The Nyquist Stability Criterion

Control Engineering (Reguleringsteknik)

Christoffer Sloth

`chsl@mmmi.sdu.dk`

SDU Robotics
The Maersk Mc-Kinney Moller Institute
University of Southern Denmark

SDU

Agenda

Introduction

Curriculum for Reguleringsteknik (REG)

Matematiske og grafiske metoder til syntese af lineære tidsinvariante systemer:[1]

- diskret og kontinuert tilstandsbeskrivelse
- analyse i tid og frekvens
- stabilitet, reguleringshastighed, følsomhed og fejl
- digitale PI, PID, LEAD og LAG regulatorer (serieregulatorer)
- tilstandsregulering, pole-placement og tilstands-estimering (observer)
- optimal regulering (least squares) og optimal tilstands-estimation (Kalman-filter)

Færdigheder:
Efter gennemførelse af kurset kan den succesfulde studerende:

- kunne analysere, dimensionere og implementere såvel kontinuert som tidsdiskret regulering af lineære tidsinvariante og stokastiske systemer

Kompetencer:
Efter gennemførelse af kurset kan den succesfulde studerende:

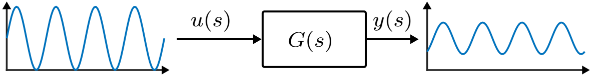
- anvende og implementere klassiske og moderne reguleringsteknikker for at kunne styre og regulere en robot hurtig og præcist

[1] Based on https://fagbesk.sam.sdu.dk/?fag_id=39673

Frequency Response

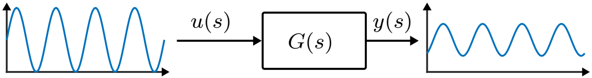
Frequency Response
Definition

The ***frequency response*** of a system is the steady-state response of a system to a sinusoidal input.



Frequency Response
Definition

The ***frequency response*** of a system is the steady-state response of a system to a sinusoidal input.



The output of a time-invariant system will have the same frequency as the sinusoidal input, but possibly with a different amplitude and phase.

Recall that the output of a time-invariant system is given by

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

where y is the output, u is the input, and h is the impulse response of the system.

Recall that the output of a time-invariant system is given by

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

where y is the output, u is the input, and h is the impulse response of the system.

To obtain the frequency response, only sinusoidal inputs are considered, i.e.,

$$u(t) = A\cos(\omega t)$$

Recall that the output of a time-invariant system is given by

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

where y is the output, u is the input, and h is the impulse response of the system.

To obtain the frequency response, only sinusoidal inputs are considered, i.e.,

$$u(t) = A\cos(\omega t)$$

By Euler's formula ($e^{ia} = \cos a + i \sin a$) the output can be written as

$$u(t) = \frac{A}{2}\left(e^{j\omega t} + e^{-j\omega t}\right)$$

Motivated by the input given as exponentials, the system response to an input $u(t) = e^{st}$ with $s = \sigma + j\omega$ is computed below

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

Motivated by the input given as exponentials, the system response to an input $u(t) = e^{st}$ with $s = \sigma + j\omega$ is computed below

$$
\begin{aligned}
y(t) &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau \\
&= \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau
\end{aligned}
$$

Motivated by the input given as exponentials, the system response to an input $u(t) = e^{st}$ with $s = \sigma + j\omega$ is computed below

$$
\begin{aligned}
y(t) &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau \\
&= \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau \\
&= \int_{-\infty}^{\infty} h(\tau)e^{st}e^{-s\tau}d\tau
\end{aligned}
$$

Motivated by the input given as exponentials, the system response to an input $u(t) = e^{st}$ with $s = \sigma + j\omega$ is computed below

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{st}e^{-s\tau}d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau e^{st} \end{aligned}$$

Motivated by the input given as exponentials, the system response to an input $u(t) = e^{st}$ with $s = \sigma + j\omega$ is computed below

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{s(t-\tau)}d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{st}e^{-s\tau}d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{-s\tau}d\tau e^{st} \\ &= H(s)e^{st} \end{aligned}$$

The expression for $y(t)$ is then given as follows for $u(t) = A\cos(\omega t)$

$$y(t) = \frac{A}{2}\left(H(j\omega)e^{j\omega t} + H(-j\omega)e^{-j\omega t}\right)$$

The expression for $y(t)$ is then given as follows for $u(t) = A\cos(\omega t)$

$$y(t) = \frac{A}{2}\left(H(j\omega)e^{j\omega t} + H(-j\omega)e^{-j\omega t}\right)$$

The complex number $H(j\omega)$ is expressed in polar form as

$$H(j\omega) = M(\omega)e^{j\varphi(\omega)}$$

where $M(\omega)$ is the magnitude of $H(j\omega)$ and $\varphi(\omega)$ is the phase of $H(j\omega)$.

