Control Theory (00350188)
lecture no. 1

Leonid Mirkin

Faculty of Mechanical Engineering
Technion — IIT

Outline

General info

- Course site: http://leo.technion.ac.il/Courses/CT/
- Credit points: 3.5
- Prerequisite: Introduction to Control (00340040), a must
- Grading policy:
  2 midterm projects (tokef): 20% each (provided exam is passed)
  Final exam ("closed"): 60% (or 100% is the grade is < 55)
- Passing policy:
  minimum passing grade is 55
  only those who pass both projects are eligible to take the final exam

General info

- Course site: `http://leo.technion.ac.il/Courses/CT/`
- Credit points: 3.5
- Prerequisite: Introduction to Control (00340040), a must
- Grading policy:
  2 midterm projects (tokef): 20% each (provided exam is passed)
  Final exam ("closed"): 60% (or 100% is the grade is < 55)
- Passing policy:
  minimum passing grade is 55
  only those who pass both projects are eligible to take the final exam

General info

- Course site: `http://leo.technion.ac.il/Courses/CT/`
- Credit points: 3.5
- Prerequisite: Introduction to Control (00340040), **a must**
- Grading policy:
  2 midterm projects (tokef): 20% each (provided exam is passed)
  Final exam ("closed"): 60% (or 100% is the grade is < 55)
- Passing policy:
  minimum passing grade is 55
  only those who pass both projects are eligible to take the final exam

General info

- Course site: `http://leo.technion.ac.il/Courses/CT/`
- Credit points: 3.5
- Prerequisite: Introduction to Control (00340040), **a must**
- Grading policy:
  2 midterm projects (tokef): 20% each (provided exam is passed)
  Final exam ("closed"): 60% (or 100% is the grade is < 55)
- Passing policy:
  minimum passing grade is 55
  only those who pass both projects are eligible to take the final exam

General info

- Course site: http://leo.technion.ac.il/Courses/CT/
- Credit points: 3.5
- Prerequisite: Introduction to Control (00340040), **a must**
- Grading policy:
  2 midterm projects (tokef): 20% each (provided exam is passed)
  Final exam ("closed"): 60% (or 100% is the grade is < 55)
- Passing policy:
  minimum passing grade is 55
  only those who pass both projects are eligible to take the final exam

Syllabus

1. Advanced single loop design
   1.1 More on loop-shaping
   1.2 More on on dead-time systems
   1.3 More on pole placement (Sylvester matrix etc.)
   1.4 Industrial control (saturation & anti-windup, reference signal generation)
   1.5 Robustness of control systems

2. Introduction to state-space methods
   2.1 Structural properties (controllability, observability, etc)
   2.2 State feedback control
   2.3 State observers
   2.4 Observer-based output feedback
   2.5 Introduction to optimization-based methods (LQR, Kalman filter, LQG)

3. Introduction to sampled-data systems
   3.1 Digital redesign of analog controllers
   3.2 Digital design

Syllabus

1. Advanced single loop design
    1.1 More on loop-shaping
    1.2 More on on dead-time systems
    1.3 More on pole placement (Sylvester matrix etc.)
    1.4 Industrial control (saturation & anti-windup, reference signal generation)
    1.5 Robustness of control systems

2. Introduction to state-space methods
    2.1 Structural properties (controllability, observability, etc)
    2.2 State feedback control
    2.3 State observers
    2.4 Observer-based output feedback
    2.5 Introduction to optimization-based methods (LQR, Kalman filter, LQG)

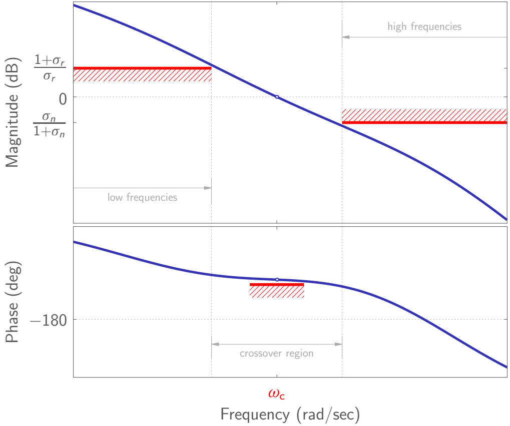
3. Introduction to sampled-data systems
    3.1 Digital redesign of analog controllers
    3.2 Digital design

Syllabus

1. Advanced single loop design
   1.1 More on loop-shaping
   1.2 More on on dead-time systems
   1.3 More on pole placement (Sylvester matrix etc.)
   1.4 Industrial control (saturation & anti-windup, reference signal generation)
   1.5 Robustness of control systems
2. Introduction to state-space methods
   2.1 Structural properties (controllability, observability, etc)
   2.2 State feedback control
   2.3 State observers
   2.4 Observer-based output feedback
   2.5 Introduction to optimization-based methods (LQR, Kalman filter, LQG)
3. Introduction to sampled-data systems
   3.1 Digital redesign of analog controllers
   3.2 Digital design

Outline

Typical operations with $L(\mathrm{j}\omega)$



Typical course of action:

- choose crossover, ω_c
- shape high-freq. roll-off
- set required crossover, ω_c
- shape phase around ω_c
- shape low-frequency gain

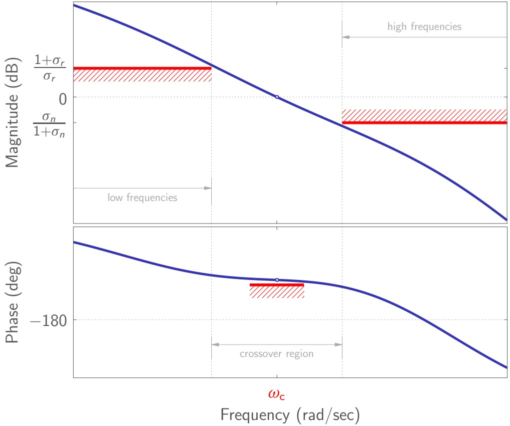
by cascade adjustments of $C(s)$.

But

- one should not be religious about that,

the steps may be skipped, reordered, or altered, depending on the situation.

Typical operations with $L(j\omega)$



Typical course of action:

- choose crossover, ω_c
- shape high-freq. roll-off (means: low-pass filter)
- set required crossover, ω_c (means: proportional controller)
- shape phase around ω_c (means: lead controller)
- shape low-frequency gain (means: lag controller)

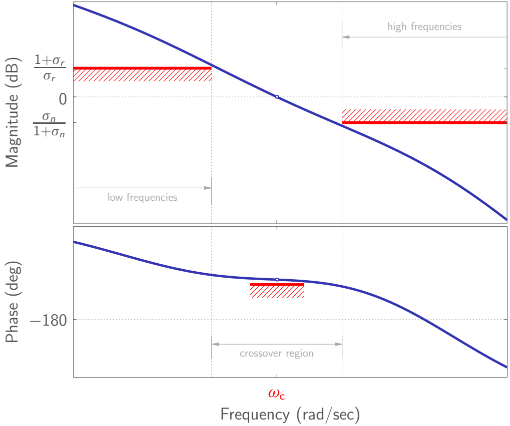
by cascade adjustments of $C(s)$.

But

- one should not be religious about that,

the steps may be skipped, reordered, or altered, depending on the situation.

Typical operations with $L(j\omega)$



Typical course of action:

- choose crossover, ω_c
- shape high-freq. roll-off (means: low-pass filter)
- set required crossover, ω_c (means: proportional controller)
- shape phase around ω_c (means: lead controller)
- shape low-frequency gain (means: lag controller)

by cascade adjustments of $C(s)$.

But

- one should not be religious about that,

the steps may be skipped, reordered, or altered, depending on the situation.

Low-pass filters: Butterworth

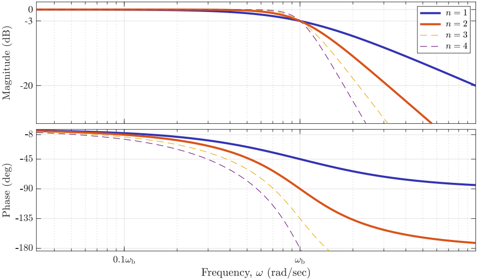
The n-order Butterworth filter[1] with bandwidth ω_{b} is the stable t.f. such that

$$|F_{\mathrm{b},n}(\mathrm{j}\omega)|^2 = \frac{1}{1 + (\omega/\omega_{\mathrm{b}})^{2n}},$$

like

$$F_{\mathrm{b},1}(s) = \frac{\omega_{\mathrm{b}}}{s + \omega_{\mathrm{b}}} \quad \text{or} \quad F_{\mathrm{b},2}(s) = \frac{\omega_{\mathrm{b}}^2}{s^2 + \sqrt{2}\omega_{\mathrm{b}} s + \omega_{\mathrm{b}}^2}$$

with

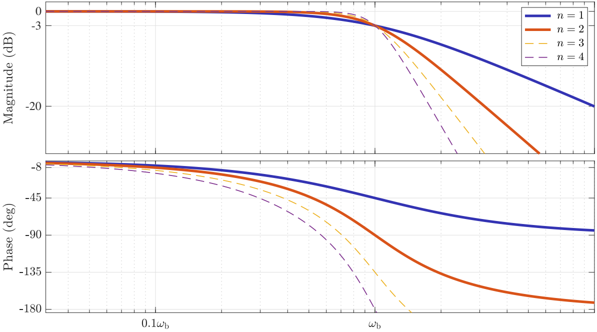


[1] MATLAB: `[num,den] = butter(n,wb,'s'); Fb = tf(num,den);`

Low-pass filter: usage

Main problem is that

- the phase lags before the magnitude starts to decay:



