Laplace Transform

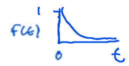
Let f(t) be a function of time t, such that f(t) = 0, for t < 0, then

$$\mathcal{L}[f(t)] = F(s) = \int_0^\infty f(t)e^{-st}dt$$

is the Laplace transform of f(t).

s is a complex variable

e^{-at} $a>0$



Example:

$$f(t) = e^{-at}$$

$$F(s) = \int_0^\infty e^{-at}e^{-st}dt = \int_0^\infty e^{-(a+s)t}dt$$

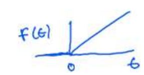
$$= -\frac{e^{-(a+s)t}}{s+a}\Bigg|_0^\infty$$

$$= -\left[\frac{e^{-(a+s)\infty}}{s+a} - \frac{1}{s+a}\right] = \frac{1}{s+a}$$

1

Laplace Transform

$a>0$



Example:

$$f(t) = at$$

$$F(s) = \int_0^\infty ate^{-st}dt$$

$$= -\frac{ate^{-st}}{s}\Bigg|_0^\infty + \int_0^\infty \frac{ae^{-st}}{s}dt$$

$$= -\frac{ae^{-st}}{s^2}\Bigg|_0^\infty = -\left[\frac{ae^{-\infty}}{s^2} - \frac{a}{s^2}\right] = \frac{a}{s^2}$$

Example (Step function):

Unit step: $f(t) = \begin{cases} 1 & \text{for } t>0 \\ 0 & \text{for } t<0 \end{cases}$

Derivative of ramp is step.

$\left(\frac{a}{s^2}\right)(s) = \frac{a}{s}$

ramp derivative

$$f(t) = \begin{cases} 0 & t<0 \\ A & t>0 \end{cases}$$

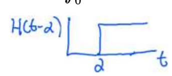
$$\mathcal{L}[f(t)] = F(s) = \int_0^\infty Ae^{-st}dt = \frac{A}{s}$$

2

Laplace Transform (Impulse Function)

Example (Heaviside Step function):

$$\mathcal{H}(t-\alpha) = \begin{cases} 0 & t<\alpha \\ A & t>\alpha \end{cases} \qquad \mathcal{L}[\mathcal{H}(t-\alpha)] = \int_0^\infty \mathcal{H}(t-\alpha)e^{-st}dt$$



Define $\tau = t-\alpha \Rightarrow t = \tau + \alpha$

$dt = d\tau$

$$\mathcal{L}[\mathcal{H}(t-\alpha)] = \int_0^\infty \mathcal{H}(t-\alpha)e^{-st}dt \longrightarrow \mathcal{L}[\mathcal{H}(\tau)] = \int_{-\alpha}^\infty \mathcal{H}(\tau)e^{-s(\tau+\alpha)}d\tau$$

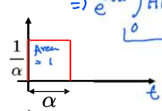
$\int_{-\alpha}^{\infty} H e^{-s(\tau+\alpha)} = \int_{-\infty}^{0} H e^{-s(\tau+\alpha)} d\tau + \int_0^\infty H e^{-s(\tau+\alpha)} d\tau$

$\Rightarrow e^{-s\alpha} \int_0^\infty H(\tau)\, e^{-s\tau}\, d\tau = \frac{e^{-s\alpha}}{s}$

pure delay

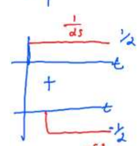
$$\mathcal{L}[\mathcal{H}(t-\alpha)] = \frac{e^{-s\alpha}}{s}$$

Define a pulse function as:



Addition of Heaviside and step function

$$f(t) = \begin{cases} 0 & t<0 \\ \frac{1}{\alpha} & t<\alpha \\ 0 & t>\alpha \end{cases} \qquad \mathcal{L}[f(t)] = \frac{1}{\alpha}\left(\frac{1}{s} - \frac{e^{-s\alpha}}{s}\right)$$



$= \frac{-e^{-s\alpha}}{\alpha s}$

In the limit as α goes to zero, the Laplace transform goes to unity.

Unit pulse: $\lim_{\alpha\to 0}\left(\frac{1-e^{-s\alpha}}{\alpha s}\right) = \infty/\infty \Rightarrow$ l'Hopital's $\lim_{\alpha \to 0}\left(\frac{s e^{-s\alpha}}{s}\right) = 1$

3

Derivative of step is unit impulse. $\left(\frac{1}{s}\right)(s) = 1$ (step)(derivative)

Laplace Transform of the Derivative of a Function

$$\mathcal{L}\left[\frac{df(t)}{dt}\right] = ?$$

Consider $\displaystyle\int_0^\infty f(t)e^{-st}dt = f(t)\frac{e^{-st}}{-s}\Bigg|_0^\infty - \int_0^\infty \left[\frac{df(t)}{dt}\right]\frac{e^{-st}}{-s}dt$

$$\Rightarrow F(s) = \frac{f(0)}{s} + \frac{1}{s}\mathcal{L}\left[\frac{df(t)}{dt}\right]$$

$$\Rightarrow \mathcal{L}\left[\frac{df(t)}{dt}\right] = sF(s) - f(0)$$

$$\frac{df(t)}{dt} = g(t)$$

$$\mathcal{L}\left[\frac{d^2f(t)}{dt^2}\right] = \mathcal{L}\left[\frac{dg(t)}{dt}\right] = s\mathcal{L}[g(t)] - g(0) = s\left(\mathcal{L}\left[\frac{df(t)}{dt}\right]\right) - \dot{f}(0)$$

$$= s^2F(s) - sf(0) - \dot{f}(0)$$

4

Laplace Transform of the Derivative of a Function

Similarly

$$\mathcal{L}\left[\frac{d^n f(t)}{dt^n}\right] = s^n F(s) - s^{n-1} f(0) - s^{n-2} \dot{f}(0) - \ldots - \frac{d^{n-1} f}{dt^{n-1}}(0)$$

$s^3 F - s^2 f(0) - s\dot{f}(0) - \ddot{f}(0)$

5

Laplace Transform of a Function multiplied by e^{-at}

Let f(t) be Laplace transformable, its Laplace transform being F(s), then the Laplace transform of $e^{-at} f(t)$

$= \int_0^\infty e^{-(s+a)t} f(t)\,dt$

$s = s+a$: $\int_0^\infty f(t) e^{-st} dt$

$$\mathcal{L}\left(e^{-at} f(t)\right) = \int_0^\infty e^{-at} f(t) e^{-st} dt = F(s+a)$$

We see that multiplication of f(t) by e^{-at} *has the effect of replacing s by (s+a) in the Laplace Transform*

Example:

$$\mathcal{L}[\ sin(\omega t)] = \frac{\omega}{s^2 + \omega^2}$$

$$\mathcal{L}[\ e^{-at} sin(\omega t)] = \frac{\omega}{(s+a)^2 + \omega^2}$$

If $f(t) = e^{-3t} \sin(5t)$

$\Rightarrow \frac{5}{(s+3)^2 + 25}$

$e^{t} \sin(5t) \Rightarrow \frac{5}{(s-1)^2 + 25}$

6

Complex-Differentiation Theorem

If f(t) is Laplace transformable, then

$$\mathcal{L}[tf(t)] = -\frac{dF(s)}{ds}$$

$$\mathcal{L}\left[t^2 f(t)\right] = \frac{d^2 F(s)}{ds^2}$$

In general

$$\mathcal{L}[t^n f(t)] = (-1)^n \frac{d^n F(s)}{ds^n}$$

Example:

$$\mathcal{L}[\ sin(\omega t)] = \frac{\omega}{s^2 + \omega^2}$$

$(\omega)(s^2 + \omega^2)^{-1}$

$$\mathcal{L}[tsin(\omega t)] = -\frac{d}{ds}\left(\frac{\omega}{s^2 + \omega^2}\right) = \frac{2s\omega}{(s^2 + \omega^2)^2}$$

$\frac{(-1)(\omega)(2s)}{(s^2 + \omega^2)^2} (-1)$

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

7

Final and Initial Value Theorem

<u>*Final Value Theorem:*</u>

If f(t) and $\frac{df(t)}{dt}$ *are Laplace transformable, if F(s) is the Laplace transform of f(t), and if* $\lim_{t\to\infty} f(t)$ *exists, then*

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

Consider

$$\lim_{s\to 0} \int_0^\infty \left[\frac{df(t)}{dt}\right] e^{-st} dt = \lim_{s\to 0} [sF(s) - f(0)]$$

Since

$$\lim_{s\to 0} e^{-st} = 1,$$

8

Final and Initial Value Theorem

we have

$$\int_0^\infty \left[\frac{df(t)}{dt}\right] dt = f(t)|_0^\infty = f(\infty) - f(0) = \lim_{s\to 0} sF(s) - f(0)$$

Thus,

$$f(\infty) = \lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s)$$

Final Value theorem is applicable only if

1. *F(s) has no roots of the denominator (poles) in the complex right half plane*
2. *F(s) should have no poles on the imaginary axis, except at most one pole at s=0*

Initial Value Theorem:

If f(t) and $\frac{df(t)}{dt}$ *are Laplace transformable, if F(s) is the Laplace transform of*

f(t), and if $\lim_{s\to\infty} sF(s)$ *exists, then*

$$f(0^+) = \lim_{s\to\infty} sF(s)$$

9

Inverse Laplace Transform

Example: (Distinct Poles)

$$F(s) = \frac{B(s)}{A(s)} = \frac{k(s+z_1)(s+z_2)\dots(s+z_m)}{(s+p_1)(s+p_2)\dots(s+p_n)} \text{ where } m < n$$

We can write

$$F(s) = \frac{B(s)}{A(s)} = \frac{a_1}{s+p_1} + \frac{a_2}{s+p_2} + \dots + \frac{a_n}{s+p_n}$$

where a_k – residue at the pole $s=p_k$.

Example:

$$F(s) = \frac{(s+3)}{(s+1)(s+2)} = \frac{a_1}{(s+1)} + \frac{a_2}{(s+2)}$$

$$a_1 = \left[\cancel{(s+1)}\frac{(s+3)}{\cancel{(s+1)}(s+2)}\right]\Bigg|_{s=-1} = 2$$

$$a_2 = \left[\cancel{(s+2)}\frac{(s+3)}{(s+1)\cancel{(s+2)}}\right]\Bigg|_{s=-2} = -1$$

$s+3 = a_1(s+2) + a_2(s+1)$

$\left(\frac{s+3}{s+2}\right) = a_1 + \frac{a_2(s+1)}{(s+2)}$

Let $s=-1$

$\frac{2}{1} = a_1 + 0$

$\frac{s+3}{s+1} = \frac{a_1(s+2)}{(s+1)} + a_2$

$s = -2$

$\frac{1}{-1} = a_2 + 0$

10

Inverse Laplace Transform

therefore,

$$f(t) = \mathcal{L}^{-1}[F(s)] = \mathcal{L}^{-1}[\frac{2}{(s+1)}] - \mathcal{L}^{-1}[\frac{-1}{(s+2)}]$$

$$f(t) = 2e^{-t} - e^{-2t} \text{ for } t > 0$$

Example:

$$F(s) = \frac{s^3 + 5s^2 + 9s + 7}{(s+1)(s+2)}$$

This has to be written as

$s^2 + 3s + 2\,)\,s^3 + 5s^2 + 9s + 7\,(\,s+2$

$s^3 + 3s^2 + 2s$

– – – – – –

$2s^2 + 7s + 7$

$2s^2 + 6s + 4$

– – – – – –

$s + 3$

11

Inverse Laplace Transform

therefore,

$$F(s) = (s+2) + \frac{(s+3)}{(s+1)(s+2)}$$

$$f(t) = \frac{d}{dt}(\delta(t)) + 2\delta(t) + 2e^{-t} - e^{-2t} \text{ for } t > 0$$

Example: (Multiple Poles) - Repeated Roots

$$F(s) = \frac{s^2 + 2s + 3}{(s+1)^3}$$

$$F(s) = \frac{B(s)}{A(s)} = \frac{a_1}{(s+1)} + \frac{a_2}{(s+1)^2} + \frac{a_3}{(s+1)^3}$$

$$a_3 = \left[(s+1)^3 F(s)\right]\Big|_{s=-1}$$

$= s^2 + 2s + 3 = a_1(s+1)^2 + \frac{a_2}{(s+1)} + a_3$

$$(s+1)^3\frac{B(s)}{A(s)} = a_1(s+1)^2 + a_2(s+1) + a_3$$

$$\therefore a_3 = \left[(s+1)^3\frac{B(s)}{A(s)}\right]\Bigg|_{s=-1}$$

$\Rightarrow a_3 = 2$

12

Inverse Laplace Transform

$a_1(s^2+2s+1) + a_2(s+1)$

$a_1 s^2 + (2a_1+a_2)s + (a_1+a_2)$

$a_1 = 1$

$2(1) + a_2 = 2 \quad a_2 = 0$

Check

$3 = a_1 + a_2 + 2$

$3 = (1) + 0 + 2$ ✓

$s^2+2s+3 = a_1(s+1)^2 + a_2(s+1) + 2$

$\frac{d}{ds} \Rightarrow \quad 2s+2 = 2a_1(s+1) + a_2$

$a_2 = 0$

$\frac{d}{ds} \Rightarrow \quad 2 = 2a_1$

$a_1 = 1$

Consider,

$$\frac{d}{ds}\left[(s+1)^3\frac{B(s)}{A(s)}\right] = 2a_1(s+1) + a_2$$

$$\therefore a_2 = \frac{d}{ds}\left[(s+1)^3\frac{B(s)}{A(s)}\right]\bigg|_{s=-1}$$

$$\therefore a_2 = \frac{d}{ds}\left[(s^2+2s+3)\right]\Big|_{s=-1} = 0$$

Similiarly

$$2a_1 = \frac{d^2}{ds^2}\left[(s+1)^3\frac{B(s)}{A(s)}\right]\bigg|_{s=-1}$$

$$2a_1 = \frac{d^2}{ds^2}\left[(s^2+2s+3)\right]\Big|_{s=-1} \Rightarrow a_1 = 1$$

