Loop Shaping

Bo Bernhardsson and Karl Johan Åström

Department of Automatic Control LTH,
Lund University

Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Introduction

- A powerful classic design method
- Electronic Amplifiers (Bode, Nyquist, Nichols, Horowitz)

  Command signal following

  Robustness to gain variations, phase margin φ_m

  Notions of minimum and non-minimum phase

  Bode *Network Analysis and Feedback Amplifier Design* 1945
- Servomechanism theory

  Nichols chart

  James Nichols Phillips *Theory of Servomechanisms* 1947
- Horowitz (see QFT Lecture)
  - Robust design of SISO systems for specified process variations
  - 2DOF, cost of feedback, QFT
  - Horowitz *Quantitative Feedback Design Theory - QFT* 1993
- $\mathcal{H}_\infty$ - Loopshaping (see $\mathcal{H}_\infty$ Lecture)

  Design of robust controllers with high robustness

  Mc Farlane Glover *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions* 1989

Harry Nyquist 1889-1976

From farm life in Nilsby Värmland to Bell Labs

Dreaming to be a teacher

- Emigrated 1907
- High school teacher 1912
- MS EE U North Dakota 1914
- PhD Physics Yale 1917
- Bell Labs 1917

Key contributions

- Johnson-Nyquist noise
- The Nyquist frequency 1932
- Nyquist's stability theorem

Hendrik Bode 1905-1982

- Born Madison Wisconsin
- Child protégé, father prof at UIUC, finished high school at 14
- Too young to enter UIUC
- Ohio State BA 1924, MA 1926 (Math)
- Bell Labs 1929

  Network theory
  Missile systems
  Information theory

- PhD Physics Columbia 1936
- Gordon McKay Prof of Systems Engineering at Harvard 1967 (Bryson and Brockett held this chair later)

Bode on Process Control and Electronic Amplifiers

The two fields are radically different in character and emphasis. ... The fields also differ radically in their mathematical flavor. The typical regulator system can frequently be described, in essentials, by differential equations by no more than perhaps the second, third or fourth order. On the other hand, the system is usually highly nonlinear, so that even at this level of complexity the difficulties of analysis may be very great. ... As a matter of idle, curiosity, I once counted to find out what the order of the set of equations in an amplifier I had just designed would have been, if I had worked with the differential equations directly. It turned out to be 55

Bode *Feedback - The History of and Idea* 1960

Nathaniel Nichols 1914 - 1997

- B.S. in chemistry in 1936 from Central Michigan University,
- M.S. in physics from the University of Michigan in 1937
- Taylor Instruments 1937-1946
- MIT Radiation Laboratory Servo Group leader 1942-46
- Taylor Instrument Company Director of research 1946-50

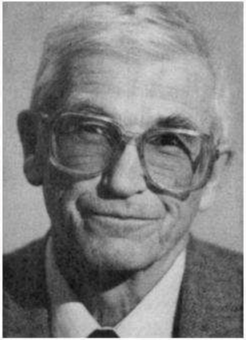
- Aerospace Corporation, San Bernadino, Director of the sensing and information division

http://ethw.org/Archives:Conversations_with_the_Elders_-_Nathaniel_Nichols

Start part 1 at Taylor: 26 min, at MIT:36 min

Isaac Horowitz 1920 - 2005

- B.Sc. Physics and Mathematics University of Manitoba 1944.
- B.Sc. Electrical Engineering MIT 1948
- Israel Defence Forces 1950-51
- M.E.E. and D.E.E. Brooklyn Poly 1951-56 (PhD supervisor Truxal who was supervised by Guillemin)
- Prof Brooklyn Poly 1956-58
- Hughes Research Lab 1958-1966
- EE City University of New York 1966-67
- University of Colorado 1967-1973
- Weizmann Institute 1969-1985
- EE UC Davis 1985-91
- Air Force Institute of Technology 1983-92



Horowitz on Feedback

Horowitz IEEE CSM 4 (1984) 22-23

It is amazing how many are unaware that the primary reason for feedback in control is uncertainty. ...

And why bother with listing all the states if only one could actually be measured and used for feedback? If indeed there were several available, their importance in feedback was their ability to drastically reduce the effect of sensor noise, which was very transpared in the input-output frequency response formulation and terribly obscure in the state-variable form. For these reasons, I stayed with the input-output description.

Important Ideas and Theory

Concepts

- Architecture with two degrees of freedom
- Effect and cost of feedback
- Feedforward and system inversion
- The Gangs of Four and Seven
- Nyquist, Hall, Bode and Nichols plots
- Notions of minimum and non-minimum phase

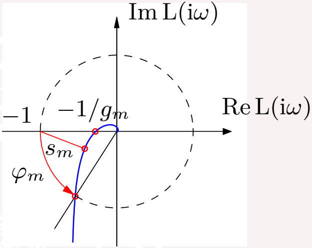
Theory

- Bode's relations
- Bode's phase area formula
- Fundamental limitations
- Crossover frequency inequality

Tools

- Bode and Nichols charts, lead, lag and notch filters

The Nyquist Plot



- Strongly intuitive
- Stability and Robustness

  Stability margins φ_m, g_m,
  $s_m = 1/M_s$
  Frequencies ω_{ms}, ω_{gc}, ω_{pc}
- Disturbance attenuation

  Circles around -1, ω_{sc}
- Process variations

  Easy to represent in the Nyquist plot
  Parameters sweep and level curves of $|T(i\omega)|$
- Measurement noise not easily visible
- Command signal response

  Level curves of complementary sensitivity function
- Bode plot similar but easier to use for design because its wider frequency range

Impact of the Nyquist Theorem at ASEA

Free translation from seminar by Erik Persson ABB in Lund 1970.