The rule of thumb:

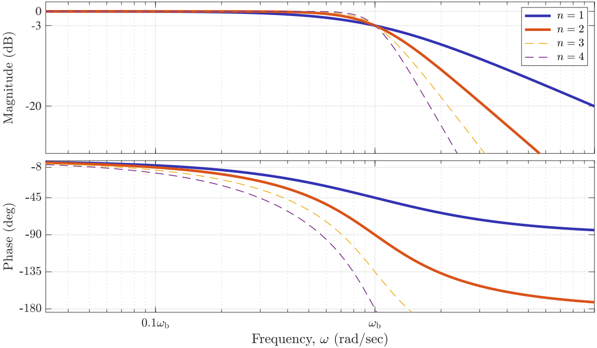
- use $\omega_b = 10\omega_c$ (decade above the intended crossover),

with $\angle F_{b,1}(j\omega_c) \approx -5.7°$, $\angle F_{b,2}(j\omega_c) \approx -8.1°$, $\angle F_{b,3}(j\omega_c) \approx -11.5°$,

Low-pass filter: usage

Main problem is that

– the phase lags before the magnitude starts to decay:



The rule of thumb:

– use $\omega_b = 10\omega_c$ (decade above the intended crossover),

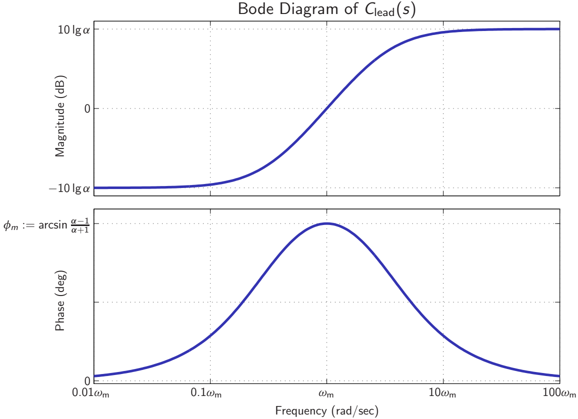
with $\angle F_{b,1}(j\omega_c) \approx -5.7^\circ$, $\angle F_{b,2}(j\omega_c) \approx -8.1^\circ$, $\angle F_{b,3}(j\omega_c) \approx -11.5^\circ$, ...

1-order lead

General form:

$$C_{\mathrm{lead}}(s) = \frac{\sqrt{\alpha}\, s + \omega_{\mathrm{m}}}{s + \sqrt{\alpha}\omega_{\mathrm{m}}}, \qquad \text{with } \alpha = \frac{1 + \sin\phi_m}{1 - \sin\phi_m},$$

where $\phi_m \in (0, 90^\circ)$ is the maximal phase lead (occurs at $\omega = \omega_{\mathrm{m}}$).

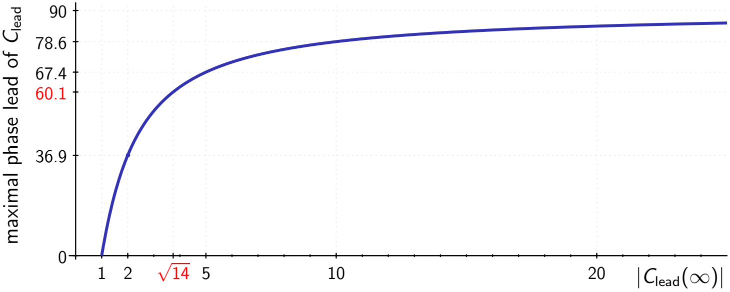


1-order lead: cost of phase lead

Phase lead is expensive, it leads to

- the decrease of the low-frequency gain and
- the increase of the high-frequency gain

of the controller (both by the factor $\sqrt{\alpha}$). Quantitatively,



whose slope decreases. A rule of thumb is that

- the phase lead above 60° might be too expensive.

2-order lead

General form:

$$C_{\mathrm{lead2}}(s) = \frac{\alpha\, s^2 + 2\zeta\sqrt{\alpha}\omega_{\mathrm{m}} s + \omega_{\mathrm{m}}^2}{s^2 + 2\zeta\sqrt{\alpha}\omega_{\mathrm{m}} s + \alpha\omega_{\mathrm{m}}^2}, \qquad \text{with } \zeta \in \Bigl[\frac{1}{\sqrt{2}}, \sqrt{2}\Bigr],$$

and

$$\alpha = 1 + 2\zeta\Bigl(\zeta + \sqrt{\zeta^2 + \cot^2 \tfrac{\phi_m}{2}}\Bigr)\tan^2 \tfrac{\phi_m}{2}$$

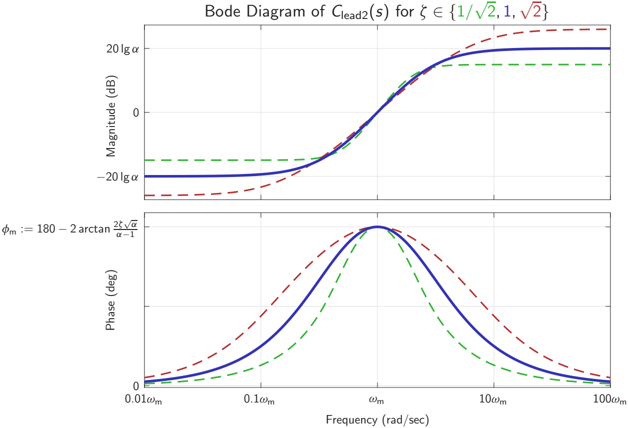
where $\phi_m \in (0, 180^\circ)$ is the maximal phase lead (occurs at $\omega = \omega_{\mathrm{m}}$). Here

- the case $\zeta = 1$ corresponds to $C_{\mathrm{lead2}} = C_{\mathrm{lead}}^2$,
- if $\zeta < 1/\sqrt{2}$, then $|C_{\mathrm{lead2}}(\mathrm{j}\omega)|$ is not monotonic, so might be trickier,
- if $\zeta > \sqrt{2}$, then it might be that $\angle C_{\mathrm{lead2}}(\mathrm{j}\omega) > \phi_m$ for some $\omega \neq \omega_{\mathrm{m}}$.

2-order lead (contd)

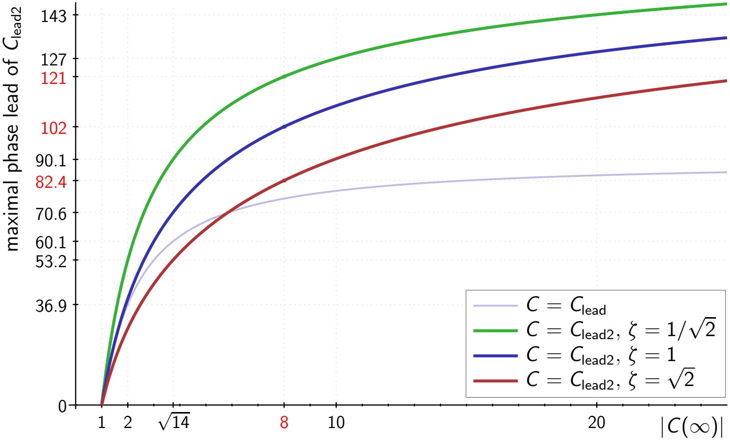
As ζ increases for the same ϕ_m,

- phase lead becomes wider
- α increases



2-order lead: cost of phase lead

Quantitatively,



A rule of thumb is that

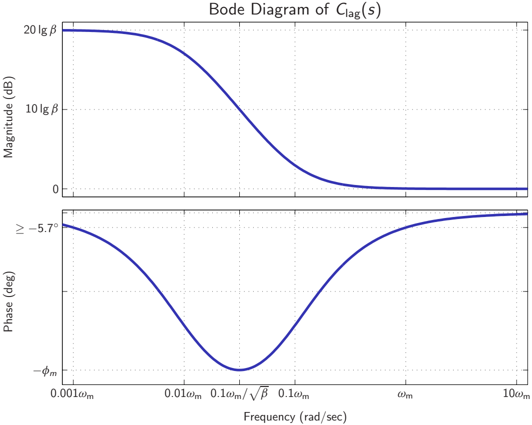
– the phase lead above 120° might be too expensive.

Lag

General form:

$$C_{\mathrm{lag}}(s) = \frac{10s + \omega_{\mathrm{m}}}{10s + \omega_{\mathrm{m}}/\beta}, \quad \text{with } \beta > 1,$$

where the phase lag at $\omega = \omega_{\mathrm{m}}$ is at most 5.7°.



Outline

M-circles

M-circles are contours of constant closed-loop magnitude on Nyquist plane.