$F(s) = \frac{1}{s+1} + \frac{2}{(s+1)^3}$

OR Compare Coefficients $\mathcal{L}^{-1}(F) = e^{-t} + t^2 e^{-t}$

Solving Differential Equations

Consider the differential equation

$$\ddot{x} + 3\dot{x} + 2x = 0, \quad x(0) = a, \dot{x}(0) = b$$

Define $$X(s) = \mathcal{L}[x(t)]$$

$$\mathcal{L}[\dot{x}(t)] = sX(s) - x(0)$$

$$\mathcal{L}[\ddot{x}(t)] = s^2X(s) - sx(0) - \dot{x}(0)$$

Therefore

$$\left[s^2X(s) - sx(0) - \dot{x}(0)\right] + 3\left[sX(s) - x(0)\right] + 2X(s) = 0$$

$$(s^2+3s+2)X(s) = sx(0) + \dot{x}(0) + 3x(0)$$

$a_1 = 2a+b$

$a_2 = -(a+b)$

$$X(s) = \frac{(s+3)a+b}{s^2+3s+2} = \frac{2a+b}{s+1} - \frac{a+b}{s+2} = \frac{a_1}{s+1} + \frac{a_2}{s+2}$$

$$x(t) = (2a+b)e^{-t} - (a+b)e^{-2t} \quad \forall t > 0$$

Solving Differential Equations

Consider the differential equation

$$\ddot{x} + 2\dot{x} + 5x = 3, \quad x(0) = 0, \dot{x}(0) = 0$$

Laplace transformation leads to:

$$(s^2+2s+5)X(s) = \frac{3}{s}$$

$X(s) = \frac{a_1}{s} + \frac{a_2 s + a_3}{s^2+2s+5}$

$\frac{3}{s^2+2s+5} = a_1 + \frac{a_2 s^2 + a_3 s}{s^2+2s+5}\Big|_{s=0}$

$a_1 = \frac{3}{5}$

Therefore

$$X(s) = \frac{3}{s(s^2+2s+5)} = \frac{3}{5s} - \frac{3}{10((s+1)^2+2^2)} - \frac{3(s+1)}{5((s+1)^2+2^2)}$$

$$x(t) = \frac{3}{5} - \frac{3}{10}e^{-t}sin(2t) - \frac{3}{5}e^{-t}cos(2t) \quad \forall t > 0$$

$3 = a_1(s^2+2s+5) + (a_2 s + a_3)s$

$3 = \left(\frac{3}{5} + a_2\right)s^2 + \left(\frac{6}{5} + a_3\right)s + 3$

$X = \frac{3}{5s} - \frac{\frac{3}{5}s + \frac{6}{5}}{(s+1)^2+4} = \frac{3}{5s} - \frac{\frac{3}{5}(s+1) + \frac{3}{5}}{(s+1)^2+4}$

Transfer Function

The transfer function of a linear time-invariant system is defined to be the ratio of the Laplace transform of the output to the Laplace transform of the input under the assumption of zero initial conditions.

Consider the differential equation:

$$a_0 y^n + a_1 y^{n-1} + \ldots + a_n y = b_0 u^m + b_1 u^{m-1} + \ldots + b_m u,$$

Where y is the output of the system and u is the input. The transfer function is given by taking the Laplace transform of the above equation:

$$\left(a_0 s^n + a_1 s^{n-1} + \ldots + a_n\right)Y(s) = \left(b_0 s^m + b_1 s^{m-1} + \ldots + b_m\right)U(s),$$

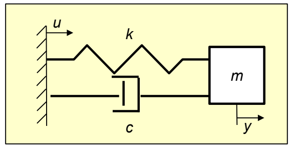
$$G(s) = \frac{Y(s)}{U(s)} = \frac{\left(b_0 s^m + b_1 s^{m-1} + \ldots + b_m\right)}{\left(a_0 s^n + a_1 s^{n-1} + \ldots + a_n\right)}$$

•The Transfer function is a property of the system itself and not a function of the input function

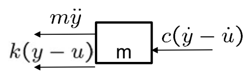
•Transfer function of many physically different systems can be identical. eg. Spring-mass-dashpot and RLC circuit.

Example

For the spring-mass dashpot system



Force balance leads to the free body diagram



$$m\ddot{y} + c\dot{y} + ky = ku + c\dot{u}$$

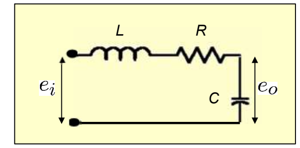
Assuming the initial conditions to be zero, and taking the Laplace transform of the above equation leads to:

$$\frac{Y(s)}{U(s)} = G(s) = \frac{cs+k}{ms^2+cs+k}$$

17

Example

For the R-L-C system



Kirchoff's law gives us

$$L\frac{di}{dt} + Ri + \frac{1}{C}\int i\,dt = e_i$$
$$\frac{1}{C}\int i\,dt = e_o$$

Assuming the initial conditions to be zero, and taking the Laplace transform of the above equation leads to:

$$\left(Ls + R + \frac{1}{Cs}\right) I(s) = E_i(s)$$
$$\frac{1}{Cs} I(s) = E_o(s)$$

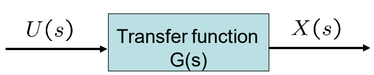
If e_i and e_o are the inputs and outputs respectively, we have

$$\frac{I(s)}{E_i(s)} = \frac{1}{Ls + R + \frac{1}{Cs}} = \frac{Cs}{LCs^2 + RCs + 1} \qquad \frac{E_o(s)}{I(s)} = \frac{1}{Cs}$$

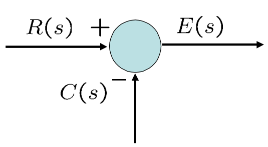
18

Block Diagrams

A block diagram of a system is a pictorial representative of the function performance by each component and of the flow of signals. The block is a symbol for the mathematical operation on the input signal to the block.

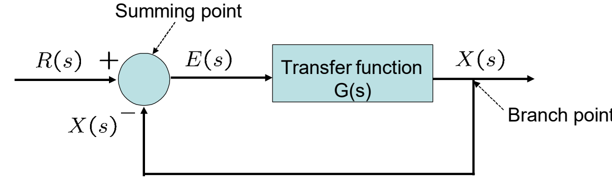


Error Detector: The error detector produces a signal which is the difference between the reference input and the feedback signal of the control system.

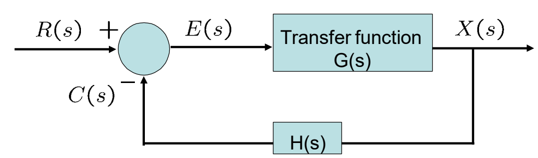


19

Block Diagram (Closed Loop)



When the output signal is fed back to the summing point for comparison with the input, it is necessary to convert the form of the output signal to that of the input signal. This is accomplished by the feedback element whose transfer function is H(S)



20

Block Diagram (Closed Loop)

The ratio of the feedback signal to the actuating error signal E(s) is called the open-loop transfer function. That is

$$\frac{C(s)}{E(s)} = G(s)H(s) \qquad \text{Open Loop TF}$$

The ratio of the output X(s) to the actuating signal E(s) is called the feedforward transfer function:

$$\frac{X(s)}{E(s)} = G(s) \qquad \text{Feedforward TF}$$

We can see from the closed loop block diagram:

$$X(s) = G(s)E(s)$$

$$E(s) = R(s) - C(s) = R(s) - H(s)X(s)$$

Eliminating E(s) from the above equations, we have

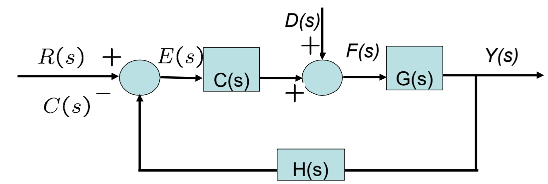
$$X(s) = G(s)\left(R(s) - H(s)X(s)\right)$$

$$\Rightarrow \frac{X(s)}{R(s)} = \frac{G(s)}{1 + G(s)H(s)} \qquad \text{Closed Loop TF}$$

This is called the closed loop transfer function

21

Block Diagram Reduction



$$E = R - HY \qquad U = CE$$

$$F = D + U \qquad Y = GF$$

$$E = R - HGF = R - HG(D + U)$$

$$E = R - HGD - HGCE$$

$$E = \frac{1}{1 + HGC}R - \frac{HG}{1 + HGC}D$$

22

Block Diagram Reduction

$$Y = GD + CGE$$

$$Y = GD + CG\left(\frac{1}{1 + HGC}R - \frac{HG}{1 + HGC}D\right)$$

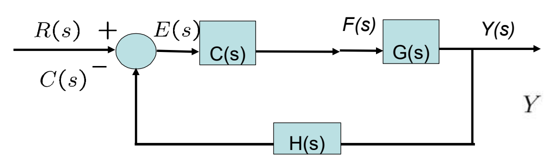
$$Y = G\frac{1 + HGC}{1 + HGC}D + CG\left(\frac{1}{1 + HGC}R - \frac{HG}{1 + HGC}D\right)$$

$$Y = \frac{G}{1 + HGC}D + \frac{CG}{1 + HGC}R$$

23

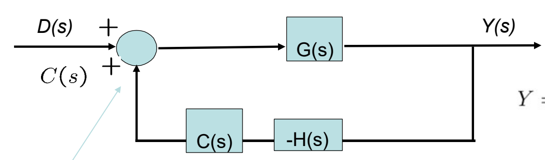
Block Diagram Reduction (Superposition)

Assume D = 0,



$$Y = \frac{CG}{1 + HGC}R$$

Assume R = 0,



$$Y = \frac{G}{1 - (-H)GC}D$$

Note: Positive Feedback)

$$Y = \frac{G}{1 + HGC}D + \frac{CG}{1 + HGC}R$$

24

First Order System

Transfer function:

$$\frac{Y(s)}{R(s)} = \frac{1}{Ts+1}$$

Response to a unit step input is: step: $R = \frac{1}{s}$

$$Y(s) = \frac{1}{Ts+1}\frac{1}{s}$$

Partial Fraction Expansion leads to: $Y = \frac{a_0}{s} + \frac{a_1}{Ts+1}$

$$Y(s) = \frac{1}{s} - \frac{T}{Ts+1} = \frac{1}{s} - \frac{1}{s+1/T}$$

Inverse Laplace transform leads to:

$$y(t) = 1 - e^{-\frac{t}{T}}$$ time constant

25

First Order System

At t = T, the output is:

$$y(t) = 1 - exp(-1) = 0.632$$ 1 time constant

*T represents the time required for the system response to reach 63.2% of the final value. T is referred to as the **Time Constant** of the system*

The slope of the system response at time = 0 is:

$$\frac{dy(t)}{dt} = \frac{1}{T}e^{-\frac{0}{T}} = \frac{1}{T}$$

Response of the first order system to a unit ramp is: $R(s) = \frac{1}{s^2}$

$$Y(s) = \frac{1}{Ts+1}\frac{1}{s^2} = \frac{1}{s^2} - \frac{T}{s} + \frac{T^2}{Ts+1}$$

$$y(t) = t - T + Te^{-\frac{t}{T}}$$

Tracking error as time tends to infinity is: $\to 0$ as $t \to \infty$

$$e(t \to \infty) = t - y(t) = T - T\cancel{e^{-\frac{t}{T}}} = T$$

26

First Order System

***Half-Life**: Time it takes for a quantity to reduce to half its initial value. For this example, velocity to decrease to half of its initial value.*

The log of the evolving output v(t) is given by the equation:

$$log(v(t)) = log(v_0) - \frac{c}{m}t$$

which can be rewritten as the equation:

$$log(v(t)) - log(v_0) = log(\frac{v(t)}{v_0}) = -\frac{c}{m}t$$

Have life is the time required for the output to reach half its initial value, therefore:

$$log(\frac{v(t)}{v_0}) = log(\frac{1}{2}) = -\frac{c}{m}t$$

27

Second Order System

Transfer function:

$$\frac{Y(s)}{R(s)} = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$

ω	natural frequency
ζ	damping ratio

$$\frac{Y(s)}{R(s)} = \frac{\omega^2}{(s+\zeta\omega)^2 + (\omega\sqrt{1-\zeta^2})^2}$$

Response to a unit step input is: $R = \frac{1}{s}$

$$Y(s) = \frac{\omega^2}{(s+\zeta\omega)^2 + (\omega\sqrt{1-\zeta^2})^2}\frac{1}{s}$$ $= \frac{a_0}{s} + \frac{a_1 s + a_2}{s^2 + 2\zeta\omega s + \omega^2}$

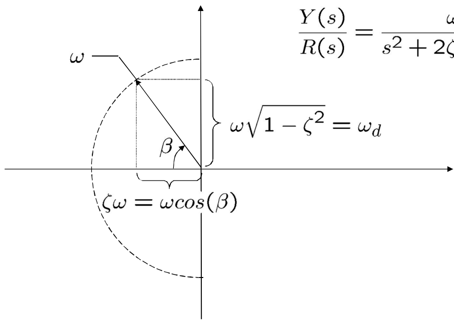
$$y(t) = 1 - e^{-\zeta\omega t}\left(cos(\omega_d t) + \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\omega_d t)\right)$$

28

S-plane representation

Laplace transform with zero initial conditions leads to:

$$\frac{Y(s)}{R(s)} = \frac{\omega^2}{s^2 + 2\zeta\omega s + \omega^2}$$



Poles Located at: $s = -\zeta\omega \pm j\omega\sqrt{1-\zeta^2}$

29

Transient Response

In analyzing and designing control systems, we must have a basis of comparison of performance of various control systems. This basis may be set up by specifying particular test input signals and by comparing the response of various systems to these input signals. Typical test signals: Step function, ramp function, impulse function, sinusoid function.

The time response of a control system consists of two parts: the transient and the steady-state response. Transient response corresponds to the behavior of the system from the initial state to the final state. By steady state, we mean the manner in which the system output behaves as time approaches infinity.

For a step input, the transient response can be characterized by:

Delay time t_d: time to reach half the final value for the first time.

Rise time t_r: time required for the response to rise from 10% to 90% for overdamped systems, and from 0% to 100% for underdamped systems

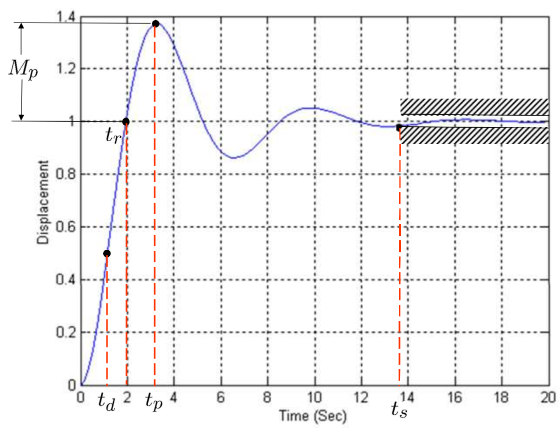
Peak time t_p: time required to reach the first peak of the overshoot

Percent Overshoot M_p: $M_p = \frac{y(t_p) - y(\infty)}{y(\infty)}$

Settling time t_s: time required for the response curve to reach and stay within 2% or 5% of the final value. Is a function of the largest time constant of the control system.