We had designed controllers by making simplified models, applying intuition and analyzing stability by solving the characteristic equation. (At that time, around 1950, solving the characteristic equation with a mechanical calculator was itself an ordeal.) If the system was unstable we were at a loss, we did not know how to modify the controller to make the system stable. The Nyquist theorem was a revolution for us. By drawing the Nyquist curve we got a very effective way to design the system because we know the frequency range which was critical and we got a good feel for how the controller should be modified to make the system stable. We could either add a compensator or we could use an extra sensor.

Why did it take 18 years? Nyquist's paper was published 1932!

Example: ASEA Depth Control of Submarine

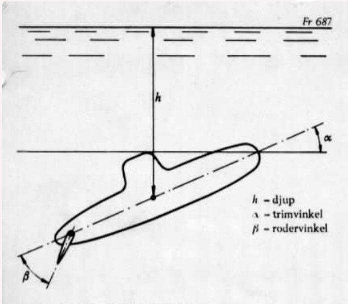


Fig. 1. Fartyget i undervattensläge.

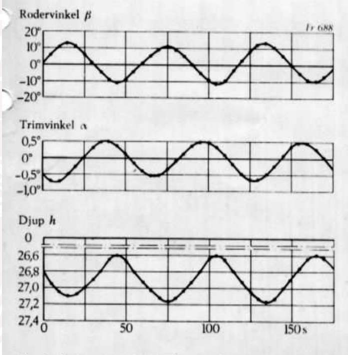


Fig. 2. Registreringar från frekvensanalysproven.

- Toolchain: measure frequency response design by loopshaping
- Fearless experimentation
- Generation of sine waves and measurement
- Speed dependence

Example: ASEA Multivariable Design

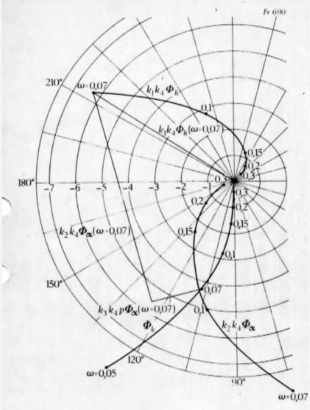


Fig. 4. Nyquist-diagram för regleringssystemet.

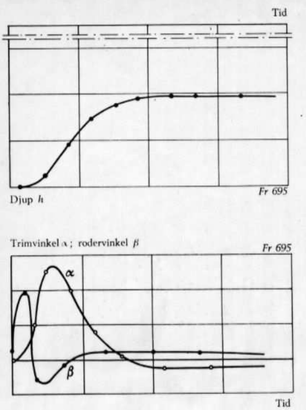


Fig. 9. Regleringsförlopp vid omställning av djupet.

Control System Design - Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Loop Shaping Design

- Determine transfer function from experiments or physics
- Translate specifications to requirements on the loop transfer function $L = PC$
- Important parameters

  Gain crossover frequency ω_{gc} and slope n_{gc} at crossover

  Low frequency slope of loop transfer function n

  High frequency roll off

  Watch out for fundamental limitations
- The controller is given by $C = L_{desired}/P$
- Design can also be done recursively

  Lag compensation

  Lead compensation

  Notch filters

Requirements

- Stablity and robustness

  Gain margin g_m, phase margin φ_m, maximum sensitivity M_s

  Stability for large process variations: $\dfrac{|\Delta P|}{|P|} < \dfrac{|1+PC|}{|PC|}$,

- Load disturbance attenuation

  $$\frac{Y_{cl}(s)}{Y_{ol}(s)} = \frac{1}{1+PC}$$

  - Can be visualized in Hall and Nichols charts

- Measurement Noise

  $$-\frac{U(s)}{N(s)} = \frac{C}{1+PC}$$

- Command signal following (system with error feedback)

  $T = \dfrac{PC}{1+PC}$ can be visualized in Hall and Nichols charts

  Fix shape with feedforward F

How are these quantities represented in loop shaping when we typically explore Bode, Nyquist or Nichols plots?

The Bode Plot

- Stability and Robustness

  Gain and phase margins g_m, φ_m, delay margins
  Frequencies ω_{gc}, ω_{pc}

- Disturbance attenuation

  Sensitivity function $S = \dfrac{1}{1+PC}$
  $P/(1+PC) \approx 1/C$ for low frequencies

- Process variations

  Can be represented by parameter sweep

- Measurement noise

  Visible if process transfer function is also plotted
  Useful to complement with gain curves of GoF

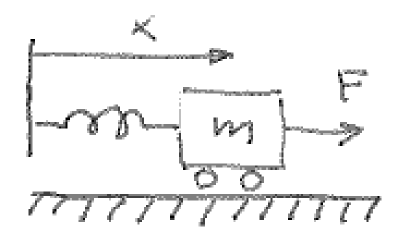
- Command signal response

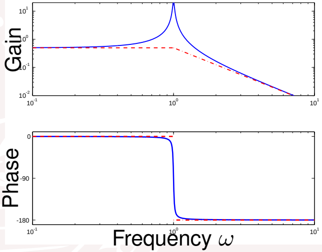
  Level curves of T in Nichols plot

- Wide frequency range

Physical Interpretations of the Bode Plot

- Logarithmic scales gives an overview of the behavior over wide frequency and amplitude ranges
- Piece-wise linear approximations admit good physical interpretations, observe units and scales