Let $L(\mathrm{j}\omega) = x + \mathrm{j}y$. Then $T(\mathrm{j}\omega) = \frac{x+\mathrm{j}y}{1+x+\mathrm{j}y}$. Hence,

$$|T(\mathrm{j}\omega)|^2 = M^2 \iff M^2(1+x)^2 + M^2y^2 = x^2 + y^2$$
$$\iff (1-M^2)x^2 - 2M^2x + (1-M^2)y^2 = M^2.$$

Then two cases are possible:

$M = 1$ then $x = -\frac{1}{2}$ (vertical line)

$M \neq 1$ then $(1-M^2)\left(x^2 - 2\frac{M^2}{1-M^2}x + \frac{M^4}{(1-M^2)^2} + y^2\right) = M^2$, so we get:

$$\left(x - \frac{M^2}{1-M^2}\right)^2 + y^2 = \left(\frac{M}{1-M^2}\right)^2$$

(circle centered at $\frac{M^2}{1-M^2}$ with radius $\frac{M}{|1-M^2|}$)

M-circles

M-circles are contours of constant closed-loop magnitude on Nyquist plane.

Let $L(j\omega) = x + jy$. Then $T(j\omega) = \frac{x+jy}{1+x+jy}$. Hence,

$$
\begin{aligned}
|T(j\omega)|^2 = M^2 &\iff M^2(1+x)^2 + M^2y^2 = x^2 + y^2 \\
&\iff (1-M^2)x^2 - 2M^2x + (1-M^2)y^2 = M^2.
\end{aligned}
$$

Then two cases are possible:

$M = 1$ then $x = -\frac{1}{2}$ (vertical line)

$M \neq 1$ then $(1-M^2)\left(x^2 - 2\frac{M^2}{1-M^2}x \pm \frac{M^4}{(1-M^2)^2} + y^2\right) = M^2$, so we get:

$$\left(x - \frac{M^2}{1-M^2}\right)^2 + y^2 = \left(\frac{M}{1-M^2}\right)^2$$

(circle centered at $\frac{M^2}{1-M^2}$ with radius $\frac{M}{|1-M^2|}$)

M-circles

M-circles are contours of constant closed-loop magnitude on Nyquist plane.

Let $L(\mathrm{j}\omega) = x + \mathrm{j}y$. Then $T(\mathrm{j}\omega) = \frac{x+\mathrm{j}y}{1+x+\mathrm{j}y}$. Hence,

$$\begin{aligned} |T(\mathrm{j}\omega)|^2 = M^2 &\iff M^2(1+x)^2 + M^2y^2 = x^2 + y^2 \\ &\iff (1-M^2)x^2 - 2M^2x + (1-M^2)y^2 = M^2. \end{aligned}$$

Then two cases are possible:

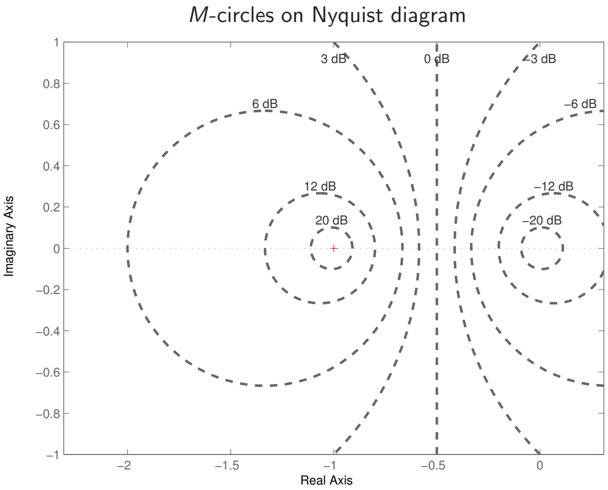
$M = 1$ then $x = -\frac{1}{2}$ (vertical line)

$M \neq 1$ then $(1-M^2)\big(x^2 - 2\frac{M^2}{1-M^2}x \pm \frac{M^4}{(1-M^2)^2} + y^2\big) = M^2$, so we get:

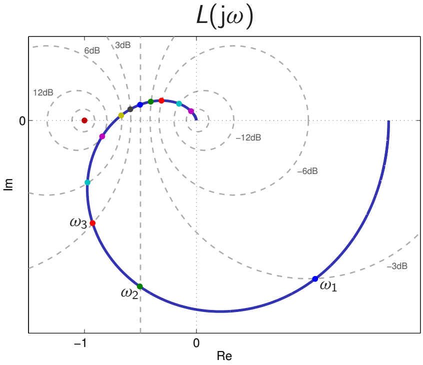
$$\Big(x - \frac{M^2}{1-M^2}\Big)^2 + y^2 = \Big(\frac{M}{1-M^2}\Big)^2$$

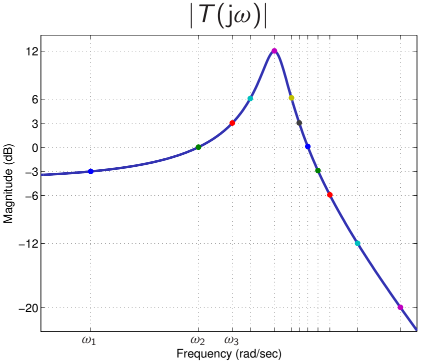
(circle centered at $\frac{M^2}{1-M^2}$ with radius $\frac{M}{|1-M^2|}$)

M-circles (contd)



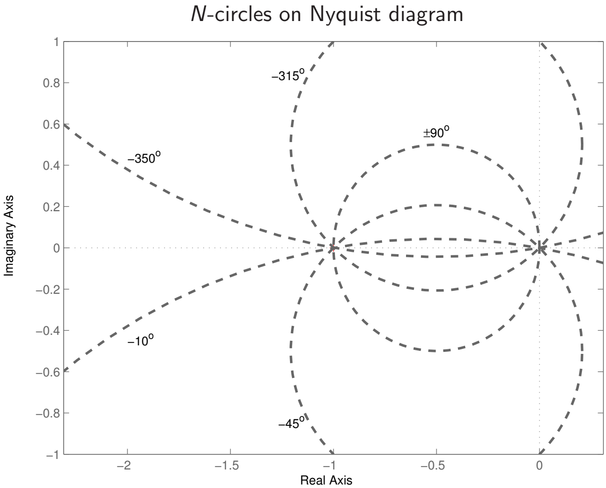
M-circles: how to read





N-circles

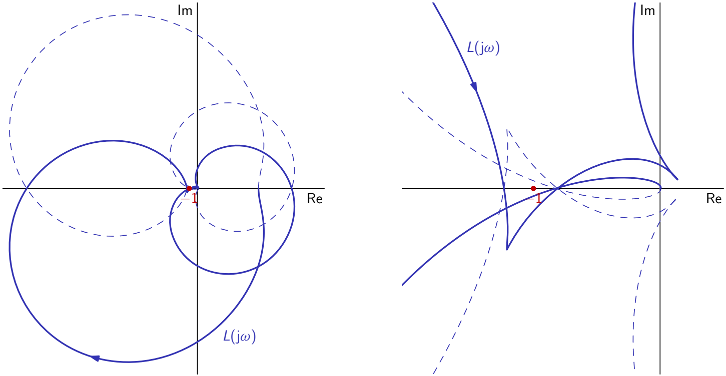
N-circles are contours of constant closed-loop phase on Nyquist plane:



Outline

Nichols chart: motivation

Let $L(s) = \frac{12960000(s+5)(s^2+0.8s+16)(s^2+1.52s+1444)e^{-0.075s}}{(s+10)(s+100)(s^2+3s+9)(s^2+0.48s+9)(s^2+0.8s+1600)^2}$. Its Nyquist plot

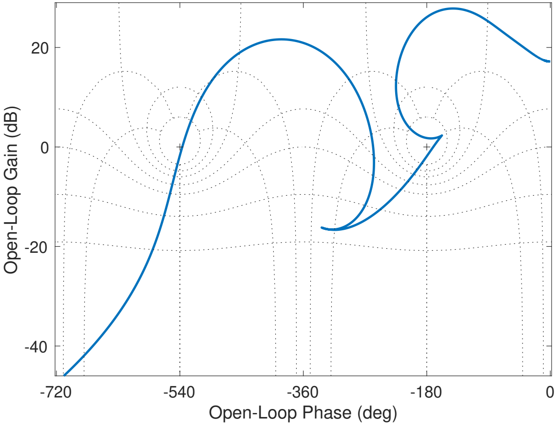


Pitfalls of the Nyquist plot:

- becomes messy for systems with multiple crossover frequencies
- crossover region is imperceptible for systems with large resonant peaks
- lacks system composition (superposition) properties of Bode diagrams

Remedy: Nichols chart

Nichols chart of transfer function $L(s)$ is plot of $|L(j\omega)|$ (in dB) vs. $\angle L(j\omega)$ (in degrees) as frequency ω changes from 0 to ∞.



Nichols chart: advantages

Since phase scale is linear rather than polar,

- Nichols chart is typically cleaner than Nyquist plot

especially for systems with large phase lags, like time-delay systems.

As magnitude scale is in dB, regions with large magnitude don't dominate, hence

- the crossover region is more visible.

Also the consequence of the logarithmic scale of $|L(j\omega)|$ is that

- multiplication of systems results in superposition

on Nichols chart, almost as easy as on the Bode diagrams.

Nichols chart: advantages

Since phase scale is linear rather than polar,

– Nichols chart is typically cleaner than Nyquist plot

especially for systems with large phase lags, like time-delay systems.

As magnitude scale is in dB, regions with large magnitude don't dominate, hence

– the crossover region is more visible.

Also the consequence of the logarithmic scale of $|L(j\omega)|$ is that

– multiplication of systems results in superposition

on Nichols chart, almost as easy as on the Bode diagrams.