30

Transient Response



31

Transient Response

For a step input, the transient response can be characterized by:

Rise time t_r: time required for the response to rise from 10% to 90% for overdamped systems, and from 0% to 100% for underdamped systems

$$y(t_r) = 1 = 1 - e^{-\zeta\omega t_r}\left(cos(\omega_d t_r) + \frac{\zeta}{\sqrt{1-\zeta^2}} sin(\omega_d t_r)\right)$$

Since $e^{-\zeta\omega t_r} \neq 0$

$$cos(\omega_d t_r) + \frac{\zeta}{\sqrt{1-\zeta^2}} sin(\omega_d t_r) = 0$$

$$\Rightarrow tan(\omega_d t_r) = -\frac{\sqrt{1-\zeta^2}}{\zeta} = -\frac{\omega_d}{\sigma}$$

Thus, the rise time is:

$$t_r = \frac{1}{\omega_d} tan^{-1}\left(\frac{\omega_d}{-\sigma}\right) = \frac{\pi - \beta}{\omega_d}$$

32

Transient Response

For a step input, the transient response can be characterized by:

Peak time t_p: time required to reach the first peak of the overshoot

$$\frac{dy}{dt} = 0 = \zeta\omega e^{-\zeta\omega t}\left(\cancel{cos(\omega_d t)} + \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\omega_d t)\right) - e^{-\zeta\omega t}\left(-\omega_d \sin(\omega_d t) + \frac{\omega_d\zeta}{\sqrt{1-\zeta^2}}\cancel{\cos(\omega_d t)}\right)$$

$$\frac{dy}{dt} = 0 = e^{-\zeta\omega t_p}sin(\omega_d t_p)\frac{\omega}{\sqrt{1-\zeta^2}}$$

Thus, the peak time is:

$$\omega_d t_p = 0, \pi, 2\pi, \ldots$$

$$t_p = \frac{\pi}{\omega_d}$$

33

Transient Response

For a step input, the transient response can be characterized by:

Percent Overshoot M_p:

$$M_p = y(t_p) - 1 = -e^{-\zeta\omega\frac{\pi}{\omega_d}}\left(cos(\pi) + \frac{\zeta}{\sqrt{1-\zeta^2}}sin(\pi)\right)$$

$$= e^{-\frac{\sigma}{\omega_d}\pi}$$

$$= e^{-\frac{\zeta}{\sqrt{1-\zeta^2}}\pi}$$

Maximum overshoot is function of damping ratio only.

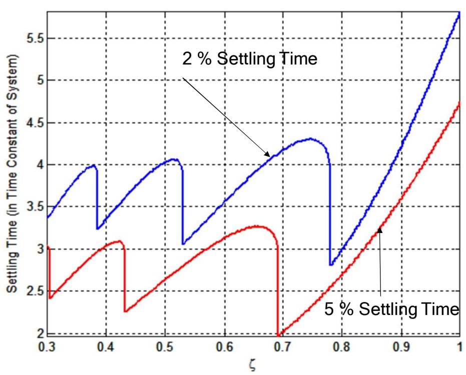
Maximum percent overshoot is: $M_p = e^{-\frac{\zeta}{\sqrt{1-\zeta^2}}\pi} \times 100$

34

Transient Response

For a step input, the transient response can be characterized by:

Settling time t_s:



2% Settling time:

$$t_s = 4T = \frac{4}{\sigma} = \frac{4}{\zeta\omega}$$

5% Settling time:

$$t_s = 3T = \frac{3}{\sigma} = \frac{3}{\zeta\omega}$$

35

Routh Stability Criterion

Consider the closed loop transfer function of the form:

$$G(s) = \frac{b_0 s^m + b_1 s^{m-1} + \ldots + b_{m-1}s + b_m}{a_0 s^n + a_1 s^{n-1} + \ldots + a_{n-1}s + a_n}$$

where a's and b's are constant and where m< n.

If any of the coefficients are zero or negative in the presence of at least one positive coefficient, there is a root or roots which are imaginary or which have positive real parts.

If all the coefficients are positive, arrange the coefficients of the polynomial in rows and columns according to the pattern:

36

Routh Stability Criterion

s^n	a_0	a_2	a_4	a_6	…
s^{n-1}	a_1	a_3	a_5	a_7	…
s^{n-2}	b_1	b_2	b_3	b_4	…
s^{n-3}	c_1	c_2	c_3	c_4	…
s^{n-4}	d_1	d_2	d_3	d_4	…
…	…	…	…	…	…
s^2	e_1	e_2			
s^1	f_1				
s^0	g_1				

$$b_1 = \frac{a_1a_2 - a_0a_3}{a_1}$$

$$b_2 = \frac{a_1a_4 - a_0a_5}{a_1}$$

$$b_3 = \frac{a_1a_6 - a_0a_7}{a_1}$$

…

$$c_1 = \frac{b_1a_3 - a_1b_2}{b_1}$$

$$c_2 = \frac{b_1a_5 - a_1b_3}{b_1}$$

…

$$d_1 = \frac{c_1b_2 - b_1c_2}{c_1}$$

37

Routh Stability Criterion

Routh's stability criterion states that the number of roots of the system G(s) with positive real parts is equal to the number of changes in the sign of the coefficients of the first column of the array.

The necessary and sufficient condition that all poles of G(s) lie in the left half plane is that all the coefficient of the denominator of G(s) be positive and all terms in the first column of the array have positive signs.

Special Case:

If a first-column term in any row is zero, but the remaining terms are not zero or there is no remaining term, then the zero term is replaced by a very small positive number ϵ and the rest of the array is evaluated.

$$s^3 + 2s^2 + s + 2 = 0$$

s^3	1	1
s^2	2	2
s	$0 \approx \epsilon$	
s^0	2	

If the sign of the coefficient above the zero is the same as that below it, it indicates that there are a pair of poles on the imaginary axis.

38

Routh Stability Criterion

Special Case:

If all the coefficients in any derived row are zero, it indicates that there are roots of equal magnitude lying radially opposite in the s-plane, eg. two real roots with equal magnitudes and opposite signs and/or two conjugate imaginary roots

$$s^5 + 2s^4 + 24s^3 + 48s^2 - 25s - 50 = 0$$

s^5	1	24	-25
s^4	2	48	-50
s^3	0	0	

In such a case, the evaluation of the rest of the array can be continued by forming an auxiliary polynomial with the coefficient of the last row and the coefficients of the derivative of this polynomial in the next row

$$P(s) = 2s^4 + 48s^2 - 50$$

$$\frac{dP(s)}{ds} = 8s^3 + 96s$$

39

Routh Stability Criterion

The closed loop system can be represented as:

$$\frac{X_1(s)}{U(s)} = \frac{(K_1 + K_2s)(s^2+1)}{s^4 + 2s^2 + (K_2s + K_1)(s^2+1)}$$

The Routh table for the closed loop system is:

s^4	1	$2 + K_1$	K_1
s^3	K_2	K_2	
s^2	$\dfrac{K_2 + K_1K_2}{K_2}$	K_1	
s^1	$\dfrac{K2}{1+K1}$		
s^0	K_1		

The second and fifth row indicate that :

$$K_1 > 0 \text{ and } K_2 > 0$$

40

Routh Stability Criterion

The third and fourth row requires:

$$(1 + K_1) > 0 \text{ or } K_1 > -1$$

The requirement:

$$K_1 > 0 \text{ and } K_1 > -1$$
$$\Rightarrow \quad K_1 > 0$$

Results in a stable controller for all gains which satisfy the inequality constraints:

$$K_1 > 0 \text{ and } K_2 > 0$$

41

Routh Stability Criterion (Relative Stability)

Relative Stability: The Routh criteria is a tools which provides a binary answer to the question of *absolute stability*, i.e., whether the system is stable or not. Relative stability permits comparing two system and gauging which system is relatively more stable. A simple characterization of relative stability is the distance of the pole with the largest real value from the imaginary axis. Closer the pole is to the imaginary axis, smaller is its relative stability. Once can use the Routh Criteria to determine the number of poles that lie to the right of a shifted imaginary axis. Substitute s = z − σ (σ = constant) into the characteristic equation of the system and rewrite the characteristic equation in terms of z. Applying the Routh Criteria to new polynomial in z permits one to determine the number of poles to the right of the vertical line s = −σ. Thus, this test reveals the number of roots which lie to the right of the vertical line s = −σ

42

Static Error Coefficients

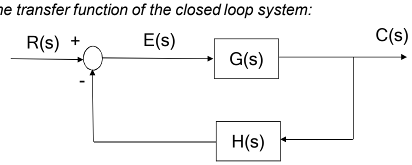
Consider the open loop transfer function:

$$G(s)H(s) = \frac{K(T_a s + 1)(T_b s + 1)...(T_{m-1}s + 1)(T_m s + 1)}{s^N (T_1 s + 1)(T_2 s + 1)...(T_n s + 1)}$$

which includes N poles at the origin of the s-plane.

A system is called type 0, type 1, type 2,…., if N = 0,1,2,…, respectively.

The transfer function of the closed loop system:



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

43

Static Error Coefficients

and the error transfer function can be calculated as:

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

The steady state error can be calculated as:

$$e_{ss} = \lim_{t \to \infty} e_{ss}(t) = \lim_{s \to 0} \frac{sR(s)}{1 + G(s)H(s)}$$

44

Static Position Error Coefficients K_p

The steady-state error of a system subject to an unit step input is:

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

The static position error coefficient K_p is defined as:

$$K_p = \lim_{s\to 0} G(s)H(s) = G(0)H(0)$$

The steady state error is given by the equation:

$$e_{ss} = \frac{1}{1+K_p}$$

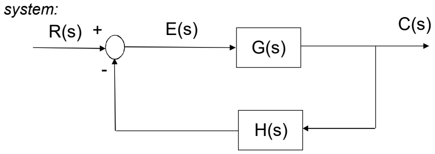
45

Root Locus

The locus of the roots of the closed loop system as a function of a system parameter, as it is varied from 0 to infinity results in the moniker, Root-Locus.

The root-locus permits determination of the closed-loop poles given the open-loop poles and zeros of the system.

For the system:



the closed loop transfer function is:

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

The characteristic equation:

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

46

Root Locus

when solved results in the poles of the closed-loop system.

The characteristic equation can be written as:

$$G(s)H(s) = -1$$
$$G(s)H(s) = e^{\pi j}$$

Since G(s)H(s) is a complex quantity, the equation can be rewritten as:

Angle criterion: $\angle G(s)H(s) = \pm\pi(2k+1),\ k=0,1,2,3,\ldots$

Magnitude criterion: $|G(s)H(s)| = 1$

The values of s which satisfy the magnitude and angle criterion lie on the root locus.

Solving the angle criterion alone results in the root-locus. The magnitude criterion locates the closed loop poles on the locus.

47

Root Locus

Often the open-loop transfer function G(s)H(s) involved a gain parameter K, resulting in the characteristic equation:

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

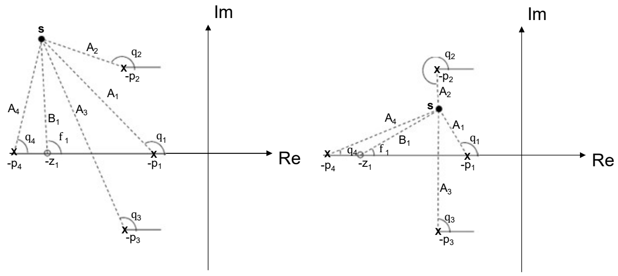
Then the root-loci are the loci of the closed loop poles as K is varied from 0 to infinity.

To sketch the root-loci, we require the poles and zeros of the open-loop system. Now, the angle and magnitude criterion can be schematically represented as, for the system:

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

48

Root Locus



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

$$|G(s)H(s)| = \frac{KB_1}{A_1A_2A_3A_4}$$

49

Root Locus (Example)

Asymptotes of the root-loci as s tends to infinity can be determined as follows: Select a test point far from the origin:

$$\lim_{s\to\infty} G(s) = \lim_{s\to\infty} \frac{K}{s(s+1)(s+2)} = \lim_{s\to\infty} \frac{K}{s^3}$$

and the angle criterion is:

$$-3\angle s = \pm\pi(2k+1) \text{ k=0,1,2,3,...}$$

or, the angle of the asymptotes are:

$$\text{angle of asymptotoes } = \pm\frac{\pi(2k+1)}{3} \text{ k=0,1,2,3,...}$$

For this example, they are 60º, -60º, -180º.

To draw the asymptotes, we must find the point where they intersect the real-axis. Since, the transfer function for a test point s far from the origin permits us to approximate the transfer function as:

$$G(s) = \frac{K}{s(s+1)(s+2)} = \frac{K}{s^3+3s^2+\ldots} \approx \frac{K}{(s+1)^3}$$

50

Root Locus (Example)

The angle criterion for the approximated system is:

$$\angle\frac{K}{(s+1)^3} = -3\angle(s+1) = \pm\pi(2k+1)$$

$$\Rightarrow \angle(s+1) = \pm\frac{\pi}{3}(2k+1) \text{ k=0,1,2,3,...}$$

Substituting $s = \sigma + j\omega$ *, we have*

$$\angle\sigma + j\omega + 1 = \pm\frac{\pi}{3}(2k+1)$$

$$tan^{-1}\left(\frac{\omega}{\sigma+1}\right) = \frac{\pi}{3}, -\frac{\pi}{3}, -\pi$$

Taking the tangent of both sides, we have:

$$\left(\frac{\omega}{\sigma+1}\right) = \sqrt{3}, -\sqrt{3}, 0$$

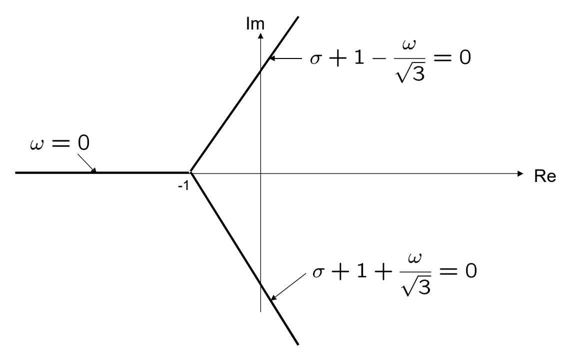
which can be written as:

$$\sigma + 1 - \frac{\omega}{\sqrt{3}} = 0, \sigma + 1 + \frac{\omega}{\sqrt{3}} = 0, \omega = 0$$

51

Root Locus (Example)

The three equations result is three straight lines shown below, which are the asymptotes, which meet at the point s=-1. Thus, the abscissa of the intersection of the asymptotes and the real-axis is determined by setting the denominator of the approximated transfer function to zero and solving for s.