The expression for $y(t)$ is then given as follows for $u(t) = A\cos(\omega t)$

$$y(t) = \frac{A}{2}\left(H(j\omega)e^{j\omega t} + H(-j\omega)e^{-j\omega t}\right)$$

The complex number $H(j\omega)$ is expressed in polar form as

$$H(j\omega) = M(\omega)e^{j\varphi(\omega)}$$

where $M(\omega)$ is the magnitude of $H(j\omega)$ and $\varphi(\omega)$ is the phase of $H(j\omega)$.
This gives the following expression for the output

$$y(t) = \frac{A}{2}M(\omega)\left(e^{j(\varphi(\omega)+\omega t)} + e^{-j(\varphi(\omega)+\omega t)}\right)$$

The expression for $y(t)$ is then given as follows for $u(t) = A\cos(\omega t)$

$$y(t) = \frac{A}{2}\left(H(j\omega)e^{j\omega t} + H(-j\omega)e^{-j\omega t}\right)$$

The complex number $H(j\omega)$ is expressed in polar form as

$$H(j\omega) = M(\omega)e^{j\varphi(\omega)}$$

where $M(\omega)$ is the magnitude of $H(j\omega)$ and $\varphi(\omega)$ is the phase of $H(j\omega)$.
This gives the following expression for the output

$$\begin{aligned} y(t) &= \frac{A}{2}M(\omega)\left(e^{j(\varphi(\omega)+\omega t)} + e^{-j(\varphi(\omega)+\omega t)}\right) \\ &= \frac{A}{2}M(\omega)\cos(\omega t + \varphi(\omega)) \end{aligned}$$

In conclusion, the system response subject to the input $u(t) = A\cos(\omega t)$ is

$$y(t) = \frac{A}{2} M(\omega) \cos(\omega t + \varphi(\omega))$$

where

$$M(\omega) = |H(j\omega)| \text{ and } \varphi(\omega) = \angle H(j\omega)$$

In conclusion, the system response subject to the input $u(t) = A\cos(\omega t)$ is

$$y(t) = \frac{A}{2} M(\omega) \cos(\omega t + \varphi(\omega))$$

where

$$M(\omega) = |H(j\omega)| \text{ and } \varphi(\omega) = \angle H(j\omega)$$

This means that the output of the system is also a sinusoidal signal with the same frequency as the input, but possibly with an other amplitude and a phase shift.

Consider the first-order differential equation

$$\dot{y} + ky = u$$

Consider the first-order differential equation

$$\dot{y} + ky = u$$

The Laplace transform of the differential equation leads to

$$H(s) = \frac{y(s)}{u(s)} = \frac{1}{s+k}$$

Consider the first-order differential equation

$$\dot{y} + ky = u$$

The Laplace transform of the differential equation leads to

$$H(s) = \frac{y(s)}{u(s)} = \frac{1}{s+k}$$

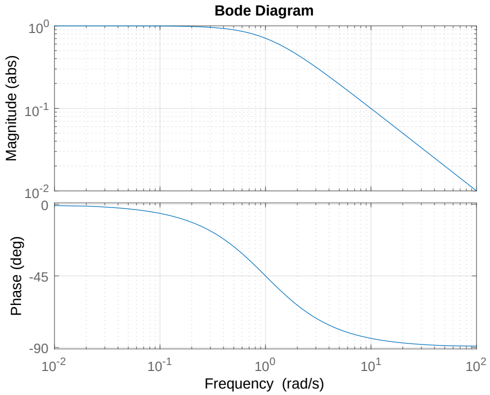
The ***frequency response*** of $H(s)$ is given by the magnitude and phase of $H(j\omega) = \frac{1}{j\omega + k}$

$$M = |H(j\omega)| = \frac{1}{\sqrt{\omega^2 + k^2}} \qquad \text{and} \qquad \varphi = -\tan^{-1}\left(\frac{\omega}{k}\right)$$

where M is the ***amplitude ratio*** and φ is the ***phase***.