Nichols chart: advantages

Since phase scale is linear rather than polar,

– Nichols chart is typically cleaner than Nyquist plot

especially for systems with large phase lags, like time-delay systems.

As magnitude scale is in dB, regions with large magnitude don't dominate, hence

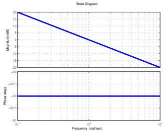
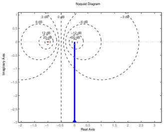
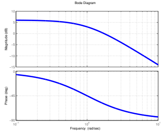
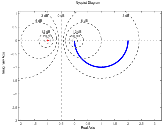
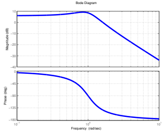
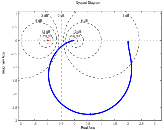
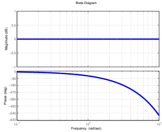
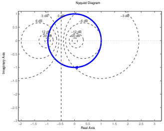
– the crossover region is more visible.

Also the consequence of the logarithmic scale of $|L(j\omega)|$ is that

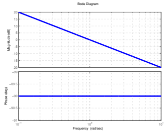
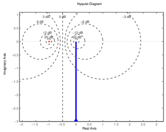
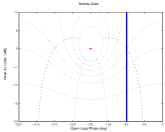
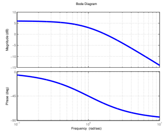
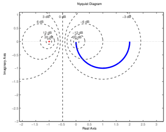
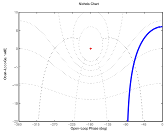
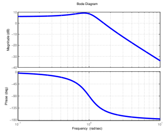
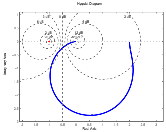
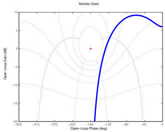
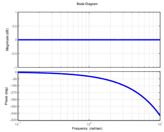
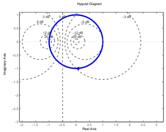
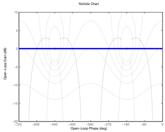
– multiplication of systems results in superposition

on Nichols chart, almost as easy as on the Bode diagrams.

Nichols charts of elementary systems

$L(s)$	Bode	Nyquist
$\frac{k}{s}$		
$\frac{k}{\tau s+1}$		
$\frac{k\omega_{\mathrm{n}}^2}{s^2+2\zeta\omega_{\mathrm{n}}s+\omega_{\mathrm{n}}^2}$		
e^{-sh}		

Nichols charts of elementary systems

$L(s)$	Bode	Nyquist	Nichols
$\dfrac{k}{s}$			
$\dfrac{k}{\tau s+1}$			
$\dfrac{k\omega_{\mathrm{n}}^2}{s^2+2\zeta\omega_{\mathrm{n}}s+\omega_{\mathrm{n}}^2}$			
e^{-sh}			

Nyquist criterion on Nichols chart

The same idea as with the Nyquist plot, we should

- count encirclements of the critical point by the frequency-response plot.

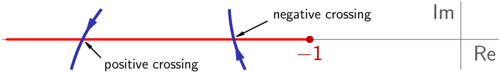
This procedure might be less tangible with the Nichols charts as

- the critical point is not unique there

(any point with $|L(\mathrm{j}\omega)| = 0\,\mathrm{dB}$ & $\arg L(\mathrm{j}\omega) = -180 \pmod{360}$ is critical).

Nyquist criterion on Nichols chart (contd)

Remember (maybe; IC, Lect. 9):



Then the

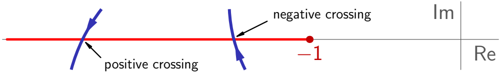
- number of counterclockwise encirclements of $-1 + \mathrm{j}0$ by the Nyquist plot of $L(\mathrm{j}\omega)$ equals *twice* the net sum of crossings the ray $(-\infty, -1]$ by the polar plot of $L(\mathrm{j}\omega)$ (plot direction is with the increase of ω).

The Nichols chart counterpart uses rays $[-180 + 360k, -180 + 360k + \mathrm{j}\infty)$ for $k \in \mathbb{Z}$:

and the rest remains the same. . .

Nyquist criterion on Nichols chart (contd)

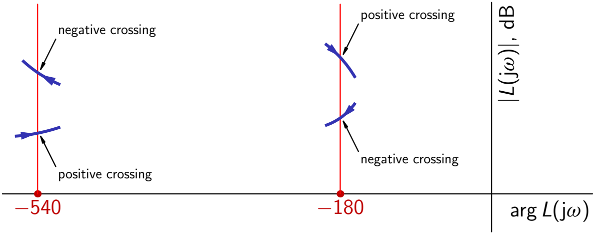
Remember (maybe; IC, Lect. 9):



Then the

- number of counterclockwise encirclements of $-1 + \mathrm{j}0$ by the Nyquist plot of $L(\mathrm{j}\omega)$ equals *twice* the net sum of crossings the ray $(-\infty, -1]$ by the polar plot of $L(\mathrm{j}\omega)$ (plot direction is with the increase of ω).

The Nichols chart counterpart uses rays $[-180 + 360k, -180 + 360k + \mathrm{j}\infty)$ for $k \in \mathbb{Z}$:



and the rest remains the same . . .

Nyquist criterion on Nichols chart: handling integrators

Polar plot: Each integrator adds a counterclockwise arc of 90° with infinite radius, starting at $L(\mathrm{j}\omega)|_{\omega=0^+}$

Nichols chart: An arc centered at the origin has a constant magnitude and changing phase $\implies$ an arc translates to a horizontal line on Nichols chart:

$L(s) = \dfrac{1}{2s(s+1)}$: $\quad L(s) = \dfrac{3(s+1)}{s^2(10s+3)}$: $\quad L(s) = \dfrac{3s+1}{4s^2(s+1)}$:

Each integrator needs -90° of the line.

Nyquist criterion on Nichols chart: handling integrators

Polar plot: Each integrator adds a counterclockwise arc of 90° with infinite radius, starting at $L(\mathrm{j}\omega)|_{\omega=0^+}$

Nichols chart: An arc centered at the origin has a constant magnitude and changing phase $\implies$ an arc translates to a horizontal line on Nichols chart:

$L(s) = \dfrac{1}{2s(s+1)}$: $\quad L(s) = \dfrac{3(s+1)}{s^2(10s+3)}$: $\quad L(s) = \dfrac{3s+1}{4s^2(s+1)}$:

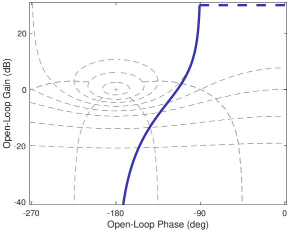
Each integrator needs -90° of the line.

Nyquist criterion on Nichols chart: handling integrators

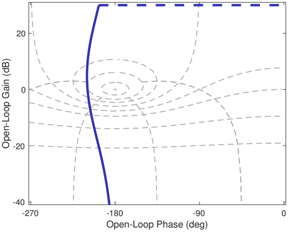
Polar plot: Each integrator adds a counterclockwise arc of 90° with infinite radius, starting at $L(j\omega)|_{\omega=0^+}$

Nichols chart: An arc centered at the origin has a constant magnitude and changing phase $\implies$ an arc translates to a horizontal line on Nichols chart:

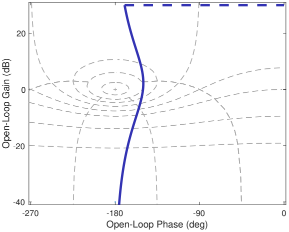
$L(s) = \dfrac{1}{2s(s+1)}$:



$L(s) = \dfrac{3(s+1)}{s^2(10s+3)}$:



$L(s) = \dfrac{3s+1}{4s^2(s+1)}$:

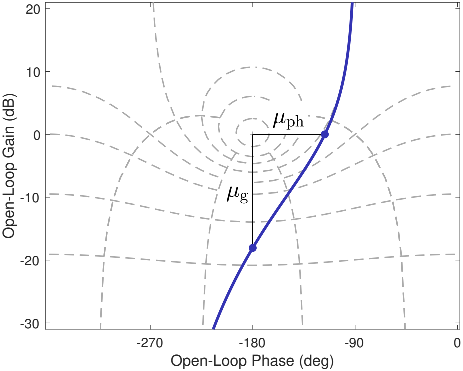


Each integrator needs -90° of the line.

Gain and phase margins on Nichols chart

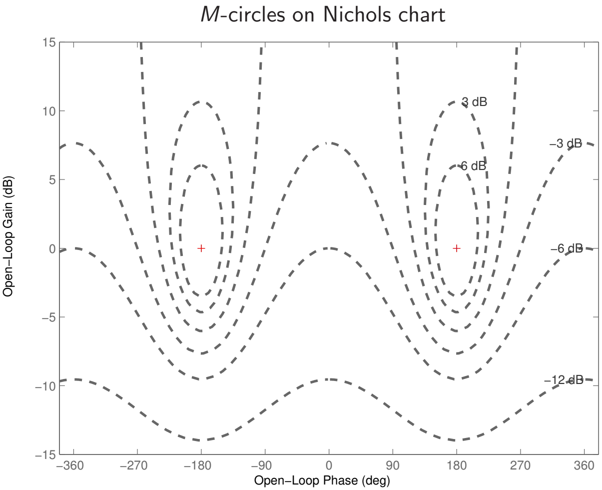
Gain margin μ_{g} and phase margin μ_{ph} are easily calculable from Nichols charts:

- μ_{g} is the vertical distance from the critical point;
- μ_{ph} is the horizontal distance from the critical point.