52

Root Locus (Example)

The root loci starting from 0 and -1, break away from the real axis as K is increased. The breakaway point corresponds to a point in the s-plane where multiple roots of the characteristic equation occur.

Representing the characteristic equation as:

$$f(s) = B(s) + KA(s) = 0 \qquad (1)$$

We know that f(s) has multiple roots at $s = s_1$, if

$$\left.\frac{df(s)}{ds}\right|_{s=s_1} = \frac{dB(s)}{ds} + K\frac{dA(s)}{ds} = 0$$

$$\Rightarrow K = -\frac{B'(s)}{A'(s)}, \text{ where } A'(s) = \frac{dA(s)}{ds}, B'(s) = \frac{dB(s)}{ds}$$

Which when substituted into Equation 1, gives us:

$$f(s) = B(s) - \frac{B'(s)}{A'(s)}A(s) = 0$$

$$B(s)A'(s) - B'(s)A(s) = 0$$

53

Root Locus (Example)

Consider the equation:

$$f(s) = B(s) + KA(s) = 0$$

which can be rewritten as:

$$K = -\frac{B(s)}{A(s)}$$

The derivative of K with respect to s gives:

$$\frac{dK}{ds} = \frac{B(s)A'(s) - B'(s)A(s)}{A^2(s)}$$

Which when set to zero gives the same constraint as before:

$$B(s)A'(s) - B'(s)A(s) = 0$$

Therefore, the breakaway point can be determined by setting:

$$\frac{dK}{ds} = \frac{B(s)A'(s) - B'(s)A(s)}{A^2(s)} = 0$$

54

Root Locus (Example)

For the current example, we have:

$$\frac{K}{s(s+1)(s+2)} + 1 = 0$$

$$K = -(s^3 + 3s^2 + 2) = 0$$

The derivative of K with respect to s gives:

$$\frac{dK}{ds} = -(3s^2 + 6s + 2) = 0$$

or

$$s = -0.4226, s = -1.5774$$

Since the breakaway point must lie between 0 and -1, s = -0.4226 corresponds to the actual breakaway point and the corresponding gain is:

$$K = 0.3849$$

The gain corresponding to the point s = -1.5774 is:

$$K = -0.3849$$

Which belongs to the complementary root-locus.

55

Root Locus (Example)

Locations where the Root-loci crosses the imaginary axis can be determined by substituting s = jwin the characteristic equation as solving for K and s.

For the current example, we have

$$(j\omega)^3 + 3(j\omega)^2 + 2(j\omega) + K = 0$$

or

$$K - 3\omega^2 + j(2\omega - \omega^3) = 0$$

Equating the real and imaginary parts to zero, we have

$$K - 3\omega^2 = 0, \text{ and } (2\omega - \omega^3) = 0$$

which results in the solutions:

$$\omega = \pm\sqrt{2}, K = 6, \text{ or } \omega = 0, K = 0$$

Thus, the root locus cross the imaginary axis at $s = \pm j\sqrt{2}$

56

Root Locus (Example)

Consider the system with an open-loop transfer function:

$$G(s) = \frac{K(s+2)}{s^2+2s+3}$$

which has open loop poles at:

$$s = -1 + j\sqrt{2} \quad s = -1 - j\sqrt{2}$$

and a zero at $s = -2$

From the rule about existence of root-loci on the real axis, it is clear that to the left of the point s=-2, the entire real axis is part of the root-loci.

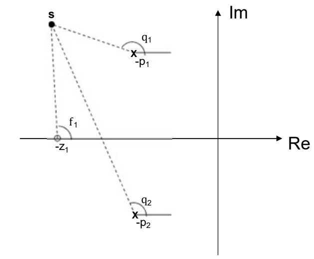
Since root-loci start at open-loop poles and end at open-loop zeros, we know that the two root-loci which start at the complex open loop poles should break in on the real axis and one of the root-loci should move towards s = -2, and the other towards $-\infty$

How does the root locus depart from $s = -1 \pm j\sqrt{2}$ *?*

57

Root Locus (Example)

Consider a test point s in the vicinity of the complex open-loop pole p_1*:*



The angular contribution of the pole p_2 *and the zero* z_1 *to the test point s can be considered to be the same as the angle made by the lines connecting the pole* p_2 *and zero* z_1 *to the pole* p_1 *(if the test point is close to* p_1*).*

Then, the angle of departure t1 is given by the angle criterion:

$$\phi_1 - (\theta_1 + \theta_2) = \pm\pi(2k+1)$$

$$\theta_1 = \pi - \theta_2 + \phi_1 = \pi - \frac{\pi}{2} + atan(\frac{\sqrt{2}}{1})$$

58

Root Locus (sketching rules)

Rule 1: Root locus is symmetric about the real axis

- *roots of characteristic equations are either real or complex conjugates*
- *Can draw upper half plane only, although typically full s-plane loci are sketched*

Rule 2: Each branch of the root locus originates at the open-loop pole with k=0, and terminates either at an open-loop zero or at infinity as k tends to infinity.

Characteristic Equation: $$\prod_{j=1}^{n}(s+p_j) + k\prod_{i=1}^{m}(s+z_i) = 0$$

When k=0, poles at: $s = -p_j$ j=1,2,3,…,n *which are open-loop poles*

Rewrite Characteristic Equation: $$\frac{1}{k}\prod_{j=1}^{n}(s+p_j) + \prod_{i=1}^{m}(s+z_i) = 0$$

When k-> infinity, poles at: $s = -z_i$ i=1,2,3,…,m *which are open-loop zeros*

-m branches of the root loci terminate at open-loop zeros

- remaining (n-m) branches go to infinity along asymptotes as k-> infinity

59

Root Locus (sketching rules)

Rule 3: A point on the real axis lies on a locus if number of open-loop poles plus zeros on the real axis to the right of this point are odd.

Rule 4: The (n-m) branches of the root loci which tend to infinity do so along straight line asymptotes whose angles are:

$$\phi_A = \frac{(2k+1)\pi}{(n-m)} \quad k = 0,1,2,3,\ldots$$

Rule 5: The asymptotes cross the real axis at a point given by:

$$-\sigma_A = \frac{\sum_{j=1}^{n}(-p_j) - \sum_{i=1}^{m}(-z_i)}{(n-m)}$$

$$= \frac{\sum \text{real part of poles} - \sum \text{real part of zeros}}{\text{number of poles - number of zeros}}$$

60

Root Locus (sketching rules)

Rule 6: Break points (points at which multiple roots of the characteristic equation occur) of root locus are solutions of dk/ds = 0.

$\phi_p = \pm\pi(2k+1) + \phi$ k = 0,1,2,3,...

Rule 7: The angle of departure from an open-loop pole is given by:

$\phi_z = \pm\pi(2k+1) - \phi$ k = 0,1,2,3,...

where f is the net angle contribution at the pole of all other open-loop poles and zeros. Similarly, the angle of arrival at an open-loop zero is given by:

where f is the net angle contribution at the pole of all other open-loop poles and zeros

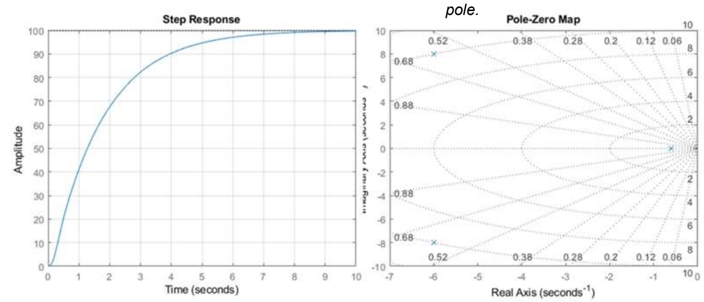
61

Dominant Poles

- The "slowest poles" of a system (those closest to the imaginary axis in the s-plane) give rise to the longest lasting terms in the transient response of the system.
- If a pole or set of poles are very slow compared to others in the transfer function, then they generally dominate the transient response and are referred to as the "Dominant Poles".

Example: $G(s) = \frac{6000}{(s^2+12s+100)(s+0.6)}$

The step response appears like a first order system response since the pole at s=-0.6 is the dominant pole.



62

Frequency Response Method

Transfer function:

$$\frac{Y(s)}{R(s)} = G(s) = \frac{p(s)}{q(s)} = \frac{p(s)}{(s+s_1)(s+s_2)(s+s_3)\ldots(s+s_n)}$$

For a stable system, the real parts of s_i lie in the left half of the complex plane. The response of the system to a sinusoidal input of amplitude X, is:

$$Y(s) = \frac{p(s)}{q(s)}\frac{\omega X}{s^2+\omega^2}$$

$$Y(s) = \frac{a}{s+j\omega} + \frac{\bar{a}}{s-j\omega} + \frac{b_1}{(s+s_1)} + \frac{b_2}{(s+s_2)} + \frac{b_3}{(s+s_3)} + \ldots + \frac{b_n}{(s+s_n)}$$

Where a and $\bar{a}$ are complex conjugate constants and b_i are constants.

63

Bode Diagram

The complex quantity G(jω) can be represented as:

$$G(j\omega) = |G(j\omega)|e^{j\phi}$$

where

$$\phi = tan^{-1}\left(\frac{Im(G(j\omega))}{Re(G(j\omega))}\right)$$

Thus, the transfer function can be represented by two plots: one of the magnitude of G(jω) vs. frequency and j vs frequency.

By plotting the logarithmic value of the magnitude of G(jω), multiplication operations are replaced by additions. This is the motivation for plotting

$20log(|G(j\omega)|)$ (decibels) vs. frequency which helps is rapidly plotting asymptotic approximation of the magnitude plot.

The basic factors which occur in a transfer function are:

-gain K

-Integral and derivative factor $(j\omega)^{-1}$

-First order factors $(1+j\omega T)^{-1}$

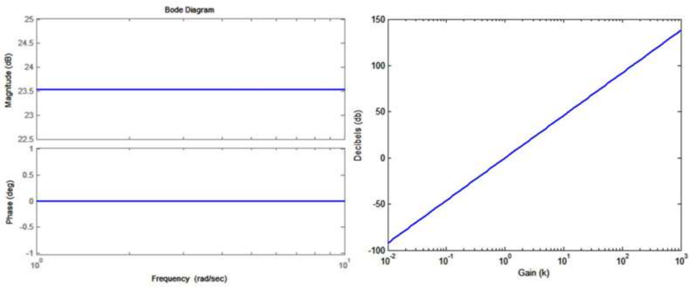
-quadratic factors $(1+2\zeta\frac{j\omega}{\omega_n}+(\frac{j\omega}{\omega_n})^2)^{-1}$

64

Bode Diagram

Gain K: K is a real number and is not a function of frequency, so the magnitude plot does not change with frequency. Moreover, since the imaginary part of K is zero, the phase is zero for all frequency.

(NOTE: the matlab command to determine the db is 20*log10(K))



65

Bode Diagram

Integral factor : The magnitude of an integrator transfer function in db is:

$$20log\left(\left|\frac{1}{j\omega}\right|\right) = -20log(\omega)$$

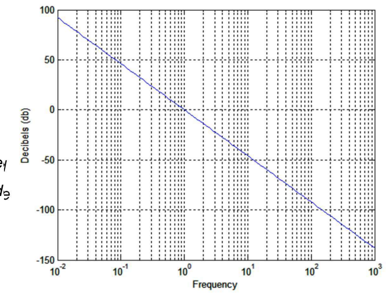
The phase of an integrator transfer function is: $\phi = -tan^{-1}\left(\frac{\omega}{0}\right) = -\frac{\pi}{2}$

The magnitude for w= 1 is 0

The magnitude for w= 10 is -20

The magnitude for w= 100 is -40

It is clear if the magnitude if plotted on a semi-log graph, the magnitude plot is linear.



66

Bode Diagram

Derivative factor : The magnitude of an derivative transfer function in db is:

$$20log\left(|j\omega|\right) = 20log(\omega)$$

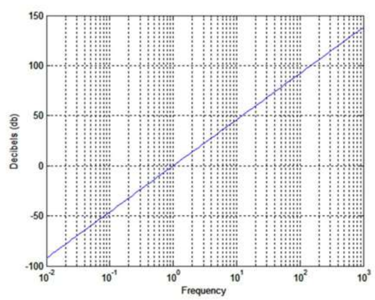
The phase of an differentiator transfer function is: $\phi = tan^{-1}\left(\frac{\omega}{0}\right) = \frac{\pi}{2}$

The magnitude for w= 1 is 0

The magnitude for w= 10 is 20

The magnitude for w= 100 is 40

It is clear if the magnitude if plotted on a semi-log graph, the magnitude plot is linear.



67

Bode Diagram

First order factor : The magnitude of a first order transfer function in db is:

$$20log\left(\left|\frac{1}{1+j\omega T}\right|\right) = -20log(\sqrt{1+\omega^2T^2})$$

For wT << 1, the log magnitude can be approximated by $-20log(\sqrt{1}) = 0$

and the phase is zero

For wT >> 1, the log magnitude can be approximated by

$$-20log(\sqrt{\omega^2T^2}) = -20log(\omega T)$$

and the phase is -90º.

*For w= 1/T, called the **Corner Frequency**, the log magnitude can be approximated by*

$$-20log(\sqrt{\omega^2T^2}) = -20log(1) = 0$$

and for large frequencies, the rest of the magnitude plot is the same as that of an integrator and the phase transitions from 0 to -90º.

Straight line approximation for the phase plot of a first order transfer function:

$$G(j\omega) \approx \begin{cases} 0 & \text{if } \omega T < \frac{1}{10}\omega T \\ -45 - 45(\log(\omega T)) & \text{if } \frac{1}{10}\omega T < \omega T < 10\omega T \\ -90 & \text{if } > 10\omega T \end{cases}$$

68

Bode Diagram

First order factor : The magnitude of a first order transfer function in db is:

$$20log\,(|1 + j\omega T|) = 20log(\sqrt{1 + \omega^2 T^2})$$

For wT << 1, the log magnitude can be approximated by $20log(\sqrt{1}) = 0$

and the phase is zero

For wT >> 1, the log magnitude can be approximated by

$$20log(\sqrt{\omega^2 T^2}) = 20log(\omega T)$$

and the phase is 90º.