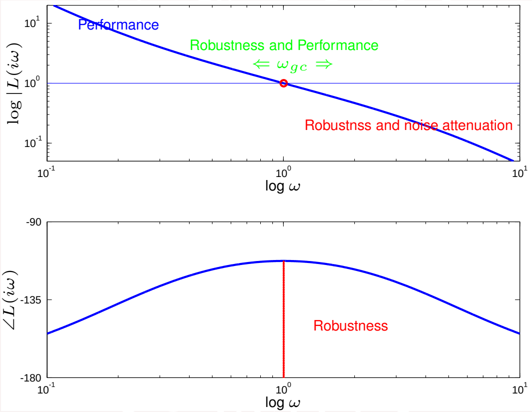




- Low frequencies $G_{xF}(s) \approx 1/k$, the spring line, system behaves like a spring for low frequency excitation
- High frequencies $G_{xF}(s) \approx 1/(ms^2)$, the mass line,, system behaves like a mass for high frequency excitation

Bode Plot of Loop Transfer Function

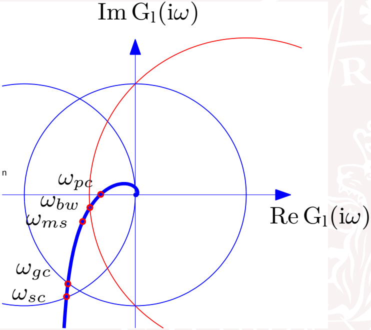
A Bode plot of the loop transfer function $P(s)C(s)$ gives a broad characterization of the feedback system

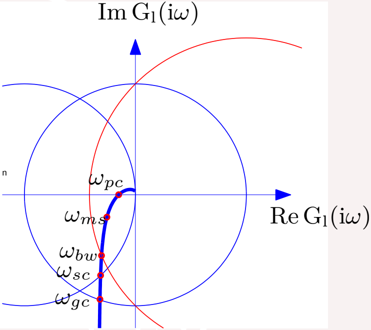


Notice relations between the frequencie $\omega_{gc} \approx \omega_{sc} \approx \omega_{bw}$

Requirements above ω_{gc}

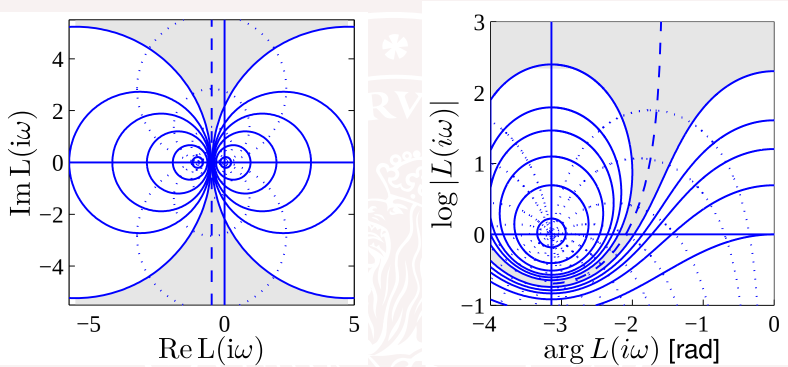
Some Interesting Frequencies





- The frequencies ω_{gc} and ω_{sc} are close
- Their relative order depends on the phase margin (borderline case $\varphi_m = 60^\circ$)

Hall and Nichols Chart



Hall is a Nyquist plot with level curves of gain and phase for the complementary sensitivity function T. Nichols=$\log$ Hall.

Both make is possible to judge T from a plot of PC

Conformality of gain and phase curves depend on scales

The Nichols chart covers a wide frequency range

The Robustness Valley $\mathrm{Re}\, \mathrm{L}(\mathrm{i}\omega) = -1/2$ dashed

Finding a Suitable Loop Transfer Function

Process uncertainty

- Add process uncertainty to the process transfer function
- Perform the design for the worst case (more in QFT)

Disturbance attenuation

- Investigate requirements pick ω_{gc} and slope that satisfies the requirements

Robustness

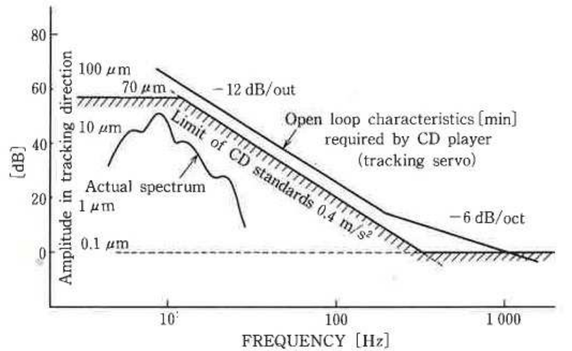
- Shape the loop transfer function around ω_{gc} to give sufficient phase margin
- Add high frequency roll-off

Measurement noise

- Not visible in L but can be estimated if we also plot P

An Example

Translate requirements on tracking error and robustness to demands on the Bode plot for the radial servo of a CD player.



From Nakajima et al Compact Disc Technology, Ohmsha 1992, page 134

Major disturbance caused by eccentricity about $70\mu m$, tracking requirements $0.1\mu m$, requires gain of 700, the RPM varies because of constant velocity read out (1.2-1.4 m/s) around 10 Hz.

Bode on Loopshaping

Bode Network Analysis and Feedback Amplifier Design p 454

The essential feature is that the gain around the feedback loop be reduced from the large value which it has in the useful frequency band to zero dB or less at some higher frequency without producing an accompanying phase shift larger than some prescribed amount. ...

