graphical Approach for Locating Lead Compensator Pole and Zero

- Check to see if all specifications have been met. Iterate if necessary. If steady-state error requirements are not met, cascade a lag compensator or alter the lead compensator to a lag-lead compensator.
- Are dominant closed-loop poles really dominant? The closed-loop poles other than the dominant ones modify the response obtained from the dominant closed-loop poles alone; the amount of modification depends on the location of these remaining closed-loop poles. Also, closed-loop zeros affect the response if they are located near the origin.

- Observations
  - Lead compensation is the approximate version of proportional-derivative control: $K_p + K_d s$. The trouble with using PD control, which has only a zero and no pole, is that the physical realization would contain a differentiator that would greatly amplify the inevitable high-frequency noise present from the sensor signal.
  - The effect of the zero and the action of the compensator will not be greatly reduced if we add a high-frequency pole.
  - Selecting the exact values of the pole and zero is done by trial and error. In general the zero is placed so as to satisfy the rise-time or settling time requirements, and the pole is located at a distance 3 to 20 times the value of the zero location.

- The choice of pole location is a compromise between the conflicting effects of noise suppression and compensation effectiveness.
- If the pole is too close to the zero, than the root locus moves back too far towards its uncompensated shape and the zero is not successful in doing its job.
- When the pole is too far to the left, the magnification of noise at the output of the compensator is too great, and the actuator will be overheated by noise energy.

Lag Compensation

- If the original system exhibits satisfactory stability and transient-response characteristics but unsatisfactory steady-state characteristics, the low-frequency gain, which affects the steady-state error, must be increased without appreciably changing the transient-response characteristics, i.e., without appreciably changing the root loci.
- In order to increase the steady-state error constant (reduce the steady-state error), the equivalent of another integration at near-zero frequency is indicated. (Similar to PI Control)

- Lag Compensation is the approximate version of PI (proportional + integral) control. PI control has the following characteristics:

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

  - PI control eliminates steady-state error by increasing the system type without appreciably changing the dominant roots of the characteristic equation.
  - m(t) continues to increase as long as an error e(t) is present. Eventually m(t) becomes large enough to produce c(t) = r(t). Error e(t) then = 0.
  - Constant K_I (generally very small) and the overall gain of the system must be selected to produce satisfactory roots of the characteristic equation.
  - The pole at the origin tends to destabilize the system.

- The improvement is thus made by a pole near $s = 0$, but usually we include a zero nearby so that the pole-zero pair does not significantly interfere with the dynamic response of the overall system as determined by the lead compensation.
- Thus we need an expression for $G_c(s)$ that will yield a significant gain at $s = 0$ to raise the steady-state error constant and that is nearly unity (no effect) at higher frequencies where dynamic response is determined.
- A lag compensator, placed in cascade with the given feedforward transfer function, will accomplish this.

- The angle contribution of the lag compensator must be limited to a small amount (e.g., 5°).
- The pole and zero are placed relatively close together and near the origin of the *s* plane. Then the closed-loop poles of the compensated system will be shifted only slightly from their original locations. Hence, the transient-response characteristics will be changed only slightly.
- There will be a closed-loop root very near the lag-compensation zero. This root will correspond to a very slowly decaying transient, which has a small magnitude because the zero will almost cancel the pole in the transfer function.

- Still, the decay is so slow that this term may seriously influence the settling time. Because of this effect it is important to place the lag pole-zero combination at as high a frequency as possible without causing major shifts in the dominant root locations.
- Also notice that the transfer function from a plant disturbance D(s) to the system error E(s) will not have a zero, and thus disturbance transients can be very long in duration in a system with lag compensation.

- Consider the Lag Compensator:

$$G_c(s) = K_c\beta\frac{Ts+1}{\beta Ts+1} = K_c\frac{s+\frac{1}{T}}{s+\frac{1}{\beta T}} \qquad (\beta > 1)$$