For w= 1/T, called the **Corner Frequency**, the log magnitude can be approximated by

$$20log(\sqrt{\omega^2 T^2}) = 20log(1) = 0$$

and for large frequencies, the rest of the magnitude plot is the same as that of an differentiator and the phase transitions from 0 to 90º.

Straight line approximation for the phase plot of a first order transfer function:

$$G(j\omega) \approx \begin{cases} 0 & \text{if } \omega T < \frac{1}{10}\omega T \\ 45 + 45(\log(\omega T)) & \text{if } \frac{1}{10}\omega T < \omega T < 10\omega T \\ 90 & \text{if } > 10\omega T \end{cases}$$

69

Bode Diagram

Second order factor : Consider a second order transfer function:

$$G(j\omega) = \frac{1}{1 + 2\zeta(\frac{j\omega}{\omega_n}) + (\frac{j\omega}{\omega_n})^2}$$

The magnitude of the transfer function is:

$$20log\left|\frac{1}{1 + 2\zeta(\frac{j\omega}{\omega_n}) + (\frac{j\omega}{\omega_n})^2}\right| = -20log\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}$$

For w<< w_n, the log magnitude can be approximated by $-20log(\sqrt{1}) = 0$

and the phase is zero

For w>> w_n, the log magnitude can be approximated by

$$-20log\left(\frac{\omega^2}{\omega_n^2}\right) = -40log\left(\frac{\omega}{\omega_n}\right)$$

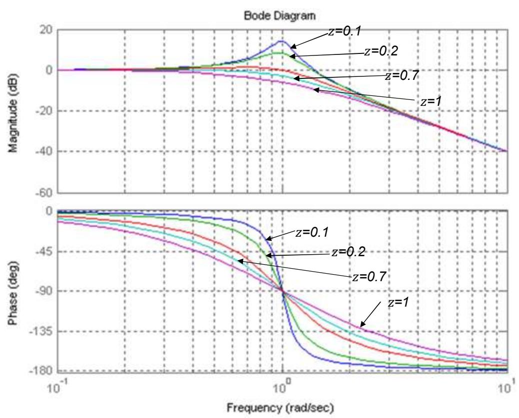
and the phase is 180º.

For $w = w_n$, called the **Corner Frequency**, the log magnitude can be approximated by $-40log(\frac{\omega_n}{\omega_n}) = -40log(1) = 0$

70

Bode Diagram

The two asymptotes determined are not function of z. Near the corner frequency a resonant peak occurs whose magnitude is a function of ζ.



71

Bode Diagram

The phase angle of a second order factor is:

$$\phi = -tan^{-1}\left(\frac{2\zeta\frac{\omega}{\omega_n}}{1 - (\frac{\omega}{\omega_n})^2}\right)$$

The phase angle is a function of w_n and z.

For w>> w_n, the log phase can be approximated by 180º.

For $w = w_n$, called the **Corner Frequency**, the phase is

$$\phi = -tan^{-1}\left(\frac{2\zeta}{0}\right) = -tan^{-1}(\infty) = -90$$

72

Bode Diagram

Resonant Peak Value: The magnitude of the transfer function

$$G(j\omega) = \frac{1}{1 + 2\zeta(\frac{j\omega}{\omega_n}) + (\frac{j\omega}{\omega_n})^2}$$

$$|G(j\omega)| = \frac{1}{\sqrt{\left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2}}$$

Which has a peak value at some frequency which is called the resonant frequency. The occurs when the denominator of the magnitude equation is a minimum.

$$g = \left(1 - \frac{\omega^2}{\omega_n^2}\right)^2 + \left(2\zeta\frac{\omega}{\omega_n}\right)^2 = 1 + \frac{\omega^4}{\omega_n^4} - 2\frac{\omega^2}{\omega_n^2} + 4\zeta^2\frac{\omega^2}{\omega_n^2}$$

$$\frac{dg}{d\omega} = 0 \Rightarrow \omega = \omega_n\sqrt{1 - 2\zeta^2}$$

Thus, there is no peak for damping ratio > 0.707

73

Bode Diagram (non-minimum Phase Systems)

Transfer function which have no poles or zeros in the right-half plane are minimum-phase systems. Systems with poles or zeros in the right-half plane are called non-minimum phase systems.

For systems with the same magnitude plot, the range of the phase angle of the minimum-phase systems is minimum among all such systems, while the range of phase angle of any nonminimum-phase systems is greater than the minimum.

Consider the example of two systems:

$$G_1(j\omega) = \frac{1 + j\omega T}{1 + j\omega T_1}, \quad G_2(j\omega) = \frac{1 - j\omega T}{1 + j\omega T_1}, \quad 0 < T < T_1$$

The magnitudes of the two transfer function are:

$$|G_1(j\omega)| = \frac{\sqrt{1 + \omega^2 T^2}}{\sqrt{1 + \omega^2 T_1^2}}, \quad |G_2(j\omega)| = \frac{\sqrt{1 + \omega^2 T^2}}{\sqrt{1 + \omega^2 T_1^2}},$$

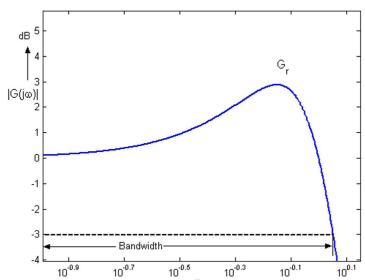
The phase of the two transfer function are:

$$\angle G_1(j\omega) = tan^{-1}(\omega T) - tan^{-1}(\omega T_1)$$
$$\angle G_2(j\omega) = -tan^{-1}(\omega T) - tan^{-1}(\omega T_1)$$

74

Bode Diagram

<u>Resonant Peak:</u> The resonant peak G_r is the maximum value of $|G(jw)|$



- *G_r is indicative of the relative stability of a stable closed loop system.*
- *A large G_r corresponds to a large maximum overshoot for a step input.*
- *Generally a desirable value for G_r is between 1.1 and 1.5*

<u>Resonant Frequency:</u> The resonant frequency w_r is the frequency at which the peak resonance G_r occurs.

75

Bode Diagram

<u>Bandwidth:</u> The bandwidth BW is the frequency at which $|G(jw)|$ drops to 70.7% or 3bD down from its zero frequency value. $20log10(\frac{1}{\sqrt{2}}) = -3db,$

- *The bandwidth is indicative of the transient response properties in the time-domain*
- *A large bandwidth corresponds to a faster rise time, since higher frequencies are more easily passed through the system*
- *Bandwidth is indicative of noise-filtering characteristics and the robustness of the system.*

<u>Second order system:</u>

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2},$$

Bandwidth is:

$$|G(j\omega)| = \frac{1}{\sqrt{(1 - \frac{\omega^2}{\omega_n^2})^2 + (2\zeta\frac{\omega}{\omega_n})^2}} = \frac{1}{\sqrt{2}} = 0.707$$

76

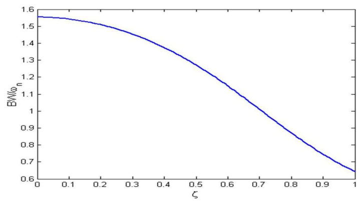
Bode Diagram

$$\sqrt{(1-\frac{\omega^2}{\omega_n^2})^2+(2\zeta\frac{\omega}{\omega_n})^2}=\sqrt{2}$$

$$\frac{\omega^2}{\omega_n^2}=(1-2\zeta^2)\pm\sqrt{4\zeta^4-4\zeta^2+2}$$

The positive sign should be chosen since the term on the left hand side is a positive real quantity for any z.

Bandwidth is: $\omega=\omega_n\sqrt{(1-2\zeta^2)\pm\sqrt{4\zeta^4-4\zeta^2+2}}$



BW/w_n decreases monotonically with z

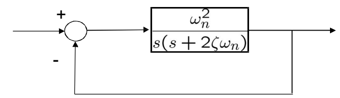
For a fixed z, BW increases with increasing w_n

77

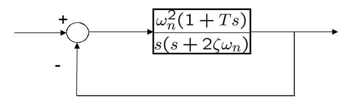
Effect of Adding a Zero to the Forward Path Transfer Function

The system: $G(s)=\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$

can be represented in unity feedback form as:



Adding a zero at s = -1/T, results in



With a closed loop transfer function

$$G(s)=\frac{\omega_n^2(1+Ts)}{s^2+(2\zeta\omega_n+T\omega_n^2)s+\omega_n^2}$$

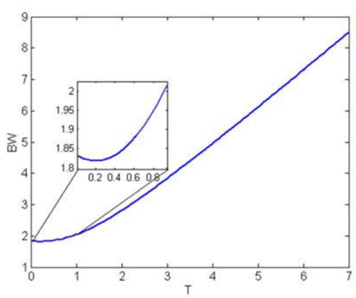
78

Effect of Adding a Zero to the Forward Path Transfer Function

In principle G_r, w_r and BW of the system can be derived as in the second order case. However since there are now three parameters z, w_n, & T, the exact expressions for G_r, w_r & BW are difficult to derive analytically. After a lengthy derivation, it can be shown that:

$$BW=(-b\pm\frac{1}{2}\sqrt{b^2+4\omega_n^4})^{0.5}$$

$$b=4\zeta^2\omega_n^2+4\zeta\omega_n^3T-2\omega_n^2-\omega_n^4T^2$$



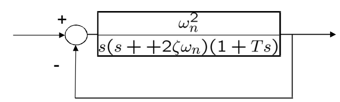
The general effect of adding a zero to the forward path transfer function is to increase the bandwidth of the closed-loop system.

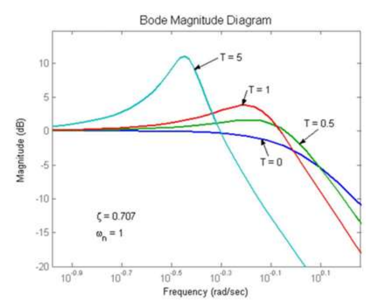
However, over a small range of T, the bandwidth actually decreases as seen in the adjacent figure.

79

Effect of Adding a Pole to the Forward Path Transfer Function

Adding a pole at s = -1/T, results in





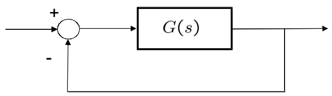
A qualitative feel can be acquired from the magnitude plot of the Bode diagram for the variation of the BW as a function of T.

The effect of adding a pole to the forward-path transfer function is to make the closed-loop system less stable, while decreasing the bandwidth.

80

Gain/Phase Margin

For the unity feedback system.



The characteristic equation is 1+G(s). When |G(s)|=1 with phase of -180°, the characteristic equation has infinite gain which corresponds to the closed loop poles lying on the imaginary axis.

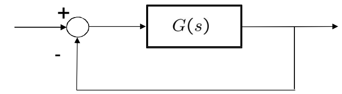
Gain margin is the amount of gain in decibels (dB) that can be added to the loop before the closed-loop system becomes unstable.

Phase margin is defined as the phase in degrees that can be added (i.e. from a transport delay), before the closed-loop system becomes unstable.

81

Bode Stability Criterion

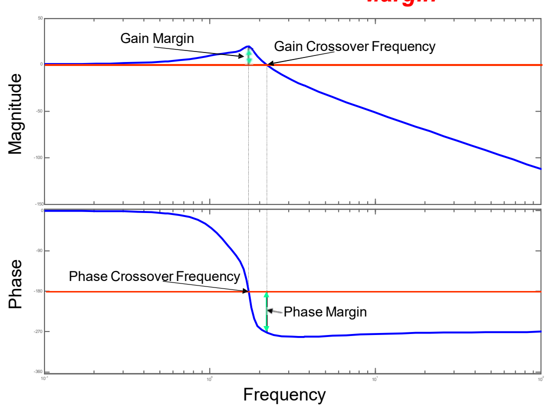
For the unity feedback system.



Bode Stability Criterion: Consider an open-loop transfer function GOL=G(s) that is strictly proper (more poles than zeros) and has no poles located on or to the right of the imaginary axis, with the possible exception of a single pole at the origin. Assume that the open-loop frequency response has only a single critical frequency and a single gain crossover frequency . Then the closed-loop system is stable if magnitude(G(s)) is < 1. Otherwise it is unstable.

82

Gain and Phase Margin

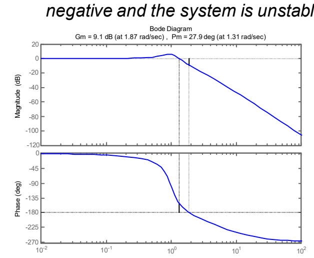


83

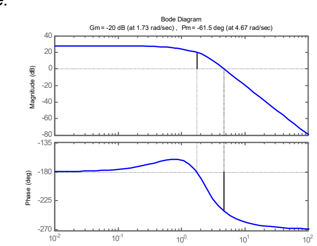
Gain Phase Margin

<u>Gain Margin:</u> The gain margin is positive and the system is stable if the magnitude of G(jw) at the phase crossover is negative in db. That is, the gain margin is measured below the 0 db line. If the gain margin is measured above the 0 db line, the gain margin is negative and the system is unstable.

<u>Phase Margin:</u> The phase margin is positive and the system is stable if the phase of G(jw) at the gain crossover is greater than -180°. That is, the phase margin is measured above the -180° axis. If the phase margin is measured below the -180° axis, the phase margin is negative and the system is unstable.



Stable system



Unstable system

84

Lead Compensator Design

The primary function of the lead compensator is to reshape the frequency response curve to provide sufficient phase lead angle to offset the excessive phase lag associated with the components of the fixed system

- Determine the open-loop gain k to satisfy the requirements on the error coefficients
- Using the gain determined k, evaluate the phase margin of the uncompensated system
- Determine the necessary phase lead angle f, to be added to the system
- Determine the attenuation factor a using the equation

$$sin(\phi_m) = \frac{1-\alpha}{1+\alpha}$$

- Determine the frequency where the magnitude of the uncompensated system is equal to -20 log(1/sqrt(a)). Select the frequency as the new gain crossover frequency. This frequency corresponds to w_m and the maximum phase shift f_m occurs at this frequency.
- Determine the corner frequency of the lead network from

$$\omega = \frac{1}{T}, \quad \omega = \frac{1}{\alpha T}$$

- Finally insert an amplifier with gain equal to $\frac{1}{\alpha}$

85

Lag Compensator Design

The primary function of the lag compensator is to provide attenuation in the high frequency range to give a system sufficient phase margin. The phase-lag characteristics are of no consequence in lag compensation.