If it were not for the phase restriction it would be desirable on engineering grounds to reduce the gain very rapidly. The more rapidly the feedback vanishes for example, the narrower we need make the region in which active design attention is required to prevent singing. ...

Moreover it is evidently desirable to secure a loop cut-off as soon as possible to avoid the difficulties and uncertainties of design which parasitic elements in the circuit introduce at high frequencies.

But the analysis in Chapter XIV (Bode's relations) shows that the phase shift is broadly proportional to the rate at which the gain changes. ... A phase margin of 30° *correspond to a slope of -5/3.*

Bode's Relations between Gain and Phase

While no unique relation between attenuation and phase can be stated for a general circuit, a unique relation does exist between any given loss characteristic and the *minimum phase shift* which must be associated with it.

$$\arg G(i\omega_0) = \frac{2\omega_0}{\pi}\int_0^\infty \frac{\log|G(i\omega)| - \log|G(i\omega_0)|}{\omega^2-\omega_0^2}d\omega$$

$$= \frac{1}{\pi}\int_0^\infty \frac{d\log|G(i\omega)|}{d\log\omega}\log\left|\frac{\omega+\omega_0}{\omega-\omega_0}\right|d\omega \approx \frac{\pi}{2}\frac{d\log|G(i\omega)|}{d\log\omega}$$

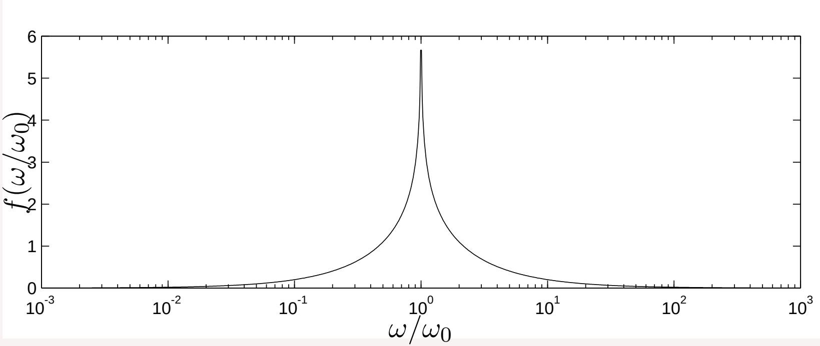
$$\frac{\log|G(i\omega)|}{\log|G(i\omega_0)|} = -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{\omega^{-1}\arg G(i\omega) - \omega_0^{-1}\arg G(i\omega_0)}{\omega^2-\omega_0^2}d\omega$$

$$= -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{d\big(\omega^{-1}\arg G(i\omega)\big)}{d\omega}\log\left|\frac{\omega+\omega_0}{\omega-\omega_0}\right|d\omega$$

Proven by contour integration

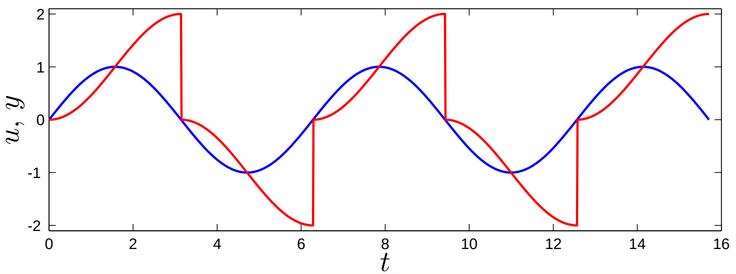
The Weighting Function

$$f\left(\frac{\omega}{\omega_0}\right) = \frac{2}{\pi^2} \log \frac{|\omega + \omega_0|}{|\omega - \omega_0|}$$



Do Nonlinearities Help?

- Can you beat Bode's relations by nonlinear compensators
- Find a compensator that gives phase advance with less gain than given by Bode's relations
- The Clegg integrator has the describing function $N(i\omega) = \frac{4}{\pi\omega} - i\frac{1}{w}$. The gain is $1.62/\omega$ and the phaselag is only 38°. Compare with integrator (J. C. Clegg A nonlinear Integrator for Servomechanisms. Trans. AIEE, part II, 77(1958)41-42)



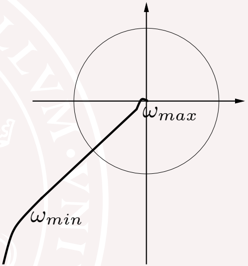
Control System Design - Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Loop Shaping for Gain Variations

- The repeater problem
- Large gain variations in vacuum tube amplifiers give distorsion



The loop transfer function

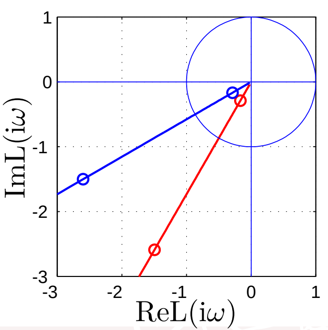
$$L(s) = \Big(\frac{s}{\omega_{gc}}\Big)^n$$

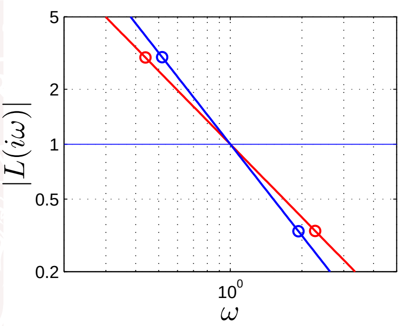
gives a phase margin that is invariant to gain variations.