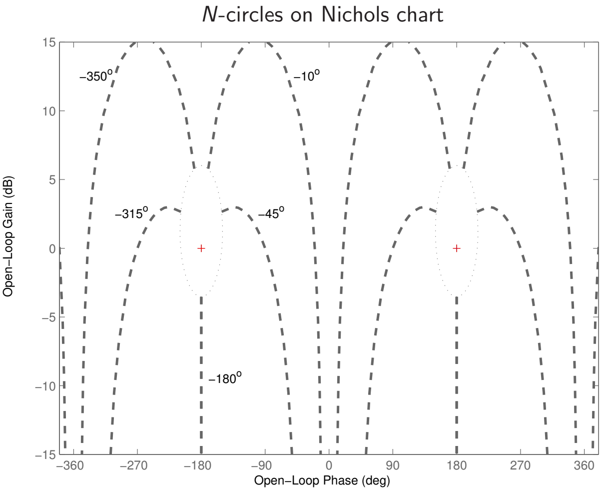


here $\mu_{\mathrm{g}} = 8 \approx 18.06\,\mathrm{dB}$ and $\mu_{\mathrm{ph}} \approx 63.36^\circ$.

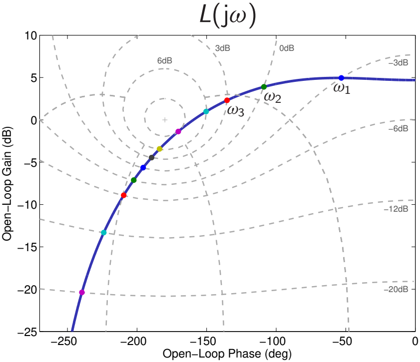
M-circles on Nichols charts

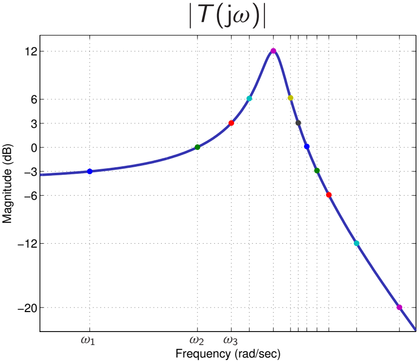


N-circles on Nichols charts



M-circles on Nichols charts: how to read

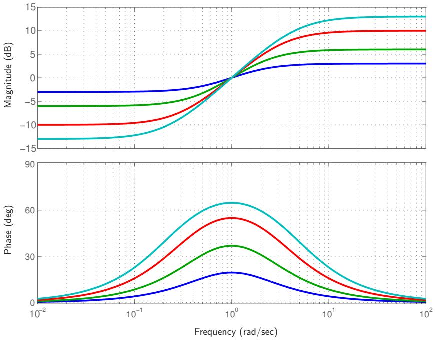




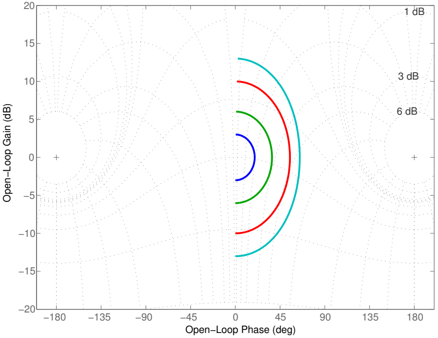
Nichols chart of lead controller

Lead controller: $C_{\mathrm{lead}}(s) = \dfrac{\sqrt{\alpha}\, s + \omega_{\mathrm{m}}}{s + \sqrt{\alpha}\omega_{\mathrm{m}}}$, $\alpha > 1$.

Bode diagrams of $C_{\mathrm{lead}}(s)$ for $\omega_{\mathrm{m}} = 1$ and $\alpha = \{2, 4, 10, 20\}$



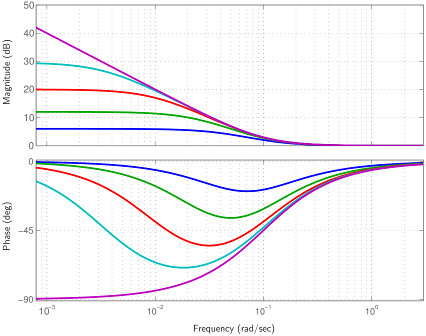
Nichols chart of $C_{\mathrm{lead}}(s)$ for $\omega_{\mathrm{m}} = 1$ and $\alpha = \{2, 4, 10, 20\}$



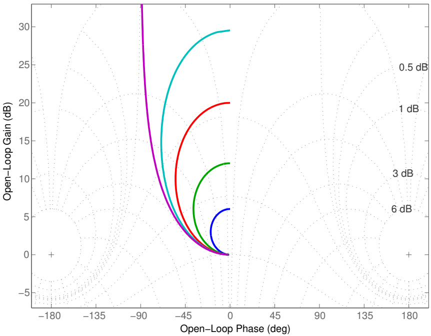
Nichols chart of lag controller

Lag controller: $C_{\text{lag}}(s) = \dfrac{10s + \omega_{\text{m}}}{10s + \omega_{\text{m}}/\beta}$, $\beta > 1$.

Bode diagrams of $C_{\text{lag}}(s)$ for $\omega_{\text{m}} = 1$ and $\beta = \{2, 4, 10, 30, \infty\}$



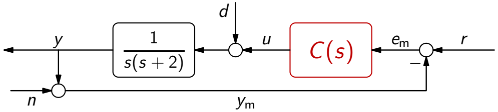
Nichols chart of $C_{\text{lag}}(s)$ for $\omega_{\text{m}} = 1$ and $\beta = \{2, 4, 10, 30, \infty\}$



Outline

System (remember IC, Lect. 11)

A DC motor controlled in closed loop:



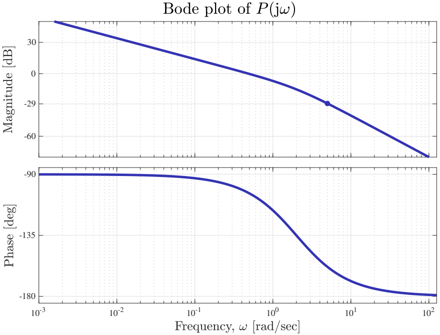
Requirements:

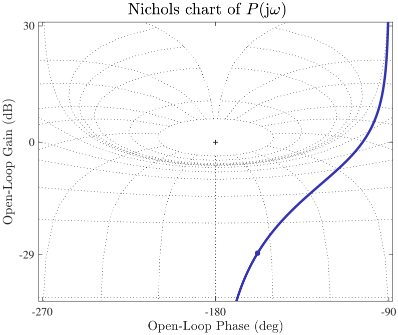
- closed-loop stability (of course)
- $\omega_c = 5\,\mathrm{rad/sec}$
- zero steady-state error for a step in r — always holds
- zero steady-state error for a step in d — integrator in $C(s)$
- $\mu_{ph} \in \{45, 60\}$

Remark: We implicitly assume that the plant is normalized, in a sense that the control amplitude $|u(t)| < 1$ is "small" and $|u(t)| > 1$ is "large".

Example 1: the plant

Let first $\mu_{\mathsf{ph}} = 45^\circ$. With $\omega_{\mathsf{c}} = 5$,

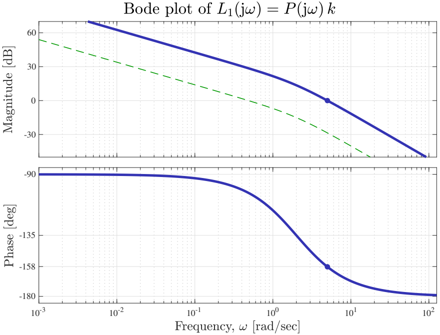


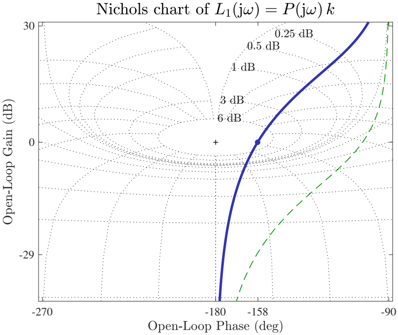


This is below the actual crossover, so can be attained by the gain $k \approx 26.9$.