Frequency Response

Example



Frequency Response

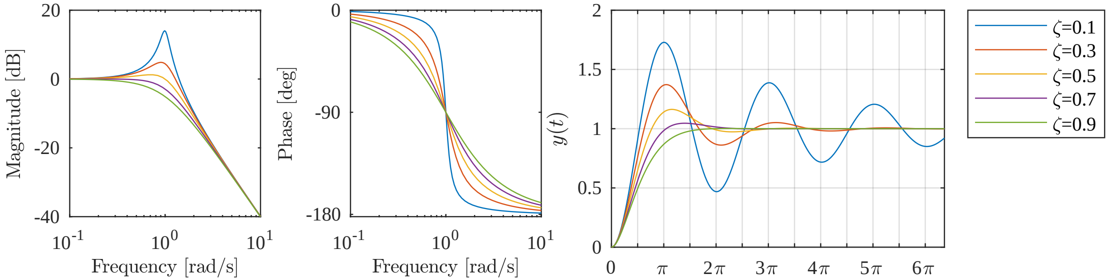
Second-Order System (1)

The transfer function of a second-order system can be written as

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

The transfer function of a second-order system can be written as

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



The ***bandwidth*** of a closed-loop system $T(s)$ is defined to be the maximum frequency at which the output y of a system will track a sinusoidal input r in a satisfactory manner (output attenuated to $1/\sqrt{2}$ times the input).
Formally, the bandwidth ω_{BW} of $T(s)$ is the maximal frequency such that

$$T(j\omega) \geq \frac{1}{\sqrt{2}} \qquad \forall \omega \in [0, \omega_{BW}]$$

Frequency Response

Terminology

The ***bandwidth*** of a closed-loop system $T(s)$ is defined to be the maximum frequency at which the output y of a system will track a sinusoidal input r in a satisfactory manner (output attenuated to $1/\sqrt{2}$ times the input).
Formally, the bandwidth ω_{BW} of $T(s)$ is the maximal frequency such that

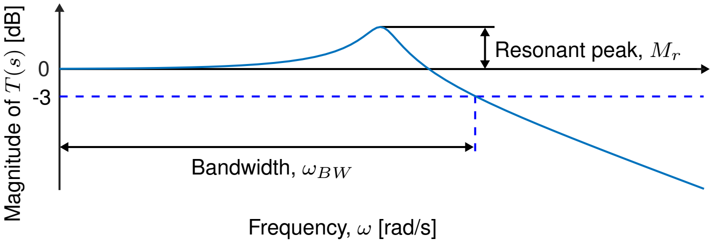
$$T(j\omega) \geq \frac{1}{\sqrt{2}} \qquad \forall \omega \in [0, \omega_{BW}]$$

The maximal value of the frequency response magnitude is called the ***resonant peak*** M_r.

The ***bandwidth*** of a closed-loop system $T(s)$ is defined to be the maximum frequency at which the output y of a system will track a sinusoidal input r in a satisfactory manner (output attenuated to $1/\sqrt{2}$ times the input).
Formally, the bandwidth ω_{BW} of $T(s)$ is the maximal frequency such that

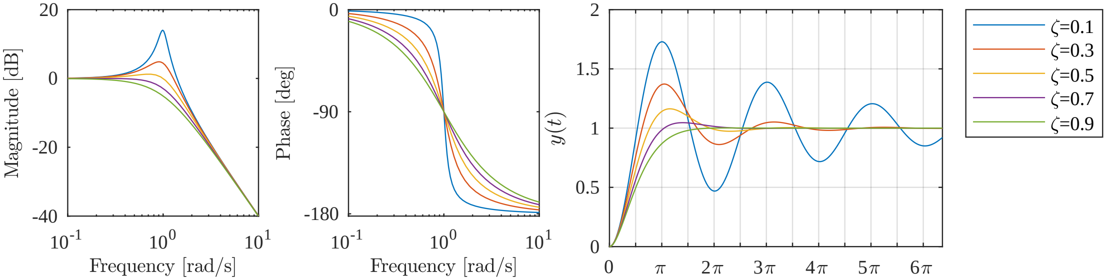
$$T(j\omega) \geq \frac{1}{\sqrt{2}} \qquad \forall \omega \in [0, \omega_{BW}]$$

The maximal value of the frequency response magnitude is called the ***resonant peak*** M_r.