The slope $n = -1.5$ gives the phase margin $\varphi_m = 45°$.

Horowitz extended Bode's ideas to deal with arbitrary plant variations not just gain variations in the QFT method.

Trade-offs





- Blue curve slope $n = -5/3$ phase margin $\varphi_m = 30^\circ$
- Red curve slope $n = -4/3$ phase margin $\varphi_m = 60^\circ$
- Making the curve steeper reduces the frequency range where compensation is required but the phase margin is smaller

A Fractional PID controller - A Current Fad

Consider the process

$$P(s) = \frac{1}{s(s+1)}$$

Find a controller that gives $L(s) = s^{-1.5}$, hence

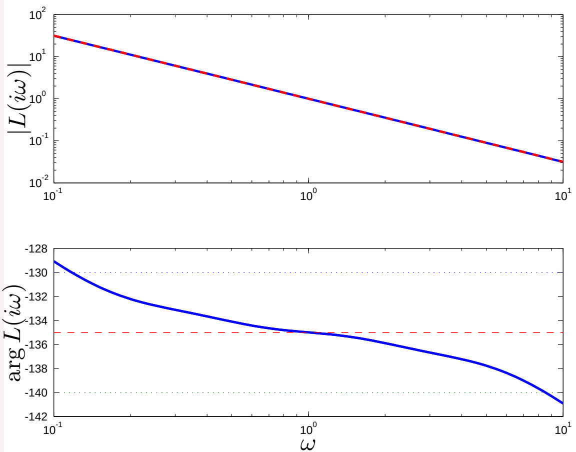
$$C(s) = \frac{L(s)}{P(s)} = \frac{s(s+1)}{s\sqrt{s}} = \sqrt{s} + \frac{1}{\sqrt{s}}$$

A controller with fractional transfer function. To implement it we approximate by a rational transfer function

$$\hat{C}(s) = k\frac{(s+1/16)(s+1/4)(s+1)(s+4)(s+16)}{(s+1/32)(s+1/8)(s+1/2)(s+2)(s+8)(s+32)}$$

High controller order gives robustness

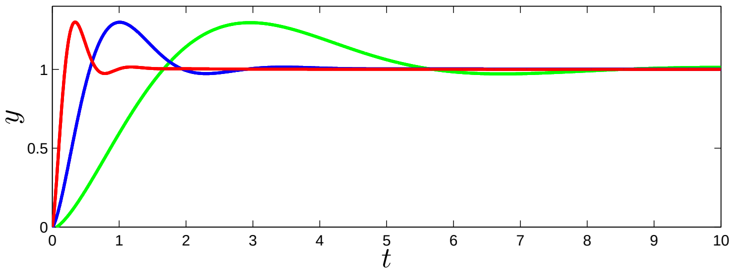
A Fractional Transfer Function



The phase margin changes only by 5° when the process gain varies in the range 0.03-30! Horowitz QFT is a generalization.

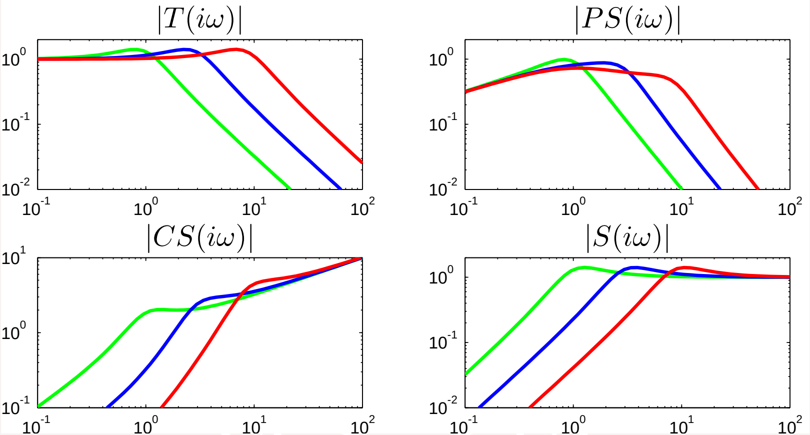
Time Responses

$$P(s) = \frac{k}{s(s+1)}, \quad L(s) = \frac{k}{s\sqrt{s}} \quad C = \sqrt{s} + \frac{1}{s\sqrt{s}}, \quad k = 1, 5, 25,$$



Notice signal shape independent in spite of 25 to 1 gain variations

Fractional System Gain Curves GOF



$$P = \frac{k}{s(s+1)}, \quad k = 1, \quad k = 5, \quad k = 25, \quad C = \sqrt{s} + \frac{1}{\sqrt{s}}$$

Control System Design - Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Requirements

Large signal behavior

- Level and rate limitations in actuators

Small signal behavior

- Sensor noise
- Resolution of AD and DA converters
- Friction

Dynamics

- Minimum phase dynamics do not give limitations

The essential limitation on loopshaping for systems with minimum phase dynamics are due to actuation power, measurement noise and model uncertainty.

Controllers for Minimum Phase Systems

The controller transfer function is given by

$$C(s) = \frac{L_{desired}(s)}{P(s)}, \qquad |C(i\omega_{gc})| = \frac{1}{|P(i\omega_{gc})|}$$

Since $|P(i\omega)|$ typically decays for large frequencies, large ω_{gc} requires high controller gain. The gain of $C(s)$ may also increase after ω_{gc} if phase advance is required. The achievable gain crossover frequency is limited by

- Actuation power and limitations
- Sensor noise
- Process variations and uncertainty

One way to capture this quantitatively is to determine the largest high frequency gain of the controller as a function of the gain crossover frequency ω_{gc}. High gain is a cost of feedback (phase advance).