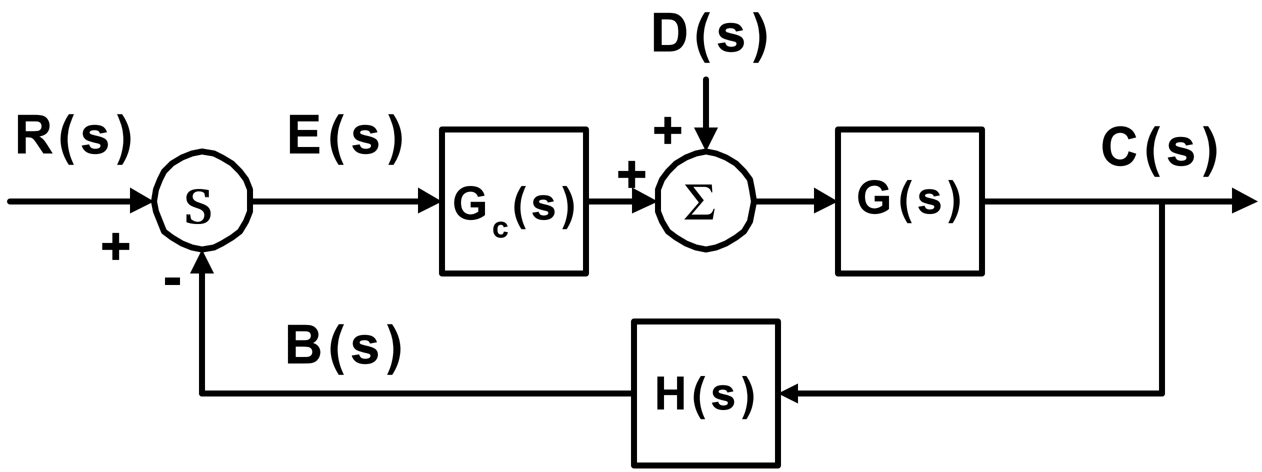
  - If we place the zero and pole of the compensator very close to each other, then at $s = s_1$, where s_1 is one of the dominant closed-loop poles, the magnitudes of the numerator and denominator are almost equal:

$$\left|s_1+\frac{1}{T}\right| \approx \left|s_1+\frac{1}{\beta T}\right|$$

  - This implies that if we set the gain of the compensator $K_c = 1$, then the transient-response characteristics will not be altered.

– The overall gain of the compensator is then β. So the overall gain of the open-loop transfer function can be increased by the factor β, where $\beta > 1$.
– If the pole and zero are placed very close to the origin, then the value of β can be made large provided physical realization of the compensator is possible.
– The value of T must be large, but its exact value is not critical. However, it should not be too large in order to avoid difficulties in realizing the phase-lag compensator by physical components.
– An increase in the gain means a reduction in the steady-state error.

- Review of Steady-State Error



  - Assume H(s) = 1 and D(s) = 0. The error is then E(s) which equals R(s) – C(s).

$$\frac{E(s)}{R(s)} = \frac{1}{1+G_c(s)G(s)} \qquad E(s) = \frac{R(s)}{1+G_c(s)G(s)}$$

$$e_{ss}(t) = \lim_{s \to 0} sE(s) = \lim_{s \to 0} s\frac{R(s)}{1+G_c(s)G(s)}$$

– Step Input: $R(s) = 1/s$

$$e_{ss}(t) = \lim_{s\to 0} s\frac{\frac{1}{s}}{1+G_c(s)G(s)} = \lim_{s\to 0}\frac{1}{1+G_c(s)G(s)} = \frac{1}{1+K_p}$$

$$K_p \equiv \lim_{s\to 0} G_c(s)G(s)$$

– Ramp Input: $R(s) = 1/s^2$

Static Error Constants: K_p and K_v

$$e_{ss}(t) = \lim_{s\to 0} s\frac{\frac{1}{s^2}}{1+G_c(s)G(s)} = \lim_{s\to 0}\frac{1}{s+sG_c(s)G(s)}$$

$$= \lim_{s\to 0}\frac{1}{sG_c(s)G(s)} = \frac{1}{K_v}$$

$$K_v \equiv \lim_{s\to 0} sG_c(s)G(s)$$

- Lag Compensation
  - improves steady-state accuracy
  - increases transient-response time
  - suppresses the effects of high-frequency noise

Lag Compensator Design Procedure

- Draw the root-locus plot for G(s)H(s) and based on the performance specifications determine the desired location for the dominant closed-loop complex-conjugate poles (e.g., $s = s_{1,2}$).
- Assume lag compensator $G_c(s)$ is:

$$G_c(s) = K_c \beta \frac{Ts+1}{\beta Ts+1} = K_c \frac{s+\frac{1}{T}}{s+\frac{1}{\beta T}} \qquad (\beta > 1)$$

- Evaluate the static error constant specified and determine the amount of increase needed to meet specifications.

- Locate the pole and zero of the lag compensator between 0 and 1 that produce the necessary increase in the error constant without appreciably altering the root loci.
- Draw the root locus for the compensated system. Locate the desired dominant closed-loop poles based on the transient-response specifications.
- Determine the open-loop gain of the compensated system from the magnitude criterion and adjust the gain K_c of the compensator so that the dominant closed-loop poles lie at the desired location.

$$\left|G_c(s_1)G(s_1)H(s_1)\right| = 1.0$$

Lead-Lag Compensator

- If improvements in both transient and steady-state response are desired, then both a lead and a lag compensator may be used simultaneously.
- It is economical to use a single lead-lag compensator rather than separate ones.