For a second-order system

- **Bandwidth**: The bandwidth $\omega_{BW} \approx \omega_n$ (the equality holds for $\zeta = 0.7$).
- **Resonant peak**: The resonant peak is $M_p = \frac{1}{2\zeta}$ for $\zeta < 0.5$.



A transfer function $G(s)$ evaluated at $s = j\omega$ can be be factored in terms of poles and zeros, and written as

$$G(j\omega) = \frac{r_1 e^{j\theta_1} r_2 e^{j\theta_2}}{r_3 e^{j\theta_3} r_4 e^{j\theta_4} r_5 e^{j\theta_5}} = \frac{r_1 r_2}{r_3 r_4 r_5} e^{j(\theta_1 + \theta_2 - \theta_3 - \theta_4 - \theta_5)}$$

A transfer function $G(s)$ evaluated at $s = j\omega$ can be be factored in terms of poles and zeros, and written as

$$G(j\omega) = \frac{r_1 e^{j\theta_1} r_2 e^{j\theta_2}}{r_3 e^{j\theta_3} r_4 e^{j\theta_4} r_5 e^{j\theta_5}} = \frac{r_1 r_2}{r_3 r_4 r_5} e^{j(\theta_1+\theta_2-\theta_3-\theta_4-\theta_5)}$$

The the magnitude of $G(j\omega)$ is

$$|G(j\omega)| = \frac{r_1 r_2}{r_3 r_4 r_5}$$

which leads to

$$\log_{10}|G(j\omega)| = \log_{10} r_1 + \log_{10} r_2 - \log_{10} r_3 - \log_{10} r_4 - \log_{10} r_5$$

Advantages of working with Bode plots

1. Dynamic compensator design can be based entirely on the Bode plots. (We will do this in the next lecture)
2. Bode plots can be determined experimentally.
3. The Bode plot of a system defined by a series connection is given by the addition of the Bode plots of the individual systems.
4. The use og log scale permits a much wider range of frequencies to be displayed on a single plot than is possible with linear scales.

Bode Plots

Advantages

Advantages of working with Bode plots

1. Dynamic compensator design can be based entirely on the Bode plots. (We will do this in the next lecture)
2. Bode plots can be determined experimentally.
3. The Bode plot of a system defined by a series connection is given by the addition of the Bode plots of the individual systems.
4. The use og log scale permits a much wider range of frequencies to be displayed on a single plot than is possible with linear scales.

Example: Consider the series connection of two systems

$$G_1(s) = \frac{1}{(s+1)^2} \qquad G_2(s) = s + 100$$

In the following, we consider the open-loop transfer function

$$KG(s) = K\frac{(s-z_1)(s-z_2)\cdots}{(s-p_1)(s-p_2)\cdots}$$

In the following, we consider the open-loop transfer function

$$KG(s) = K\frac{(s-z_1)(s-z_2)\cdots}{(s-p_1)(s-p_2)\cdots}$$

For convenience, the transfer function is evaluated at $s = j\omega$ and rewritten into *Bode form* as

$$KG(s) = K_0\frac{(j\omega\tau_1+1)(j\omega\tau_2+1)\cdots}{(j\omega\tau_a+1)(j\omega\tau_b+1)\cdots}$$

In the following, three classes of terms of the transfer function are analyzed

1. Class 1: $K_0(j\omega)^n$ for $n \in \mathbb{Z}$.
2. Class 2: $(j\omega\tau + 1)^{\pm 1}$.
3. Class 3: $\left((j\omega/\omega_n)^2 + 2\zeta(j\omega/\omega_n) + 1\right)^{\pm 1}$.

For the term $K_0(j\omega)^n$, we have

$$\log K_0|(j\omega)^n| = \log K_0 + n\log|j\omega|$$

For the term $K_0(j\omega)^n$, we have

$$\log K_0|(j\omega)^n| = \log K_0 + n\log|j\omega|$$

This means that the magnitude plot is a straight line with slope $n \cdot (20$ dB/dec$)$.

For the term $K_0(j\omega)^n$, we have

$$\log K_0|(j\omega)^n| = \log K_0 + n\log|j\omega|$$

This means that the magnitude plot is a straight line with slope $n \cdot (20$ dB/dec$)$.