- Select the form of the compensator

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

- Define $k_c\beta = K \quad \Rightarrow G_c(s) = K\frac{Ts+1}{\beta Ts+1}$
- The open loop transfer function of the compensated system is:

$$G_c(s)G(s) = K\frac{Ts+1}{\beta Ts+1}G(s) = \frac{Ts+1}{\beta Ts+1}G_1(s)$$

where $G_1(s) = KG(s)$

- Determine K to satisfy the static velocity error constant
- If the system $G_1(s)$ does not satisfy the phase and gain margin requirements, ding the frequency point where the phase margin of the open-loop system is -180 + the required phase margin. Add 5-10 degrees to compensate for the phase-lag of the compensator

86

Lag Compensator Design (cont.)

- To prevent detrimental effects of the phase lag of the compensator, the pole and zero of the compensator must be located substantially lower than the new gain crossover frequency. Choose the corner frequency 1/T to be 1 octave or 1 decade below the new gain crossover frequency.
- Determine the attentuation necessary to bring the magnitude curve down to 0 db at the new gain crossover frequency. Since this attenuation is $-20\,log(\beta)$termine the value of β which leads to the other corner frequency of $1/\beta T$
- Using the value of K determined earlier, solve for k_c

$$k_c = \frac{K}{\beta}$$

87

Comparison of Lead and Lag Compensator

- Lead compensator achieves the desired result through the phase lead contribution of the compensator.

  Lag compensator achieves the desired results through the merits of its attenuation properties at high frequencies.
- Lead compensator yields a higher gain crossover frequency than is possible with lag compensation.

  The higher gain crossover frequency means larger bandwidth. If noise is present in the measurements, then a large bandwidth may not be desirable. It does result in faster response due to increased bandwidth.
- Lead compensator requires an additional increase in gain to offset the attentuation inherent in the lead network
- Lag compensation reduces the system gain at higher frequencies without reducing the system gain at lower frequencies.

  Since the system bandwidth is reduced, the system has a slower speed to respond. Also, high frequency noise can be attenuated.
- Lag compensation will introduce a pole-zero combination near the origin that will generate a long tail with small amplitude in the transient response.

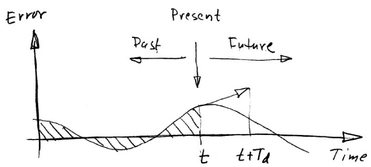
88

PID Controller

- The PID controller has the form:

$$u(t) = k_p e(t) + k_i \int e(\tau)\, d\tau + k_d \frac{de}{dt}$$

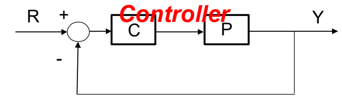
$$u(t) = k_p \left(e(t) + \frac{1}{T_r} \int e(\tau)\, d\tau + T_d \frac{de}{dt} \right)$$



89

PID Controller

- For the system



- The closed loop transfer function is:

$$\frac{Y}{R} = G_{YR} = \frac{CP}{1+CP}$$

- For a proportional controller, i.e., $u(t) = k_p e(t)$, the zero frequency gain with a proportional controller is:

$$G_{YR}(s=0) = \frac{C(0)P(0)}{1+C(0)P(0)}$$

- And the steady state error for a unit step is:

$$1 - G_{YR}(s=0) = \frac{1}{1+k_p P(0)}$$

- For example for a system: $P(s) = \frac{1}{(s+1)^3}$, with k_p = 1,2,5, the steady state error is:

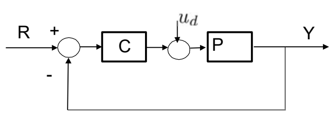
$$\frac{1}{1+k_p} = 0.5, 0.33, 0.17$$

Matlab simulations show that the error decreases, but becomes more oscillatory.

90

PID Controller

- To avoid steady state error, we design a controller: $u(t) = k_p e(t) + u_d$



for a unit step reference, the system response is:

$$Y(s) = \frac{CP}{1+CP} R(s) + \frac{P}{1+CP} u_d$$

- and the error is: $E(s) = R(s) - Y(s) = \frac{1}{s} - \frac{CP}{1+CP}\frac{1}{s} - \frac{P}{1+CP}\frac{u_d}{s}$

$$E(s) = R(s) - Y(s) = \frac{1}{s}\left(1 - \frac{CP}{1+CP} - \frac{P}{1+CP} u_d\right)$$

$$E(s) = R(s) - Y(s) = \frac{1}{s}\left(\frac{1}{1+CP} - \frac{P}{1+CP} u_d\right)$$

91

PID Controller

- The steady state error is:

$$e_{ss}(t \to \infty) = \lim_{s \to 0} E(s) = \lim_{s \to 0} s \frac{1}{s}\left(\frac{1}{1+CP} - \frac{P}{1+CP} u_d\right)$$

$$e_{ss}(t \to \infty) = \left(\frac{1}{1+C(0)P(0)} - \frac{P(0)}{1+C(0)P(0)} u_d\right)$$

If we select $u_d = \frac{1}{P(0)}$, we can force the steady state error to zero for the proportional controller. NOTE: requires exact knowledge of the plant, which is usually not available.

- PI Controller: $C(s) = k_p + \frac{k_i}{s}$

We see that the controller has infinite zero frequency gain, $C(s=0) = \infty$, therefore:

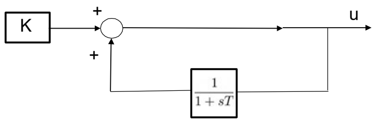
$$G_{YR} = \frac{C(0)P(0)}{1+C(0)P(0)} = 1$$

And the steady state error is 0.

92

PID Controller

- The integral part of the controller is implemented as:



$$\frac{U(s)}{E(s)} = K\frac{1}{1-\frac{1}{1+sT}} = K\frac{1+sT}{sT} = K + \frac{K}{sT}$$

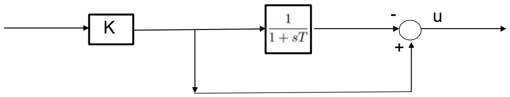
This implementation is called “automatic reset”.

- For K = 1, simulate system response for K_i = 0, 0.2, 0.5, 1.
- The steady state error is removed. For larger gain, it is faster, but the system becomes oscillatory.

93

Derivative Implementation

- The integral part of the controller is implemented as:



$$C(s) = K\left(1 - \frac{1}{1+sT}\right) = \frac{KsT}{1+sT}$$

94

Ziegler-Nichols (Z-N) Oscillation Method

Consider a PID controller parameterized as:

$$G_c(s) = K_p\left(1 + \frac{1}{T_r s} + \frac{T_d s}{\tau_D s + 1}\right)$$

Where T_r and T_d are known as the reset time and derivative time, respectively. The time constant τ_D is chosen as:

$$0.1T_d \leq \tau_D \leq 0.2T_d$$

$$\tau_D \neq 0$$

95

$\tau_D \neq 0$

The classical argument to choose was , apart from ensuing that the controller be proper, to attenuate high-frequency noise.

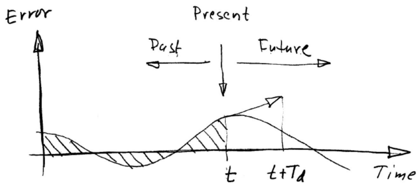
- The following procedure for selection the PID gains is only for open-loop stable plants.
  - Set the true plant under proportional control, with a very small gain
  - Increase the gain until the loop starts oscillating. Note that linear oscillation is required and that it should be detected at the controller output.
  - Record the controller critical gain $K_p = K_c$ and the oscillation period of the controller output P_c
  - Adjust the controller parameters according the table on the next viewgraph.

Control System Design, Goodwin, Graebe and Salgado, pp 162-165

96

Ziegler-Nichols (Z-N) Oscillation Method

- Consider a typical error evolution graph:



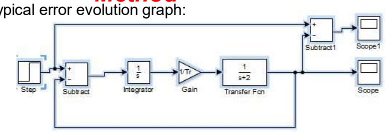
T_d, the derivative time is a measure of how far ahead into the future one can forecast the change in error based on the current slope of the error.

T_r, the reset time is approximately the time taken by the controller to overcome the steady state error.

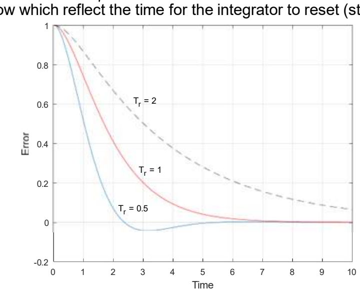
97

Ziegler-Nichols (Z-N) Oscillation Method

- Consider a typical error evolution graph:



Varying the reset time T_r results into different time for the error to reach zero as shown below which reflect the time for the integrator to reset (start from zero).



98

Ziegler-Nichols (Z-N) Oscillation Method

	K_p	T_r	T_d
P	$0.50K_c$		
PI	$0.45K_c$	$\frac{P_c}{1.2}$	
PID	$0.60K_c$	$0.5P_c$	$\frac{P_c}{8}$

99

Ziegler-Nichols (Z-N) Oscillation Method (EXAMPLE)

Consider a plant with a transfer function:

$$G(s) = \frac{1}{(s+1)^3}$$

The critical gain K_c and the critical frequency w_c can be determined by substituting s = jw in the characteristic equation and solving for K_c and w_c .

$$K_c + (j\omega_c + 1)^3 = 0$$

Equating the real and imaginary parts to zero, we have:

$$K_c - 3\omega_c^2 + 1 = 0$$
$$-\omega_c^3 + 3\omega = 0$$

Which results in: $\omega_c = \sqrt{3} \rightarrow P_c = 3.6276$

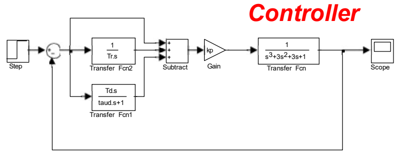
$$K_c = 8$$

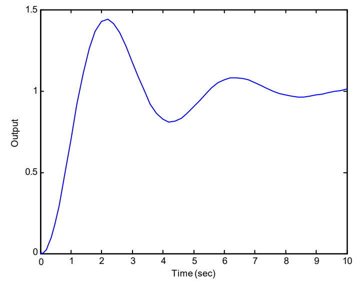
Thus, the PID gains are: $K_p = 0.6K_c = 4.8$

$$T_r = 0.5P_c = 1.81$$
$$T_d = 0.125P_c = 0.45$$

100

PID Controller





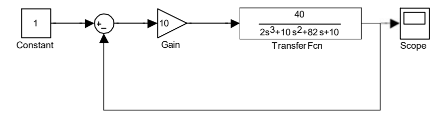
101

Ziegler-Nichols (Z-N) Oscillation Method (Cruise Control)

Consider a plant with a transfer function:

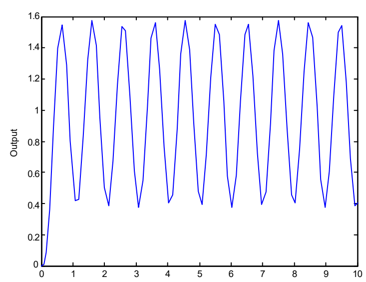
$$G(s) = \frac{40}{2s^3 + 10s^2 + 82s + 10}$$

Implement a proportional controller. Increase the gain till the system system response includes sustained oscillations. The critical gain K_c and the critical frequency w_c can be observed from the graph.



102

PID Controller



The critical gain Kc =10 and the critical frequency wc = 2pi rad/sec, or the period is 1 sec.

103

Integrator Antiwindup

- In any control system, the output of the actuator can saturate since the dynamic range of all actuators is limited.
- When the actuator saturates, the control signal stops changing and the feedback path is effectively broken.
- If the error signal continues to be applied to the integrator, the integrator output will grow (windup) until the sign of the error changes and the integration turns around.
- This can result in large overshoot, as the output must grow to produce the necessary unwinding error, resulting in poor transient response.

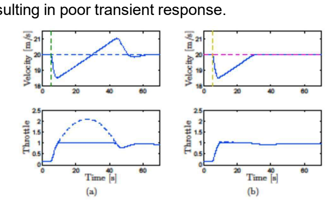
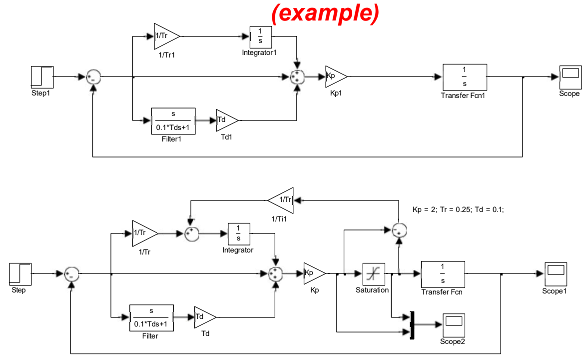


Figure 10.7: Simulation of windup (left) and anti-windup (right). The figure shows the speed v and the throttle u for a car that encounters a slope that is so steep that the throttle saturates. The controller output is dashed. The controller parameters are $k_p = 0.5$ and $k_i = 0.1$.

104

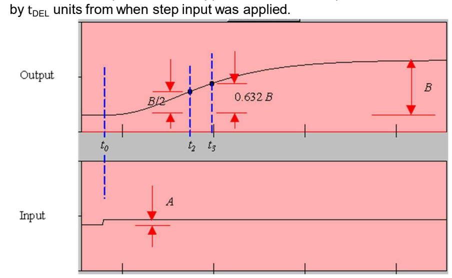
Integrator Antiwindup (example)



105

Cohen Coon Method

- The following procedure for selection the PID gains is only for open-loop stable plants.
  - Wait until process is at steady state
  - Introduce a step change in the input.
  - Based on the output, obtain an approximate first order process with a time constant τ delayed by t_{DEL} units from when step input was applied.



  - From the measurements, evaluate:

$$t_1 = \frac{(t_2 - log(2)t_3)}{(1 - log(2))}, \quad \tau = t_3 - t_1, \quad \tau_{DEL} = t_1 - t_0, \quad K = \frac{B}{A}$$

http://www.chem.mtu.edu/~tbco/cm416/cctune.html

106

Cohen Coon Tuning Method

$$G_c(s) = K_p\left(1 + \frac{1}{T_r s} + \frac{T_d s}{\tau_D s + 1}\right) \qquad r = \frac{\tau_{DEL}}{\tau}$$

	K_p	T_r	T_d
P	$\frac{1}{Kr}\left(1 + \frac{r}{3}\right)$		
PI	$\frac{1}{Kr}\left(0.9 + \frac{r}{12}\right)$	$\frac{30 + 3t}{9 + 20r}\tau_{DEL}$	
PID	$\frac{1}{Kr}\left(\frac{4}{3} + \frac{r}{4}\right)$	$\frac{32 + 6t}{13 + 8r}\tau_{DEL}$	$\frac{4}{11 + 2r}\tau_{DEL}$

107

Nyquist Stability Criterion

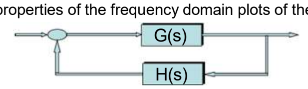
Thus far, stability has been analyzed by investigating the location of the roots of the characteristic equations by:

- Locus of the closed loop poles via the root-locus analysis
- Routh-Hurwitz criterion which provides information about the relative position of the closed loop poles with respect to the imaginary axis.