Example 1: adjusting crossover

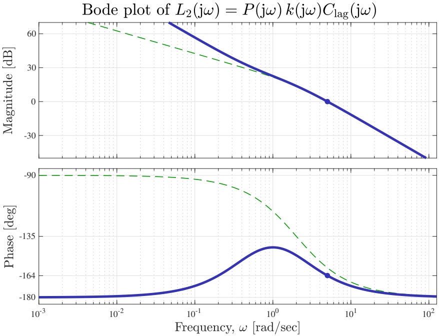
We get:

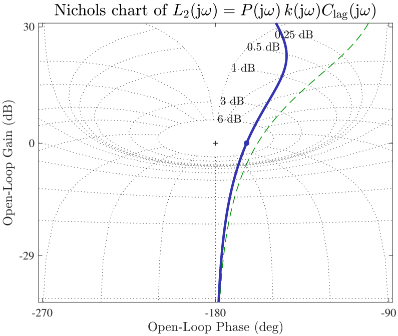




Example 1: adjusting low-frequency gain

Use the lag controller with $\omega_m = 5$ and $\beta = \infty$:





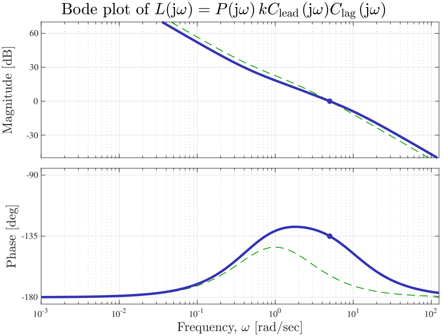
Here $\mu_{\mathrm{ph}} \approx 16^\circ$ and

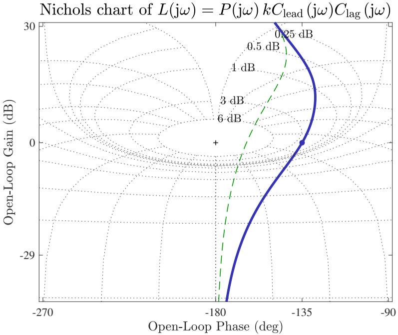
– we need a phase lead of $45^\circ - 16^\circ = 29^\circ$,

for which one lead is enough.

Example 1: adjusting phase around crossover

We get:



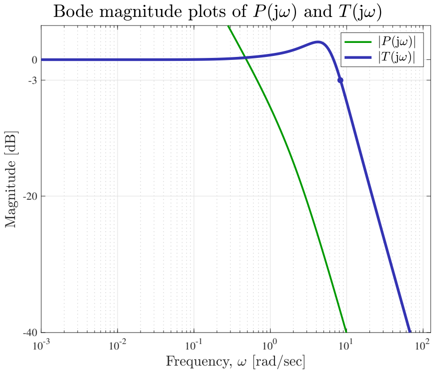


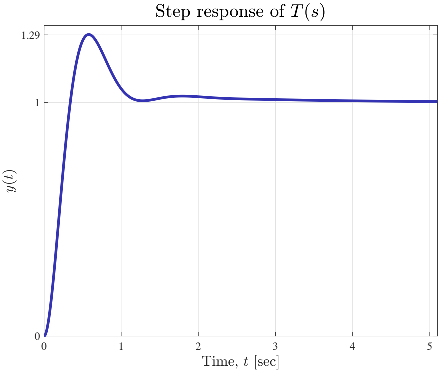
Here $\mu_{\text{ph}} \approx 45^\circ$, which is what we need. Resulting controller:

$$C(s) = kC_{\text{lead}}(s)C_{\text{lag}}(s) = \frac{45.619(s+2.951)(s+0.5)}{s(s+8.471)}.$$

Note Nichols chart location vis-à-vis M-circles (not quite "nice").

Example 1: closed-loop command response



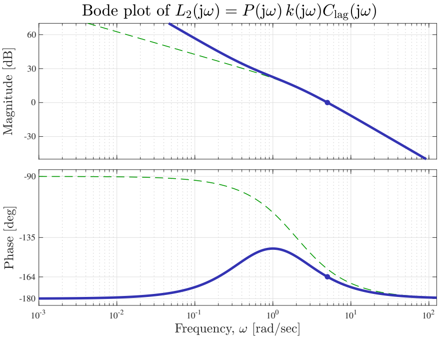


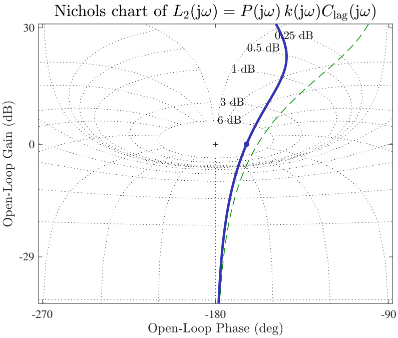
To note:

- resonance peak (agrees with M-circles) $\Longrightarrow$ OS $\approx$ 29%
- closed-loop bandwidth $\omega_\mathsf{b} \approx 8.3176$, which is a bit above the designed $\omega_\mathsf{c} = 5$ and higher than the open-loop bandwidth

Example 2: adjusting low-frequency gain

Now, let $\mu_{\text{ph}} = 60^\circ$. The first design steps, up until the addition of the lag part, remain the same and we have





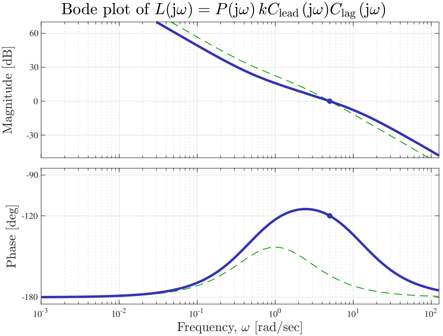
Here $\mu_{\text{ph}} \approx 16^\circ$ and

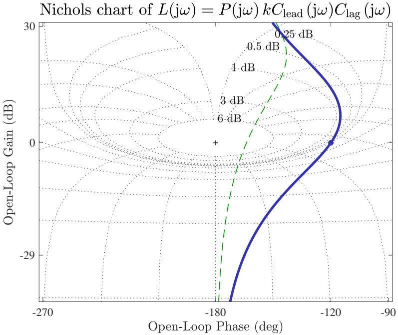
- we now need a phase lead of $60^\circ - 16^\circ = 44^\circ$,

for which one lead is enough as well.

Example 2: adjusting phase around crossover

We get:



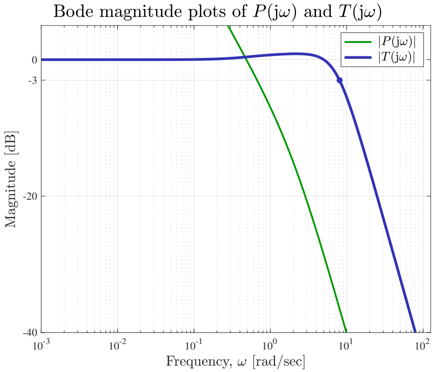


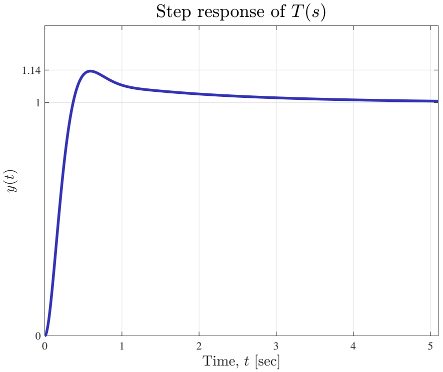
Here $\mu_{\text{ph}} \approx 60^\circ$, which is what we need. Resulting controller:

$$C(s) = kC_{\text{lead}}(s)C_{\text{lag}}(s) = \frac{62.977(s+2.127)(s+0.5)}{s(s+11.75)}.$$

Note again Nichols chart location vis-à-vis M-circles ("nicer").

Example 2: closed-loop command response



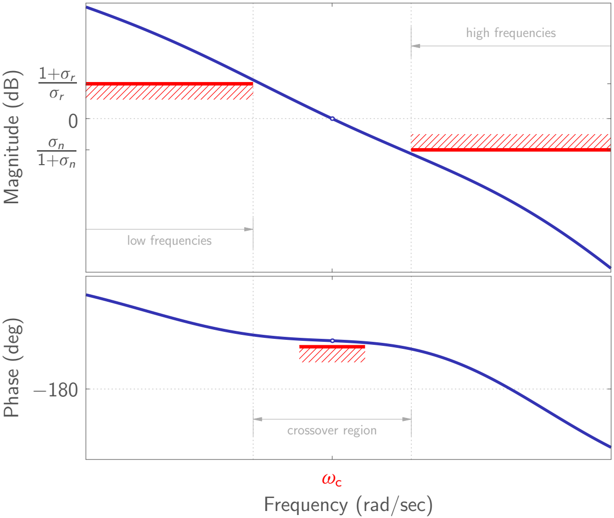


To note:

- resonance peak becomes lower $\implies$ lower OS $\approx 14\%$
- closed-loop bandwidth $\omega_b \approx 8.0649$, which is a bit above the designed $\omega_c = 5$ and higher than the open-loop bandwidth

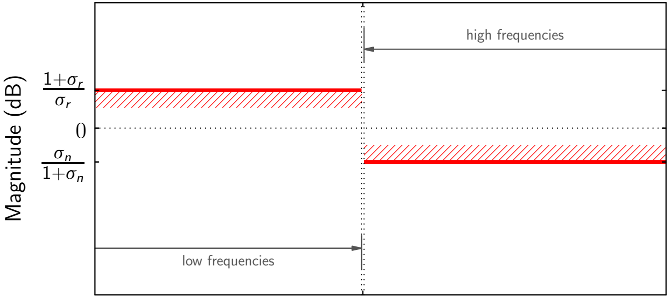
Outline

Loop shaping: big picture



Dream loop shape

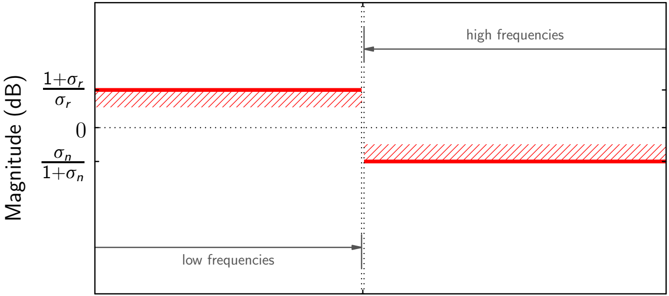
We'd prefer to have narrow crossover region, something like this:



Intuitively, it is hard to believe that this is possible (too good to be true:-).
It turns out that this "intuition" can be rigorously justified.