The phase is constant and given by

$$\varphi = n \cdot 90^\circ$$

To sketch a Bode plot for $K_0(j\omega)^n$, we start at $\omega = 1$, then from

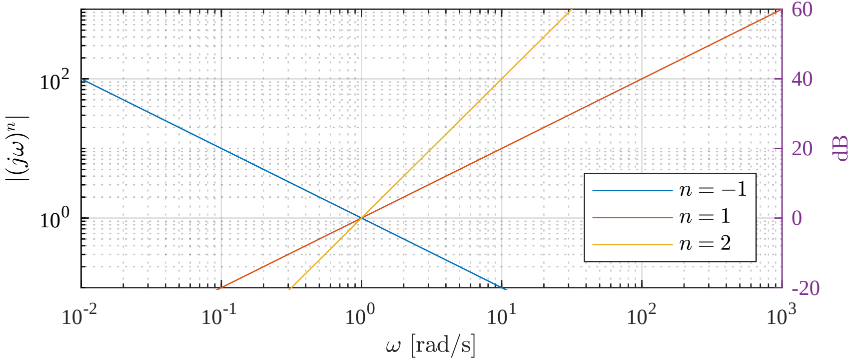
$$\log K_0|(j\omega)^n| = \log K_0 + n \log |j\omega|$$

it is seen that $\log K_0|(j)^n| = \log K_0$.

To sketch a Bode plot for $K_0(j\omega)^n$, we start at $\omega = 1$, then from

$$\log K_0|(j\omega)^n| = \log K_0 + n\log|j\omega|$$

it is seen that $\log K_0|(j)^n| = \log K_0$.



For the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

- For $\omega\tau \ll 1$, $j\omega\tau + 1 \approx 1$.
- For $\omega\tau \gg 1$, $j\omega\tau + 1 \approx j\omega\tau$.

For the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

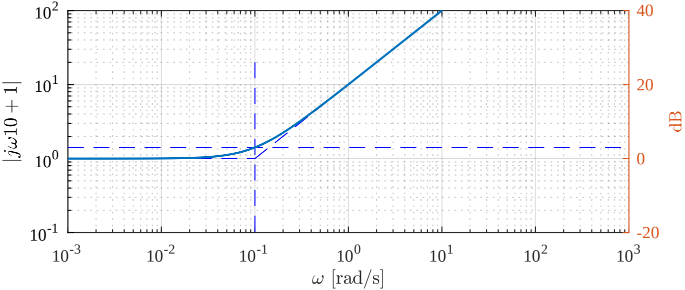
- For $\omega\tau \ll 1$, $j\omega\tau + 1 \approx 1$.
- For $\omega\tau \gg 1$, $j\omega\tau + 1 \approx j\omega\tau$.

In addition, for $\omega = 1/\tau$, the gain is $\sqrt{2}$ - an increase of 3 dB compared to the DC gain.
The point $\omega = 1/\tau$ is called the ***break point***.

For the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

- For $\omega\tau \ll 1$, $j\omega\tau + 1 \approx 1$.
- For $\omega\tau \gg 1$, $j\omega\tau + 1 \approx j\omega\tau$.

In addition, for $\omega = 1/\tau$, the gain is $\sqrt{2}$ - an increase of 3 dB compared to the DC gain.
The point $\omega = 1/\tau$ is called the ***break point***.



For the phase of the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

- For $\omega\tau \ll 1$, $\angle 1 = 0°$.
- For $\omega\tau \gg 1$, $\angle j\omega\tau = 90°$.

For the phase of the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

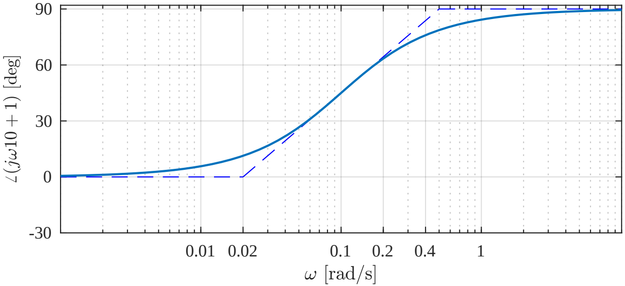
- For $\omega\tau \ll 1$, $\angle 1 = 0°$.
- For $\omega\tau \gg 1$, $\angle j\omega\tau = 90°$.

In addition, at the break point $\omega = 1/\tau$, the angle is $\angle(1 + j) = 45°$.