Gain of a Simple Lead Networks

$$G_n(s) = \left(\frac{s+a}{s/\sqrt[n]{k}+a}\right)^n, \qquad G_\infty(s) = k^{\frac{s}{s+a}}$$

Phase lead: $\varphi_n = n \arctan \dfrac{\sqrt[n]{k}-1}{2\sqrt[2n]{k}}, \quad \varphi_\infty = \dfrac{1}{2}\log k,$

$G_\infty(s) = e^{\frac{2\varphi s}{s+a}}$

Maximum gain for a given phase lead φ:

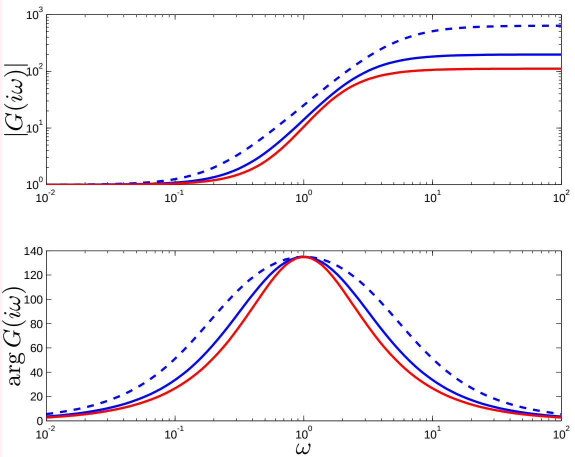
$k_n = \left(1 + 2\tan^2\frac{\varphi}{n} + 2\tan\frac{\varphi}{n}\sqrt{1+\tan^2\frac{\varphi}{n}}\ \right)^n, \ k_\infty = e^{2\varphi}$

Phase lead	*n=2*	*n=4*	*n=6*	*n=8*	*n=∞*
90°	34	25	24	24	23
180°	-	1150	730	630	540
225°	-	14000	4800	3300	2600

Same phase lead with significantly less gain if order is high!

High order controllers can be useful

Lead Networks of Order 2, 3 and ∞



Increasing the order reduces the gain significantly without reducing the width of the peak too much

Bode’s Phase Area Formula

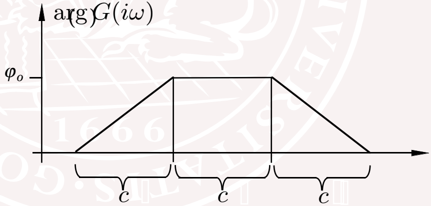
Let $G(s)$ be a transfer function with no poles and zeros in the right half plane. Assume that $\lim_{s\to\infty} G(s) = G_\infty$. Then

$$\log \frac{G_\infty}{G(0)} = \frac{2}{\pi}\int_0^\infty \arg G(i\omega)\frac{d\omega}{\omega} = \frac{2}{\pi}\int_{-\infty}^\infty \arg G(i\omega) d\log\omega$$

The gain K required to obtain a given phase lead φ is an exponential function of the area under the phase curve in the Bode plot

$$k = e^{4c\varphi_0/\pi} = e^{2\gamma\varphi_0}$$

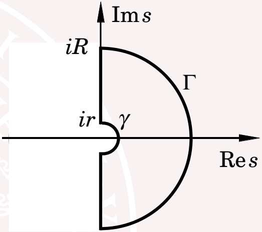
$$\gamma = \frac{2c}{\pi}$$



Proof

Integrate the function

$$\frac{\log G(s)/G(\infty)}{s}$$



around the contour, $\arg G(i\omega)/\omega$ even fcn

$$0 = \int_{-\infty}^{0} \Big(\log \frac{|G(\omega)|}{|G(\infty)|} + i \arg \frac{G(\omega)}{G(\infty)}\Big)\frac{d\omega}{\omega} +$$

$$\int_{0}^{-\infty} \left(\log \frac{|G(\omega)|}{|G(\infty)|} + i \arg \frac{G(\omega)}{G(\infty)}\right)\frac{d\omega}{\omega} + i\pi \log \frac{|G(0)|}{|G(\infty)|}$$

Hence

$$\log \frac{|G(0)|}{|G(\infty)|} = \frac{2}{\pi}\int_0^{\infty} \arg G(i\omega)\, d\log\omega$$

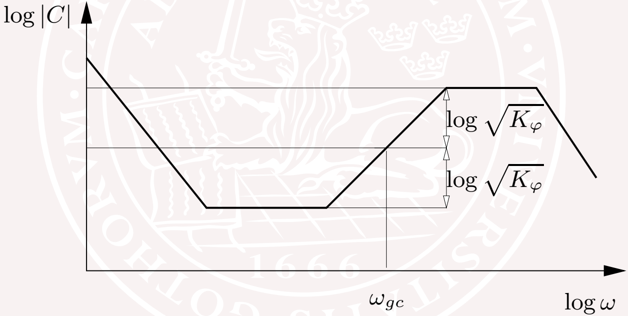
Estimating High Frequency Controller Gain 1

Required phase lead at the crossover frequency

$$\varphi_l = \max(0, -\pi + \varphi_m - \arg P(i\omega_{gc}))$$

Bode's phase area formula gives a gain increase of $K_\varphi = e^{2\gamma\varphi_l}$