Dream loop shape

We'd prefer to have narrow crossover region, something like this:



Intuitively, it is hard to believe that this is possible (too good to be true:-).
It turns out that this "intuition" can be rigorously justified.

Bode's gain-phase relation: minimum-phase loop

Let $L(s)$ be stable and minimum-phase and such that $L(0) > 0$. Then $\forall\omega_0$

$$\arg L(\mathrm{j}\omega_0) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} \ln\coth\frac{|\nu|}{2}\,\mathrm{d}\nu, \qquad \text{where } \nu := \ln\frac{\omega}{\omega_0}$$

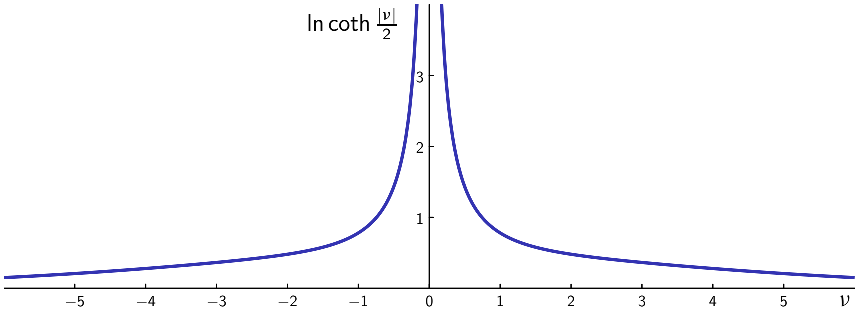
$(\coth x := \frac{\mathrm{e}^x+\mathrm{e}^{-x}}{\mathrm{e}^x-\mathrm{e}^{-x}})$.

Bode's gain-phase relation: minimum-phase loop

Let $L(s)$ be stable and minimum-phase and such that $L(0) > 0$. Then $\forall\omega_0$

$$\arg L(\mathrm{j}\omega_0) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu}\ln\coth\frac{|\nu|}{2}\,\mathrm{d}\nu, \qquad \text{where } \nu := \ln\frac{\omega}{\omega_0}$$

($\coth x := \frac{\mathrm{e}^x+\mathrm{e}^{-x}}{\mathrm{e}^x-\mathrm{e}^{-x}}$). Function $\ln\coth\frac{|\nu|}{2} = \ln\left|\frac{\omega+\omega_0}{\omega-\omega_0}\right|$:



may be thought of as a rough approximation of the Dirac delta.

Bode's gain-phase relation: what does it mean

Since $\ln\coth\frac{|\nu|}{2}$ decreases rapidly as ω deviates from ω_0,

– $\arg L(\mathrm{j}\omega_0)$ depends mostly on $\frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} = \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\ln\omega}$ near frequency ω_0.

But

– $\dfrac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\ln\omega} = \dfrac{\mathrm{d}\log|L(\mathrm{j}\omega)|}{\mathrm{d}\log\omega}$ is the roll-off[2] of the Bode plot of $|L(\mathrm{j}\omega)|$.

It can be shown that

$$\arg L(\mathrm{j}\omega_0) < \begin{cases} -N\times 65.3^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{3}\le\frac{\omega}{\omega_0}\le 3 \\ -N\times 75.3^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{5}\le\frac{\omega}{\omega_0}\le 5 \\ -N\times 82.7^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{10}\le\frac{\omega}{\omega_0}\le 10 \end{cases}$$

In other words,

– high negative slope of $|L(\mathrm{j}\omega)|$ necessarily causes large phase lag.

[2] Roll-off is the absolute value of the negative slope, scaled by 20.

Bode's gain-phase relation: what does it mean

Since $\ln \coth \frac{|\nu|}{2}$ decreases rapidly as ω deviates from ω_0,

– $\arg L(\mathrm{j}\omega_0)$ depends mostly on $\frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} = \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\ln\omega}$ near frequency ω_0.

But

– $\dfrac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\ln\omega} = \dfrac{\mathrm{d}\log|L(\mathrm{j}\omega)|}{\mathrm{d}\log\omega}$ is the roll-off[2] of the Bode plot of $|L(\mathrm{j}\omega)|$.

It can be shown that

$$\arg L(\mathrm{j}\omega_0) < \begin{cases} -N \times 65.3^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{3} \le \frac{\omega}{\omega_0} \le 3 \\ -N \times 75.3^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{5} \le \frac{\omega}{\omega_0} \le 5 \\ -N \times 82.7^\circ, & \text{if roll-off of } |L(\mathrm{j}\omega)| \text{ is } N \text{ for } \frac{1}{10} \le \frac{\omega}{\omega_0} \le 10 \end{cases}$$

In other words,

– high negative slope of $|L(\mathrm{j}\omega)|$ necessarily causes large phase lag.

---

[2] Roll-off is the absolute value of the negative slope, scaled by 20.

Bode's gain-phase relation: what does it mean

Since $\ln\coth\frac{|\nu|}{2}$ decreases rapidly as ω deviates from ω_0,

– $\arg L(j\omega_0)$ depends mostly on $\frac{d\ln|L(j\omega)|}{d\nu} = \frac{d\ln|L(j\omega)|}{d\ln\omega}$ near frequency ω_0.

But

– $\dfrac{d\ln|L(j\omega)|}{d\ln\omega} = \dfrac{d\log|L(j\omega)|}{d\log\omega}$ is the roll-off[2] of the Bode plot of $|L(j\omega)|$.

It can be shown that

$$\arg L(j\omega_0) < \begin{cases} -N\times 65.3^\circ, & \text{if roll-off of } |L(j\omega)| \text{ is } N \text{ for } \frac{1}{3}\le\frac{\omega}{\omega_0}\le 3 \\ -N\times 75.3^\circ, & \text{if roll-off of } |L(j\omega)| \text{ is } N \text{ for } \frac{1}{5}\le\frac{\omega}{\omega_0}\le 5 \\ -N\times 82.7^\circ, & \text{if roll-off of } |L(j\omega)| \text{ is } N \text{ for } \frac{1}{10}\le\frac{\omega}{\omega_0}\le 10 \end{cases}$$

In other words,

– high negative slope of $|L(j\omega)|$ necessarily causes large phase lag.

[2] Roll-off is the absolute value of the negative slope, scaled by 20.

Bode's gain-phase relation: implication

For systems with rigid loops, it is advisable to

- keep loop roll-off $\not\gg 1$ in the crossover region[3]

to guarantee that $L(\mathrm{j}\omega)$ is far enough from the critical point. This, in turn, means that

- low- and high-frequency regions should be well-separated.

This is the reason why our "dream shape" is not an option.

[3] I.e. not much smaller than $-20\,\mathrm{dB/dec}$ slope of $|L(\mathrm{j}\omega)|$.

Bode's gain-phase relation: implication

For systems with rigid loops, it is advisable to

– keep loop roll-off $\not\gg 1$ in the crossover region

to guarantee that $L(\mathrm{j}\omega)$ is far enough from the critical point. This, in turn, means that

– low- and high-frequency regions should be well-separated.

This is the reason why our "dream shape" is not an option.

Bode's gain-phase relation: implication

For systems with rigid loops, it is advisable to

– keep loop roll-off $\not\gg 1$ in the crossover region

to guarantee that $L(\mathrm{j}\omega)$ is far enough from the critical point. This, in turn, means that

– low- and high-frequency regions should be well-separated.

This is the reason why our "dream shape" is not an option.

Gain-phase relation: one nonminimum-phase zero

Let $L(s)$ has one RHP zero at $z > 0$. Then

$$L(s) = \frac{-s+z}{s+z}\,L_{\rm mp}(s)$$

for a minimum-phase $L_{\rm mp}(s)$. Since $\left|\frac{-j\omega+z}{j\omega+z}\right| \equiv 1$, $|L(j\omega)| = |L_{\rm mp}(j\omega)|$ and

$$\begin{aligned}\arg L(j\omega_0) &= \arg L_{\rm mp}(j\omega_0) + \arg\frac{-j\omega_0+z}{j\omega_0+z} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{d\ln|L(j\omega)|}{d\nu}\ln\coth\frac{|\nu|}{2}\,d\nu - 2\arctan\frac{\omega_0}{z}.\end{aligned}$$

Thus,

$\Longrightarrow$ nonminimum-phase zero adds a phase lag (especially at $\omega > z$)

imposing additional constraints on the slope of $|L(j\omega)|$ in crossover region.