For the phase of the term $(j\omega\tau + 1)^{\pm 1}$, we study two asymptotes

- For $\omega\tau \ll 1$, $\angle 1 = 0°$.
- For $\omega\tau \gg 1$, $\angle j\omega\tau = 90°$.

In addition, at the break point $\omega = 1/\tau$, the angle is $\angle(1 + j) = 45°$.



For the term $\left((j\omega/\omega_n)^2 + 2\zeta(j\omega/\omega_n) + 1\right)^{\pm 1}$, we have

- For $\omega \ll \omega_n$, the amplitude is approximately 1.

For the term $\left((j\omega/\omega_n)^2 + 2\zeta(j\omega/\omega_n) + 1\right)^{\pm 1}$, we have

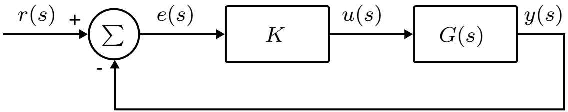
- For $\omega \ll \omega_n$, the amplitude is approximately 1.

In addition, at the break point $\omega = \omega_n$, the magnitude is $|G(j\omega)| = 1/(2\zeta)$ and the phase is $\pm 90°$.

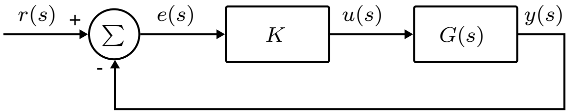
1. Rewrite the considered transfer function into Bode form.
2. Determine the value of the $K_0(j\omega)^n$ term. Plot the low frequency magnitude asymptote through the point K_0 at $\omega = 1$ and with slope of $n\times$20/db per decade.
3. Complete the composite magnitude asymptotes by extending the low-frequency asymptote until the first frequency break point. Then change the slope according to the behavior at the break point, and continue the procedure for the remaining break points.
4. Sketch the approximate magnitude curve by increasing the asymptote value by a factor $\sqrt{2}$ at first-order numerator break and decreasing it by a factor $1/\sqrt{2}$ at denominator break points.

1. Rewrite the considered transfer function into Bode form.
2. Determine the value of the $K_0(j\omega)^n$ term. Plot the low frequency magnitude asymptote through the point K_0 at $\omega = 1$ and with slope of $n\times$20/db per decade.
3. Complete the composite magnitude asymptotes by extending the low-frequency asymptote until the first frequency break point. Then change the slope according to the behavior at the break point, and continue the procedure for the remaining break points.
4. Sketch the approximate magnitude curve by increasing the asymptote value by a factor $\sqrt{2}$ at first-order numerator break and decreasing it by a factor $1/\sqrt{2}$ at denominator break points.
5. Plot the low-frequency asymptote of the phase $\phi = n \times 90°$.
6. Change the phase at the phase points, and correct the phase according to the slope at the phase point.

The stability of the closed-loop system shown below.



The stability of the closed-loop system shown below.



The stability of the system is determined from the frequency response of the *open-loop* transfer function $KG(s)$, which is usually known.

Recall the closed loop transfer function

$$T(s) = \frac{KG(s)}{1 + KG(s)}$$

The point that separates stable and unstable systems is associated to a closed-loop pole on the imaginary axis. At this point

$$|KG(s)| = 1 \quad \text{and} \quad \angle G(s) = 180^\circ$$

The point that separates stable and unstable systems is associated to a closed-loop pole on the imaginary axis. At this point

$$|KG(s)| = 1 \quad \text{and} \quad \angle G(s) = 180^\circ$$

At the point of ***neutral stability***, the root locus conditions hold for $s = j\omega$, i.e.,

$$|KG(j\omega)| = 1 \quad \text{and} \quad \angle G(j\omega) = 180^\circ$$

The point that separates stable and unstable systems is associated to a closed-loop pole on the imaginary axis. At this point

$$|KG(s)| = 1 \quad \text{and} \quad \angle G(s) = 180^\circ$$

At the point of ***neutral stability***, the root locus conditions hold for $s = j\omega$, i.e.,

$$|KG(j\omega)| = 1 \quad \text{and} \quad \angle G(j\omega) = 180^\circ$$

This leads to the stability condition

$$|KG(j\omega)| < 1 \quad \text{and} \quad \angle G(j\omega) = -180^\circ$$

which holds for systems that become unstable, when the gain K increases.