Cross-over condition: $|P(i\omega_{gc})C(i\omega_{gc})| = 1$



$$K_c = \max_{\omega \geq \omega_{gc}} |C(i\omega)| = \frac{\sqrt{K_\varphi}}{|P(i\omega_{gc})|} = \frac{\max(1, e^{\gamma(-\pi+\varphi_m-\arg P(i\omega_{gc}))})}{|P(i\omega_{gc})|} = \frac{e^{\gamma\varphi_l}}{|P(i\omega_{gc})|}$$

Estimating High Frequency Controller Gain 2

$$\frac{C}{1+PC} = CS \approx C$$

The largest high frequency gain of the controller is approximately given by ($\gamma \approx 1$)

$$K_c = \max_{\omega \geq \omega_{gc}} |C(i\omega)| = \frac{e^{\gamma\varphi_l}}{|P(i\omega_{gc})|} = \frac{\max(1, e^{\gamma(-\pi+\varphi_m-\arg P(i\omega_{gc}))})}{|P(i\omega_{gc})|}$$

Notice that K_c only depends on the process

- Compensation for process gain $1/|P(i\omega_{gc})|$
- Compensation for phase lag: $e^{\gamma\varphi_l} = e^{\gamma(-\pi+\varphi_m-\arg P(i\omega_{gc}))}$

The largest allowable gain is determined by sensor noise and resolution and saturation levels of the actuator. Results also hold for NMP systems but there are other limitations for such systems

Example - Two and Eight Lags $P = (s+1)^{-n}$

$$K_c = \frac{1}{|P(i\omega_{gc})|} e^{\gamma(-\pi+\varphi_m - \arg P(i\omega_{gc}))} = \left(1+\omega_{gc}^2\right)^{n/2} e^{\gamma(n \arctan \omega_{gc} - \pi + \varphi_m)}$$

$$\gamma = 1, \quad \varphi_m = \frac{\pi}{4}, \quad n = 2, \quad n = 8$$

ω_{gc}	10	20	50	100	200
K_c	181.5	796	5.3×10^3	2.2×10^4	8.7×10^4
φ_l	33.6	39.3	42.7	43.8	44.4
$-\arg P(i\omega_{gc})$	168	174	178	179	179

ω_{gc}	0.5	1.0	1.2	1.4	1.5
K_c	9.4	812	3.7×10^3	1.5×10^4	2.7×10^4
φ_l	78	225	266	300	315
$-\arg P(i\omega_{gc})$	212	360	401	435	450

Summary of Non-minimum Phase Systems

Non-minimum phase systems are easy to control. High performance can be achieved by using high controller gains. The main limitations are given by actuation power, sensor noise and model uncertainty.

$$\frac{PC}{1+PC} = T \qquad C = \frac{T}{P(1-T)} = \frac{L}{P}$$

The high frequency gain of the controller can be estimated by ($\gamma \approx 1$)

$$K_c = \max_{\omega \geq \omega_{gc}} |C(i\omega)| = \frac{e^{\gamma\varphi_l}}{|P(i\omega_{gc})|} = \frac{e^{\gamma(-\pi+\varphi_m-\arg P(i\omega_{gc}))}}{|P(i\omega_{gc})|}$$

Notice that K_c only depends on the process; two factors:

- Compensation for process gain $1/|P(i\omega_{gc})|$
- Gain required for phase lead: $e^{\gamma(-\pi+\varphi_m-\arg P(i\omega_{gc}))}$

Control System Design - Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Requirements

Large signal behavior

- Level and rate limitations in actuators

Small signal behavior

- Sensor noise
- Resolution of AD and DA converters
- Friction

Dynamics

- Non-minimum phase dynamics limit the achievable bandwidth
- Non-minimum phase dynamics give severe limitations

  Right half plane zeros
  Right half plane poles (instabilities)
  Time delays

Non-minimum Phase Systems

Dynamics pose severe limitations on achievable performance for systems with poles and zeros in the right half plane

- Right half plane poles
- Right half plane zeros
- Time delays

Bode introduced the concept *non-minimum phase* to capture this. A system is *minimum phase* system if all its poles and zeros are in the left half plane.

Theme: Capture limitations due to NMP dynamics quantitatively

Bode's Relations between Gain and Phase

There is a unique relation between gain and phase for a transfer function with no poles and zeros in the right half plane.

$$\arg G(i\omega_0) = \frac{2\omega_0}{\pi}\int_0^\infty \frac{\log|G(i\omega)| - \log|G(i\omega_0)|}{\omega^2 - \omega_0^2} d\omega$$

$$= \frac{1}{\pi}\int_0^\infty \frac{d\log|G(i\omega)|}{d\log\omega} \log\Big|\frac{\omega+\omega_0}{\omega-\omega_0}\Big| d\omega \approx \frac{\pi}{2}\frac{d\log|G(i\omega)|}{d\log\omega}$$

$$\frac{\log|G(i\omega)|}{\log|G(i\omega_0)|} = -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{\omega^{-1}\arg G(i\omega) - \omega_0^{-1}\arg G(i\omega_0)}{\omega^2-\omega_0^2} d\omega$$

$$= -\frac{2\omega_0^2}{\pi}\int_0^\infty \frac{d(\omega^{-1}\arg G(i\omega))}{d\omega}\log\Big|\frac{\omega+\omega_0}{\omega-\omega_0}\Big| d\omega$$

Transfer functions with poles and zeros in the right half plane have larger phase lags for the same gain. Factor process transfer function as

$$G(s) = G_{mp}(s)G_{nmp}(s), \qquad |G_{nmp}(i\omega)| = 1, \qquad \angle G_{nmp}(i\omega) < 0$$

Normalized NMP Factors 1

Factor process transfer function as $P(s) = P_{mp}(s)P_{nmp}(s)$, $|P_{nmp}(i\omega)| = 1$ and $P_{nmp}(i\omega)$ negative phase.