The Nyquist criterion is a frequency domain method which has the following desirable feature:

- It provides the same amount of information on the absolute stability as does the Routh-Hurwitz criterion
- The Nyquist criterion indicates the degree of stability of a stable system
- It gives information on the frequency-domain response of the system
- It can used to study system with time-delays
- It can be modified for use for nonlinear systems.

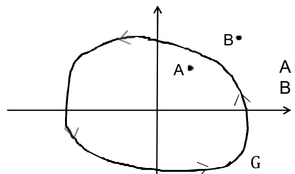
- The Nyquist criterion is a graphical method of determining the stabiltiy of a closed-loop system by investigating the properties of the frequency domain plots of the loop transfer function G(s)H(s)



108

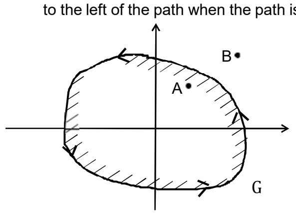
Nyquist Stability Criterion

- Encircled: A point is said to be encircled by a closed path if it is found inside the path.

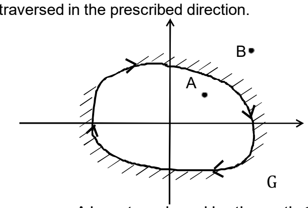


A is encircled by the path G
B is not encircled by the path G

- Enclosed: A point or a region is said to be enclosed by a closed path if it is found to lie to the left of the path when the path is traversed in the prescribed direction.



A is enclosed by the path G

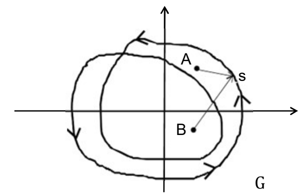


A is not enclosed by the path G
B Is enclosed by the path G

109

Nyquist Stability Criterion

- Number of Encircled and Enclosure: When a point is said to be encircled or enclosed by a closed path, a number N may be assigned to the number of encirclement or enclosure, as the case may be. The value N may be determined by drawing a vector from A to any arbitrary point s on the closed path G and let s follow the path in the prescribed direction until it returns to the starting point. The total net number of revolutions traversed by this vector is N.

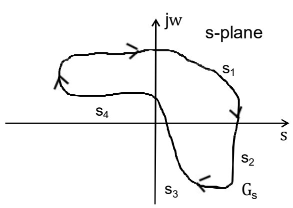


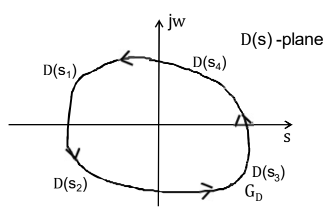
A is encircled once by the path G
B is encircled twice by the path G

110

Nyquist Stability Criterion

- Suppose that a continuous path G_s is arbitrarily chosen in the s-plane. If all points on G_s are in the specified region where D = 1+G(s)H(s) is analytic, then the curve G_D mapped by the function D into the D(s) plane is also a closed one.





Let D(s) be a single-valued rational function that is analytic in a given region in the s-plane except at a finite number of points. Suppose that an arbitrary closed path G_s is chosen in the s-plane so that D(s) is analytic at every point on G_s ; the corresponding D(s) locus mapped in the D(s)-plane will encircle the origin as many times as the difference between the number of zeros and the number of poles of D(s) that are encircled by the s-plane locus G_s .

111

Nyquist Stability Criterion

- In equation form, this statement can be expressed as:

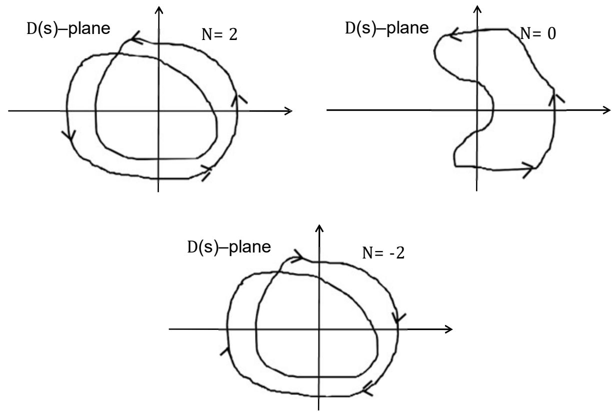
$$N = Z - P$$

where
- N = # of encirclement of the origin made by the D(s)–plane locus G_D
- Z = # of zeros of D(s) encircled by the s-plane locus G_s in the s-plane.
- P= # of poles of D(s) encircled by the s-plane locus G_s in the s-plane

- N can be positive, zero or negative
  1. N >0 (Z>P): If the s-plane locus encircles more zeros than poles of D(s)=1+G(s)H(s) in a certain prescribed direction (clockwise or counterclockwise), N is a positive integer. In this case the D(s)–plane locus will encircle the origin of the D(s)–plane N times in the same direction as that of G_s
  2. N =0 (Z=P): If the s-plane locus encircles as many zeros as poles, or no poles and zeros of D(s)=1+G(s)H(s), the D(s)–plane locus G_D will not encircle the origin of the D(s)–plane .
  3. N <0 (Z<P): If the s-plane locus encircles more poles than zeros of D(s)=1+G(s)H(s) in a certain prescribed direction (clockwise or counterclockwise), N is a negative integer. In this case the D(s)–plane locus will encircle the origin of the D(s)–plane N times in the opposite direction as that of G_s

112

Nyquist Stability Criterion



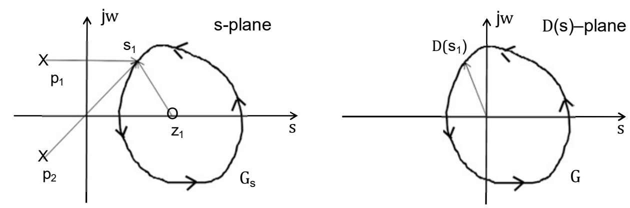
113

Nyquist Stability Criterion

- Consider the system with a positive K:

$$\Delta(s) = \frac{K(s+z_1)}{(s+p_1)(s+p_2)}$$

The poles and zeros are assumed to be located as shown below:



The function D(s) can be written as:

$$\Delta(s) = \frac{K|s+z_1|}{|s+p_1||s+p_2|}\left(\angle s+z_1 - \angle s+p_1 - \angle s+p_2\right)$$

114

Nyquist Stability Criterion

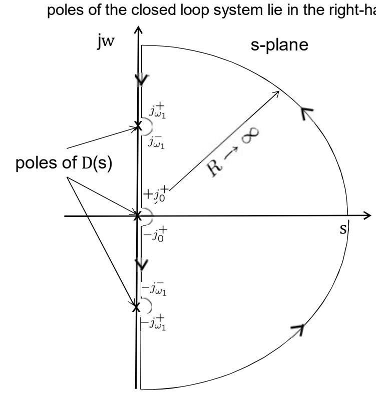
- The factors $(s+z_1)(s+p_1)(s+p_2)$ can be represented by the arrows.
- Now if the points s_1 moves along the locus of G_s in the prescribed counterclockwise direction until it returns to the starting point, the angle generated by the vectors drawn from the poles that are not encircled by G_s when s_1 completes one round trip are zero, whereas the vector drawn from z_1 which is encircled by G_s, generates a positive angle (counterclockwise sense) of 360 degrees. This is why only the poles and zeros of D(s) which are inside the G_s path in the s-plane, would contribute to the value of N.
- Since poles of D(s) correspond to negative phase angle and zeros correspond to positive phase angle, the value of N depends only on the difference between Z and P.
- In the present case Z = 1, P = 0
- Thus N = Z-P = 1

  Which means that the D(s)-plane locus should encircle the origin one in the same direction as the s-plane locus.

115

Nyquist Path

- If the G_s path in the s-plane is taken to be one that encircles the entire right-half of the s-plane , the mapping of the G_s path in the D(s)-plane can be used to determine if any poles of the closed loop system lie in the right-half of the complex plane.



The Nyquist path must not pass through any singularity of D(s), the small semi-circles shown along the y-axis, are used to indicate that the path should go around these singular points. It is apparent that if any pole or zero of D(s) lies inside the right half of the s-plane, it will be encircled by this Nyquist path.

116

Nyquist Stability Criterion

- Once the Nyquist path is specified, the stability of the closed-loop system can be determined by plotting the D = 1+G(s)H(s) locus when s takes on values along the Nyquist path and investigating the behavior of D(s) plot with respect to the origin in the D(s)-plane.
- However, since G(s)H(s) are generally known functions, it is simpler to construct the Nyquist plot of G(s)H(s), and the same conclusion of stability of the closed-loop system can be determined from the plot of G(s)H(s) with respect to the (-1,0jw) point in the G(s)H(s)—plane. This because the origin of the D(s) corresponds to the (-1,0jw) point on the G(s)H(s)—plane.
- Closed loop stability implies that D = 1+G(s)H(s) has zeros only in the left half of the s-plane. Open-loop stability implies that G(s)H(s) has poles only in the left-half of the s-plane.

- N_0=# of encirclements of the origin made by G(s)H(s)
- Z_0=#of zeros of G(s)H(s) in the right-half of the s-plane
- P_0=#of poles of G(s)H(s) in the right-half of the s-plane
- N_{-1}=# of encirclements of (-1,0jw) made by G(s)H(s)
- Z_{-1}=#of zeros of 1+G(s)H(s) in the right-half of the s-plane
- P_{-1}=#of poles of 1+G(s)H(s) in the right-half of the s-plane

117

Nyquist Stability Criterion

- Note: $P_{0=}$ $P_{-1,}$ since G(s)H(s) and (1+G(s)H(s)) always have the same poles.
- Closed loop stability implies that Z_{-1} = 0.
- Open-loop stability requires that P_0 = 0.
- The procedure of determining stability using the Nyquist plot is:
  - Define the Nyquist path according to the pole-zero properties of G(s)H(s)
  - The Nyquist plot of G(s)H(s) is constructed
  - The values of N_0 and N_{-1} are determined by observing the behavior of the Nyquist plot of G(s)H(s) with respect to the origin and the (-1,0jw) point
  - Once N_0 and N_{-1} are determined, the value of P_0 is determined from N_0= Z_0-P_0.
  - Once P_0 is determined, : $P_{0=}$ $P_{-1,}$ and Z_{-1} is determined from

$$N_{-1} = Z_{-1} - P_{-1}$$

since it has been established that for a stable closed-loop system Z_{-1} must be zero, gives

$$N_{-1} = -P_{-1}$$

Therefore, the Nyquist criterion may be formally stated as: For a closed-loop system to be stable, the Nyquist plot of 1+G(s)H(s) must encircle the (-1,0jw) point as many times as the number of poles of 1+G(s)H(s) that are in the right half of the s-plane, and the encirclements if any, must be made in the clockwise direction.

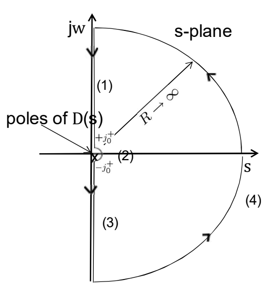
118

Nyquist Stability Criterion (Example)

- Consider a system with a transfer function (K and a are positive numbers):

$$G(s)H(s) = \frac{K}{s(s+a)}$$

- It is clear that G(s)H(s) does not have any pole in the right-half of the s-plane, thus $P_{0=}$ $P_{-1,}$ =0. The Nyquist path for the system is show below.



Since G(s)H(s) has a pole at the origin, it is necessary for the Nyquist path include a small semicircle around s=0.The entire Nyquist path is divided into four sections. section (2) of the Nyquist path can be represented by a phasor as:

$$s = \epsilon e^{j\theta}$$

Where $\epsilon \longrightarrow 0$ and q denote the magnitude and phase of the phasor, respectively. as the Nyquist path is traversed from $+j_0^+$ To $-j_0^+$, along section (2), the phasor rotates through 180 degrees starting at +90 and reaching -90 degrees.

119

Nyquist Stability Criterion (Example)

- The Nyquist plot G(s)H(s) can be determined as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} = \frac{K}{\epsilon e^{j\theta}(\epsilon e^{j\theta}+a)}$$

- Which can be approximated as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} \approx \frac{K}{a\epsilon e^{j\theta}} = \infty e^{-j\theta}$$

- Which indicates that all points on the Nyquist plot G(s)H(s) that correspond to section (2) on the Nyquist path have an infinite magnitude and the corresponding phases is opposite that of the s-plane locus. Since the phase varies from +90 to -90 degrees in the clockwise direction, the G(s)H(s) plot should have a phase that varies from -90 to +90 degrees in the counter-clockwise direction.
- The same technique can be used to determine the behavior of the G(s)H(s) plot, which corresponds to the semi-circle with infinite radius on the Nyquist path (section(4)). The points on this section can be represented by the phasor:

$$s = Re^{j\theta} \text{ where } R \to \infty$$

- Thus G(s)H(s) can be represented as:

$$G(s)H(s)|_{s=Re^{j\theta}} = \frac{K}{R^2 e^{2j\theta}} = 0e^{-2j\theta}$$

- Which corresponds to a phasor of infinitesimally small magnitude which rotates around the origin 2x180 degrees in the clockwise direction.