Consider the transfer function

$$H_1(s) = \frac{s^2 + 2s + 2}{(s+1)(s+5)}$$

Consider the transfer function

$$H_1(s) = \frac{s^2 + 2s + 2}{(s+1)(s+5)}$$

The transfer function evaluated at a point s_0 in the complex plane is

$$H_1(s_0) = |v|e^{j\alpha}$$

where

$$\alpha = \theta_1 + \theta_2 - \phi_1 - \phi_2$$

with θ_i (ϕ_i) being the angles from the zeros (poles) to s_0.

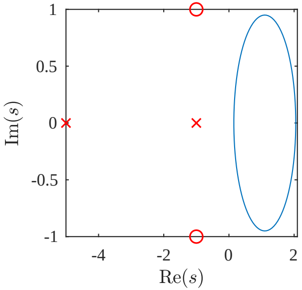
The transfer function evaluated at a point s_0 in the complex plane is

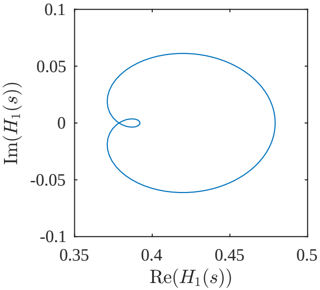
$$H_1(s_0) = |v|e^{j\alpha}$$

where

$$\alpha = \theta_1 + \theta_2 - \phi_1 - \phi_2$$

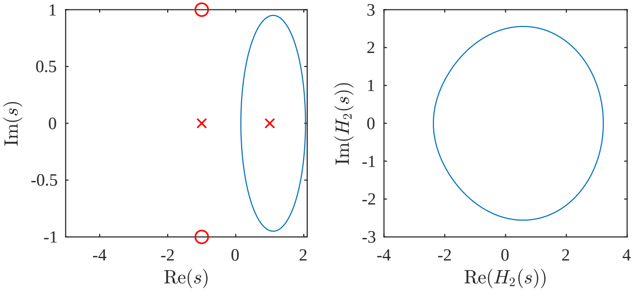
with θ_i (ϕ_i) being the angles from the zeros (poles) to s_0.





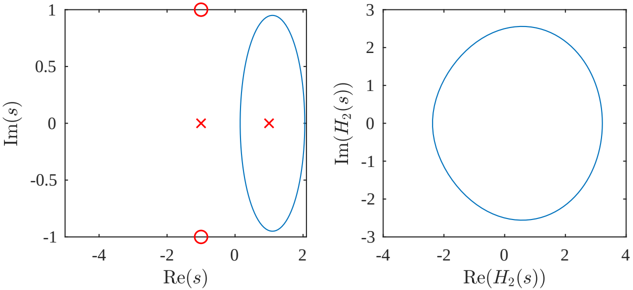
Consider the transfer function

$$H_2(s) = \frac{s^2 + 2s + 2}{(s+1)(s-1)}$$



Consider the transfer function

$$H_2(s) = \frac{s^2 + 2s + 2}{(s+1)(s-1)}$$



The contour encircles the origin in the counter clockwise direction.

Theorem. A contour map of a complex function will encircle the origin $Z - P$ times where Z is the number of zeros and P is the number of poles of the function inside the contour.

The Nyquist Stability Criterion

To verify the stability of a system, one needs to determine the number of closed-loop poles in the right half-plane. Thus, the Nyquist plot is a map of a contour that encircles the entire right-half plane.

To verify the stability of a system, one needs to determine the number of closed-loop poles in the right half-plane. Thus, the Nyquist plot is a map of a contour that encircles the entire right-half plane.

The number of closed-loop poles in the right half-plane equals the number of right half-plane zeros of

$$1 + KG(s)$$

To verify the stability of a system, one needs to determine the number of closed-loop poles in the right half-plane. Thus, the Nyquist plot is a map of a contour that encircles the entire right-half plane.

The number of closed-loop poles in the right half-plane equals the number of right half-plane zeros of

$$1 + KG(s)$$

To only work with the open-loop transfer function $KG(s)$, it suffices to count the number of encirclements of the point -1 when considering a contour that encircles the right half-plane for $KG(s)$.

Let N denote the number of clockwise encirclements of -1. Then the number of zeros in the right half plane Z (closed-loop poles) minus the number of open-loop poles in the right half-plane P is

$$N = Z - P$$