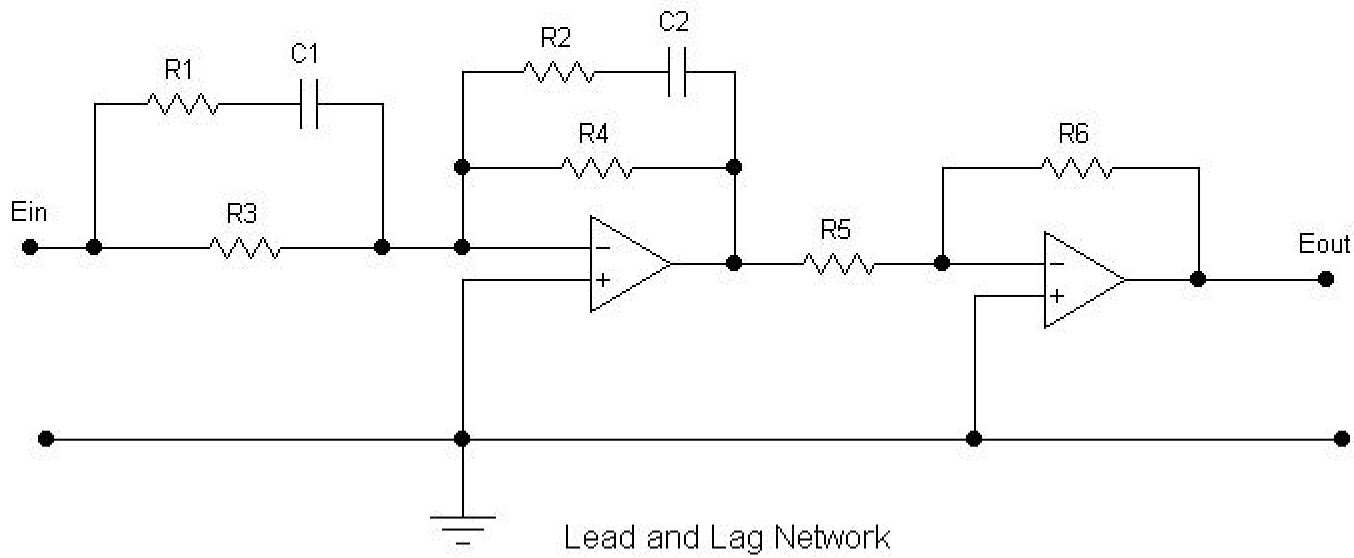
$$G_c(s) = K_c \frac{\beta}{\gamma} \left[\frac{T_1 s + 1}{\frac{T_1}{\gamma} s + 1} \right] \left[\frac{T_2 s + 1}{\beta T_2 s + 1} \right] = K_c \frac{\left(s + \frac{1}{T_1} \right) \left(s + \frac{1}{T_2} \right)}{\left(s + \frac{\gamma}{T_1} \right) \left(s + \frac{1}{\beta T_2} \right)}$$

$$\beta > 1.0$$

$$\gamma > 1.0$$

- Lead/Lag compensation is an approximate version of PID (proportional + integral + derivative) control. The PID controller has the following form:

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



$$\frac{E_o(s)}{E_i(s)} = \frac{R_4 R_6}{R_3 R_5}\left[\frac{(R_1 + R_3)C_1 s + 1}{R_1 C_1 s + 1}\right]\left[\frac{R_2 C_2 s + 1}{(R_2 + R_4)C_2 s + 1}\right]$$

Lead and Lag Electronic Compensator

Lead / Lag Compensator Design Procedure

- Design process is a combination of the design of the lead compensator and that of the lag compensator.
- Determine the desired location for the dominant closed-loop complex-conjugate poles (e.g., $s = s_{1,2}$).
- Using G(s)H(s), determine the angle deficiency ϕ_c if the dominant closed-loop poles are to be at the desired location. The phase-lead portion of the compensator must contribute this angle.

$$\phi_c = \pm 180^\circ (2k+1) - \angle G(s_1)$$

- Choose T_1 and γ from the requirement:

$$\angle \frac{s_1 + \frac{1}{T_1}}{s_1 + \frac{\gamma}{T_1}} = \phi_c$$

- Determine the value of K_c from the magnitude condition:

$$\left| K_c \frac{s_1 + \frac{1}{T_1}}{s_1 + \frac{\gamma}{T_1}} G(s_1)H(s_1) \right| = 1.0$$

- Determine the value of β to satisfy the steady-state error requirement, if specified.

- Using the value of β determined, choose the value of T_2 such that

$$\left| \frac{s+\frac{1}{T_2}}{s+\frac{1}{\beta T_2}} \right| \approx 1 \qquad -5^\circ < \angle \frac{s_1+\frac{1}{T_2}}{s_1+\frac{1}{\beta T_2}} < 0^\circ$$

- Example:

$$G(s) = \frac{4}{s(s+0.5)}$$

Desired performance specifications:

$$\omega_n = 5 \text{ rad/sec} \qquad \zeta = 0.5$$

unit-ramp-input steady-state error < 0.0125

Lead/Lag Compensator:

$$G_c(s) = 6.26\left(\frac{s+0.5}{s+5.02}\right)\left(\frac{s+0.2}{s+0.01247}\right)$$