120

Nyquist Stability Criterion (Example)

- The Nyquist plot G(s)H(s) for section (1) and (3) can now be determined. For section (3) substituting s = jw

$$G(s)H(s)|_{s=j\omega} = \frac{K}{j\omega(j\omega + a)}$$

- Which can be rewritten as:

$$G(s)H(s)|_{s=j\omega} = \frac{K(-\omega^2 - ja\omega)}{\omega^4 + a^2\omega^2}$$

- The intersect of G(s)H(s) with the real axis is determined by equating the imaginary part of G(s)H(s) to zero, which leads to:

$$IM\left(G(j\omega)H(j\omega)\right) = \frac{-Ka\omega}{\omega^4 + a^2\omega^2} = 0$$

- Which gives $\omega = \infty$. This means that the only intersect on the real-axis in the G(s)H(s) plane is at the origin with $\omega = \infty$

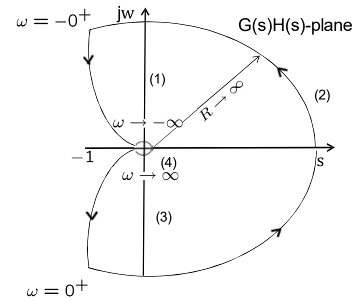
121

Nyquist Stability Criterion (Example)

- Since $Z_0 = P_0 = 0$ from G(s)H(s) and from the figure below $N_0 = N_{-1} = 0$, we have

$$Z_{-1} = N_{-1} + P_{-1} = 0$$

- since $P_0 = P_{-1}$. Therefore, the close-loop system is stable.



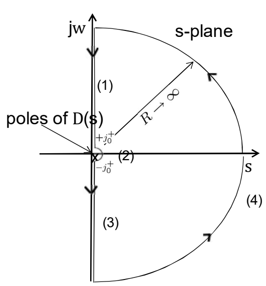
122

Nyquist Stability Criterion (Example 2)

- Consider a system with a transfer function (K and a are positive numbers):

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

- It is clear that G(s)H(s) does not have any pole in the right-half of the s-plane, thus $P_0 = P_{-1} = 0$. The Nyquist path for the system is show below.



Since G(s)H(s) has a pole at the origin, it is necessary for the Nyquist path include a small semicircle around s=0.The entire Nyquist path is divided into four sections. section (2) of the Nyquist path can be represented by a phasor as:

$$s = \epsilon e^{j\theta}$$

Where $\epsilon \to 0$ and q denote the magnitude and phase of the phasor, respectively. as the Nyquist path is traversed from $+j_0^+$ To $-j_0^+$, along section (2), the phasor rotates through 180 degrees starting at +90 and reaching -90 degrees.

123

Nyquist Stability Criterion (Example 2)

- The Nyquist plot G(s)H(s) can be determined as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} = \frac{K(\epsilon e^{j\theta} - 1)}{\epsilon e^{j\theta}(\epsilon e^{j\theta} + 1)}$$

- Which can be approximated as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} \approx \frac{-K}{\epsilon e^{j\theta}} = \infty e^{-j(\theta+\pi)}$$

- Which indicates that all points on the Nyquist plot G(s)H(s) that correspond to section (2) on the Nyquist path have an infinite magnitude and the corresponding phases starts at an angle of +90 and ends at -90 degrees and goes around the origin in the counter-clockwise direction.
- The same technique can be used to determine the behavior of the G(s)H(s) plot, which corresponds to the semi-circle with infinite radius on the Nyquist path (section(4)). The points on this section can be represented by the phasor:

$$s = Re^{j\theta} \text{ where } R \to \infty$$

- Thus G(s)H(s) can be represented as:

$$G(s)H(s)|_{s=Re^{j\theta}} = \frac{K}{Re^{j\theta}} = 0e^{-j\theta}$$

- Which corresponds to a phasor of infinitesimally small magnitude which rotates around the origin 180 degrees in the clockwise direction.

124

Nyquist Stability Criterion (Example)

- The Nyquist plot G(s)H(s) for section (1) and (3) can now be determined. For section (3) substituting s = jw

$$G(s)H(s)|_{s=j\omega} = \frac{K(j\omega - 1)}{j\omega(j\omega + 1)}$$

- Which can be rewritten as:

$$G(s)H(s)|_{s=j\omega} = \frac{K(2\omega + j(1-\omega^2))}{\omega^3 + \omega}$$

- The intersect of G(s)H(s) with the real axis is determined by equating the imaginary part of G(s)H(s) to zero, which leads to:

$$IM\,(G(j\omega)H(j\omega)) = \frac{K(1-\omega^2)}{\omega^3 + \omega} = 0$$

- Which gives $\omega = \pm 1$, which are the frequencies at which the G(s)H(s) curve crosses the real axis.

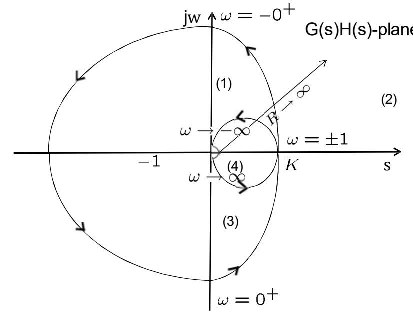
125

Nyquist Stability Criterion (Example)

- Since Z_0= 1 from G(s)H(s) and from the figure below N_{-1}=1, we have

$$Z_{-1} = N_{-1} + P_{-1} = 1$$

- since P_0= P_{-1}.Therefore, the close-loop system is unstable.

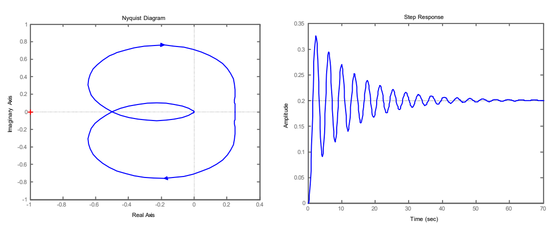


126

Relative Stability—Gain Margin, Phase Margin

- To demostrate the concept of relative stability, the Nyquist plot and the step response of a typical third-order system is shown below.

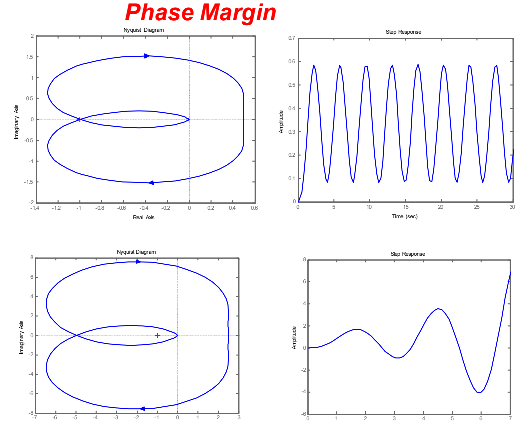
$$G(s)H(s) = \frac{K}{s^3 + 2s^3 + 3s + 4}$$



$K = 1$

127

Relative Stability—Gain Margin, Phase Margin



$K = 2$

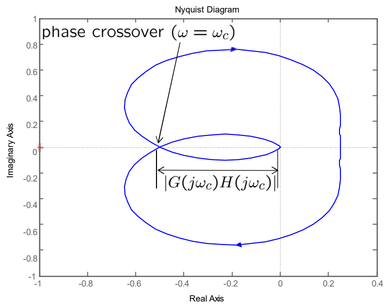
$K = 10$

128

Relative Stability—Gain Margin, Phase Margin

- Note that as K is increased, the point at which the Nyquist plot crosses the real axis (phase crossover point) moves closer to the (-1,0jw) point and once it crosses it, the system is unstable.
- GAIN MARGIN: The gain margin is a measure of the closeness of the phase-crossover point to the (-1,0jw) point. The gain margin of the closed-loop system that has G(s)H(s) as its loop transfer function is:

$$\text{GM} = 20\log_{10}\left(\frac{1}{|G(j\omega_c)H(j\omega_c)|}\right)$$



129

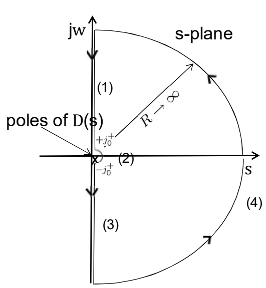
Relative Stability—Gain Margin, Phase Margin

- GAIN MARGIN: is the amount of gain in decibels that can be allowed to increase in the loop before the closed-loop system reaches instability.
- When the G(s)H(s) plot goes through the (-1,0jw) point, the gain margin is 0 db, which implies that the loop gain can no longer be increased as the system is already on the margin of stability.
- When the G(s)H(s) plot does not intersect the negative real axis at any finite nonzero frequency, and the Nyquist stability criterion indicates that the (-1,0jw) point must not be enclosed for system stability, the gain margin is infinite in decibels, which means that theoretically, the value of the loop gain can be increased to infinity before any instability occurs.
- When the (-1,0jw) point is to the right of the phase-crossover point, the magnitude of G(s)H(s) is greater than unity and the gain margin is negative in decibels, which implies that the system is unstable.
- NOTE: If G(s)H(s) has poles or zeros in the right-half of the s-plane, (-1,0jw) point must be encircled by the G(s)H(s) plot for stability. Under this condition, a stable system yields a negative gain margin.

130

Nyquist Example

- Consider the system: G(s)H(s) = 5(s+3)/(s(s-1))



Since G(s)H(s) has a pole at the origin, it is necessary for the Nyquist path include a small semicircle around s=0.The entire Nyquist path is divided into four sections. section (2) of the Nyquist path can be represented by a phasor as:

$$s = \epsilon e^{j\theta}$$

Where $\epsilon \longrightarrow 0$ and q denote the magnitude and phase of the phasor, respectively. as the Nyquist path is traversed from $+j_0^+$ To $-j_0^+$ along section (2), the phasor rotates through 180 degrees starting at +90 and reaching -90 degrees.

131

Nyquist Stability Criterion (Example 2)

- The Nyquist plot G(s)H(s) can be determined as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} = \frac{5(\epsilon e^{j\theta}+3)}{\epsilon e^{j\theta}(\epsilon e^{j\theta}-1)}$$

- Which can be approximated as:

$$G(s)H(s)|_{s=\epsilon e^{j\theta}} \approx \frac{-15}{\epsilon e^{j\theta}} = \infty e^{-j(\theta+\pi)}$$

- Which indicates that all points on the Nyquist plot G(s)H(s) that correspond to section (2) on the Nyquist path have an infinite magnitude and the corresponding phases starts at an angle of +90 and ends at -90 degrees and goes around the origin in the counter-clockwise direction.
- The same technique can be used to determine the behavior of the G(s)H(s) plot, which corresponds to the semi-circle with infinite radius on the Nyquist path (section(4)). The points on this section can be represented by the phasor:

$$s = Re^{j\theta} \text{ where } R \to \infty$$

- Thus G(s)H(s) can be represented as:

$$G(s)H(s)|_{s=Re^{j\theta}} = \frac{K}{Re^{j\theta}} = 0e^{-j\theta}$$

- Which corresponds to a phasor of infinitesimally small magnitude which rotates around the origin 180 degrees in the clockwise direction starting at +90 degrees.

132

Nyquist Stability Criterion (Example)

- The Nyquist plot G(s)H(s) for section (1) and (3) can now be determined. For section (3) substituting s = jw

$$G(s)H(s)|_{s=j\omega} = \frac{5(j\omega+3)}{j\omega(j\omega-1)}$$

- Which can be rewritten as:

$$G(s)H(s)|_{s=j\omega} = -\frac{5(4\omega^2 + j(\omega^3 - 3\omega))}{\omega^4+\omega^2}$$

- The intersect of G(s)H(s) with the real axis is determined by equating the imaginary part of G(s)H(s) to zero, which leads to:

$$IM\left(G(j\omega)H(j\omega)\right) = \frac{5(3-\omega^2)}{\omega^3+\omega} = 0$$

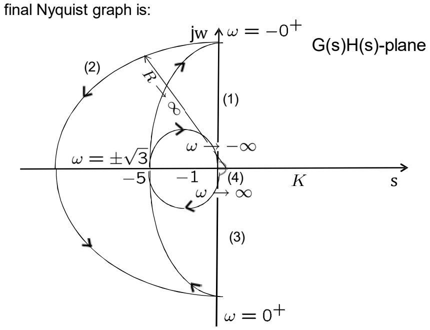
- Which gives $\omega = \pm\sqrt{3}$ which are the frequencies at which the G(s)H(s) curve crosses the real axis.
- The corresponding real part is:

$$IM\left(G(j\omega)H(j\omega)\right) = \frac{20\omega}{\omega^3+\omega} = -5$$

133

Nyquist Stability Criterion (Example)

- The final Nyquist graph is:



- Which shows that the (-1,j0) point is encircled in a clockwise manner: N_{-1} = -1
- P_{-1} = 1, which corresponds to a right half plane zero. Therefore

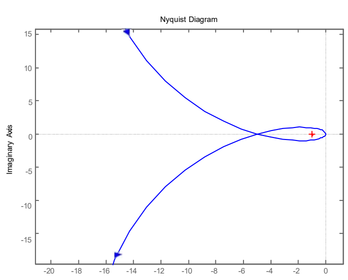
$$Z_{-1} = N_{-1} + P_{-1} = 0$$

- Which indicates that the closed loop system is stable.

134

Relative Stability—Gain Margin, Phase Margin

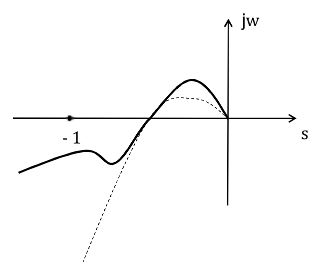
- Consider the system: G(s) = 5(s+3)/(s(s-1))
- Gain Margin = -14 db (Note system is stable, even though the gain margin is negative)



135

Relative Stability—Gain Margin, Phase Margin

- Gain margin is one way of representing the relative stability of a feedback control system. In general a system has a large gain margin, should be relatively more stable than one with a smaller gain margin.
- The gain margin of the two system shown below by the solid and the dashed lines, have the same gain margin. However, the system represented by the dashed line is more stable than the system represented by the solid line. This is due to the fact that with any change in the system parameter, other than the loop gain, it is easier for the system represented by the solid line to pass through the (-1,0jw) point.

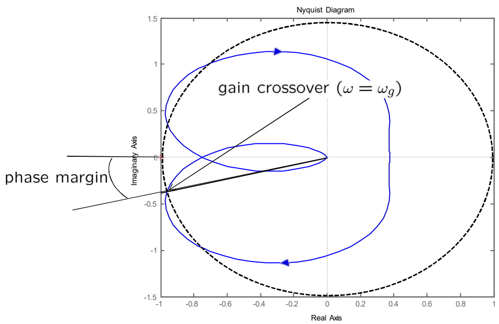


136

Relative Stability—Gain Margin, Phase Margin

- PHASE MARGIN: is defined as the angle in degrees through which G(s)H(s) plot must be rotated about the origin in order that the gain-crossover point on the locus passes through the (-1,0jw) point .
- The phase margin indicates the effect on stability of changes in system parameters which alter the phase of G(s)H(s).

$$\text{PM} = \angle G(j\omega_g)H(j\omega_g) - 180$$



137