Gain-phase relation: one nonminimum-phase zero

Let $L(s)$ has one RHP zero at $z > 0$. Then

$$L(s) = \frac{-s + z}{s + z} L_{\mathrm{mp}}(s)$$

for a minimum-phase $L_{\mathrm{mp}}(s)$. Since $\left|\frac{-\mathrm{j}\omega+z}{\mathrm{j}\omega+z}\right| \equiv 1$, $|L(\mathrm{j}\omega)| = |L_{\mathrm{mp}}(\mathrm{j}\omega)|$ and

$$\begin{aligned} \arg L(\mathrm{j}\omega_0) &= \arg L_{\mathrm{mp}}(\mathrm{j}\omega_0) + \arg \frac{-\mathrm{j}\omega_0 + z}{\mathrm{j}\omega_0 + z} \\ &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{\mathrm{d} \ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} \ln \coth \frac{|\nu|}{2} \mathrm{d}\nu - 2 \arctan \frac{\omega_0}{z}. \end{aligned}$$

Thus,

– nonminimum-phase zero adds a phase lag (especially at $\omega > z$)

imposing additional constraints on the slope of $|L(\mathrm{j}\omega)|$ in crossover region.

Gain-phase relation: complex nonminimum-phase zeros

Now, let $L(s)$ has a pair of RHP zero at $z_\mathrm{r} \pm \mathrm{j}z_\mathrm{i}$, $z_\mathrm{r} > 0$. Then

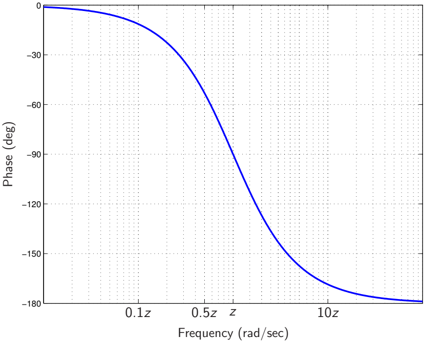
$$L(s) = \frac{-s + z_\mathrm{r} + \mathrm{j}z_\mathrm{i}}{s + z_\mathrm{r} + \mathrm{j}z_\mathrm{i}} \frac{-s + z_\mathrm{r} - \mathrm{j}z_\mathrm{i}}{s + z_\mathrm{r} - \mathrm{j}z_\mathrm{i}} L_\mathrm{mp}(s)$$

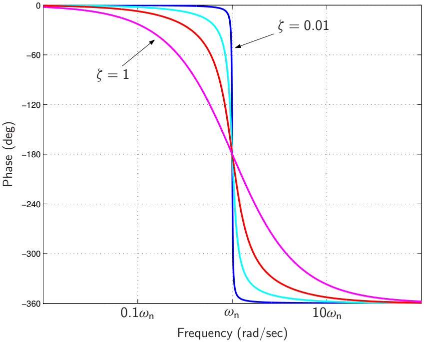
and we have:

$$\begin{aligned}\arg L(\mathrm{j}\omega_0) &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} \ln\coth\frac{|\nu|}{2}\mathrm{d}\nu + \arg\frac{-\mathrm{j}\omega_0 + z_\mathrm{r} \pm \mathrm{j}z_\mathrm{i}}{\mathrm{j}\omega_0 + z_\mathrm{r} \pm \mathrm{j}z_\mathrm{i}} \\ &= \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu} \ln\coth\frac{|\nu|}{2}\mathrm{d}\nu \\ &\quad - 2\Big(\arctan\frac{\omega_0 + z_\mathrm{i}}{z_\mathrm{r}} + \arctan\frac{\omega_0 - z_\mathrm{i}}{z_\mathrm{r}}\Big).\end{aligned}$$

This harden constraints when $\omega_0 > z_\mathrm{i}$, though may soften when $\omega_0 \ll z_\mathrm{i}$.

Phase of all-pass systems





$$\arg \left.\frac{-s+z}{s+z}\right|_{s=j\omega}$$

$$\arg \left.\frac{s^2-2\zeta\omega_n s+\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}\right|_{s=j\omega}$$

Gain-phase relation: multiple nonminimum-phase zeros

In this case

$$L(s) = \frac{-s+z_1}{s+z_1}\frac{-s+z_2}{s+z_2}\cdots\frac{-s+z_k}{s+z_k}\,L_{\mathrm{mp}}(s)$$

and we have:

$$\arg L(\mathrm{j}\omega_0) = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\mathrm{d}\ln|L(\mathrm{j}\omega)|}{\mathrm{d}\nu}\ln\coth\frac{|\nu|}{2}\mathrm{d}\nu - \sum_{i=1}^{k}\arg\frac{-\mathrm{j}\omega_0+z_i}{\mathrm{j}\omega_0+z_i},$$

which further harden constraints.

Limitations due to nonminimum-phase zeros

For systems with a single crossover frequency[3] RHP zeros near ω_c

– impose additional limitations on the roll-off in the crossover region.

Consequently, a well-known rule of thumb says that for nonminimum-phase systems

– crossover frequency ω_c should be $<$ the smallest RHP zero.

Also, it is safe to claim (regarding RHP zeros) that

– closer to the real axis $\implies$ more restrictive crossover limitations

[3] For lightly damped systems it sometimes might be desirable to inject a phase lag by adding RHP zeros. Yet this must be done with maximal care (don't try it at home!)

Limitations due to nonminimum-phase zeros

For systems with a single crossover frequency RHP zeros near ω_c

– impose additional limitations on the roll-off in the crossover region.

Consequently, a well-known rule of thumb says that for nonminimum-phase systems

– crossover frequency ω_c should be $<$ the smallest RHP zero.

Also, it is safe to claim (regarding RHP zeros) that

– closer to the real axis $\Longrightarrow$ more restrictive crossover limitations

Limitations due to nonminimum-phase zeros

For systems with a single crossover frequency RHP zeros near ω_c

- impose additional limitations on the roll-off in the crossover region.

Consequently, a well-known rule of thumb says that for nonminimum-phase systems

- crossover frequency ω_c should be $<$ the smallest RHP zero.

Also, it is safe to claim (regarding RHP zeros) that

- closer to the real axis $\implies$ more restrictive crossover limitations

Outline

Loop-shaping tools

M- and N-circles

Nichols chart

Design example: the use of Nichols charts

Bode's gain-phase relation

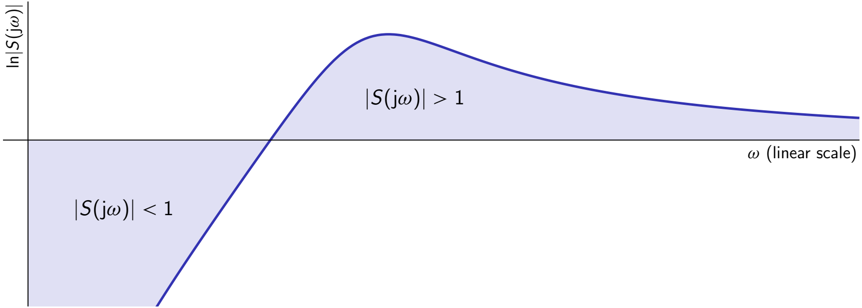
Philosophical remark: Bode's sensitivity integral

Bode's sensitivity integral

Let $L(s)$ be a loop transfer function having pole excess ≥ 2. Then, provided $S(s)$ is stable,

$$\int_0^\infty \ln|S(\mathrm{j}\omega)|\mathrm{d}\omega = \begin{cases} 0 & \text{if } L \text{ stable} \\ \pi \sum_{i=1}^m \operatorname{Re} p_i & \text{otherwise } (p_i\text{—unstable poles of } L) \end{cases}$$

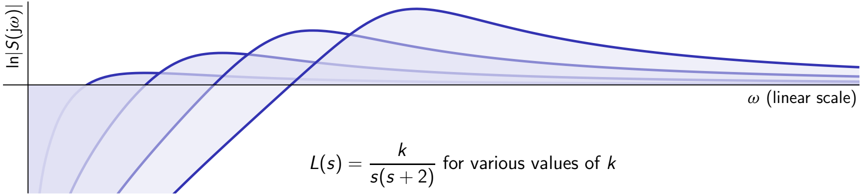
i.e.



What does it mean ?

Some conclusions:

- since $\pi \sum \operatorname{Re} p_i \geq 0$, $|S(j\omega)|$ cannot[3] be < 1 over all frequencies
- improvements in one region inevitable cause deterioration in other (so-called waterbed effect)



On qualitative level,

- controller can only redistribute $|S(j\omega)|$ over frequencies

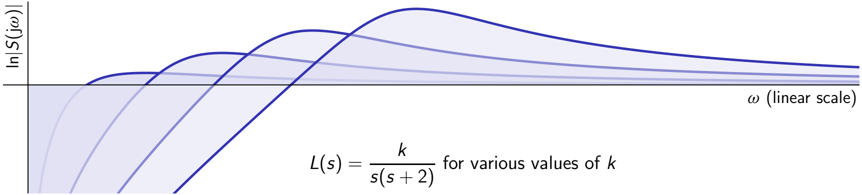
and the art of control may thus be seen as art of redistribution of $|S(j\omega)|$.

[3] If pole excess of $L(s)$ is ≥ 2, of course. Yet this is typical in applications.

What does it mean ?

Some conclusions:

- since $\pi \sum \operatorname{Re} p_i \geq 0$, $|S(j\omega)|$ cannot be < 1 over all frequencies
- improvements in one region inevitable cause deterioration in other (so-called waterbed effect)



On qualitative level,

- controller can only redistribute $|S(j\omega)|$ over frequencies

and the art of control may thus be seen as art of redistribution of $|S(j\omega)|$.