Right half plane zero $z = 1$
ω_{gc} not too large

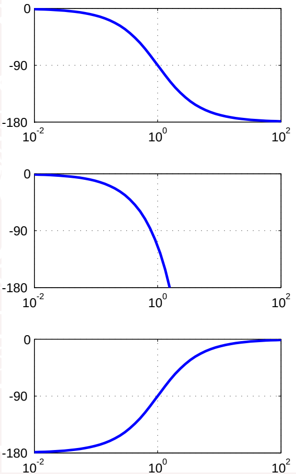
$$P_{nmp}(s) = \frac{1-s}{1+s}$$

Time delay $L = 2$
ω_{gc} not too large

$$P_{nmp}(s) = e^{-2s}$$

Right half plane pole $p = 1$
ω_{gc} must be large

$$P_{nmp}(s) = \frac{s+1}{s-1}$$



Normalized NMP Factors 2

Factor process transfer function as $P(s) = P_{mp}(s)P_{nmp}(s)$, $|P_{nmp}(i\omega)| = 1$ and $P_{nmp}(i\omega)$ negative phase.

RHP pole zero pair $z > p$
OK if you pick ω_{gc} properly

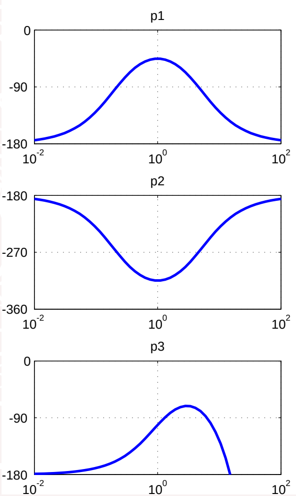
$$P_{nmp}(s) = \frac{(5-s)(s+1/5)}{(5+s)(s-1/5)}$$

RHP pole-zero pair $z < p$
Impossible with stable C

$$P_{nmp}(s) = \frac{(1/5-s)(s+5)}{(1/5+s)(s-5)}$$

RHP pole and time delay
OK if you pick ω_{gc} properly

$$P_{nmp}(s) = \frac{1+s}{1-s}\, e^{-0.2s}$$



Examples of P_{nmp}

Factor process transfer function as $P(s) = P_{mp}(s)P_{nmp}(s)$ such that each non-minimum phase factor is all-pass and has negative phase

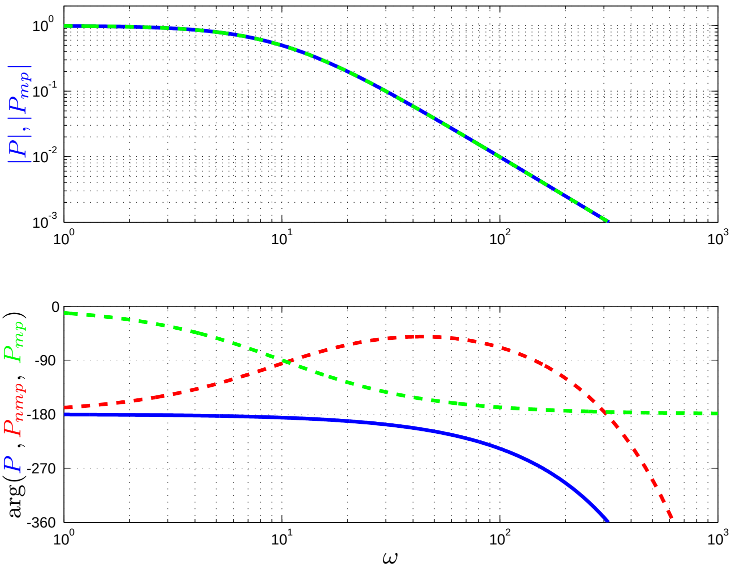
$$P(s) = \frac{1-s}{(s+2)(s+3)} = \frac{1}{(s+1)(s+2)(s+3)} \times \frac{1-s}{1+s}, \qquad P_{nmp}(s) = \frac{1-s}{1+s}$$

$$P(s) = \frac{s+3}{(s-1))(s+2)} = \frac{s+3}{(s+1)(s+2)} \times \frac{s+1}{s-1}, \qquad P_{nmp}(s) = \frac{s+1}{s-1}$$

$$P(s) = \frac{1}{s+1} \times e^{-s}, \qquad P_{nmp}(s) = e^{-s}$$

$$P(s) = \frac{s-1}{(s-2)(s+3)} = -\frac{s+1}{(s+2)(s+3)} \times \frac{1-s}{1+s}\frac{s+2}{s-2}, \quad P_{nmp} = \frac{1-s}{1+s}\frac{s+2}{s-2}$$

Bode Plots Should Look Like This



The Phase-Crossover Inequality

Assume that the controller C has no poles and zeros in the RHP, factor process transfer function as $P(s) = P_{mp}(s)P_{nmp}(s)$ such that $|P_{nmp}(i\omega)| = 1$ and P_{nmp} has negative phase. Requiring a phase margin φ_m we get

$$\begin{aligned}\arg L(i\omega_{gc}) &= \arg P_{nmp}(i\omega_{gc}) + \arg P_{mp}(i\omega_{gc}) + \arg C(i\omega_{gc}) \\ &\geq -\pi + \varphi_m\end{aligned}$$

Approximate $\arg\left(P_{mp}(i\omega_{gc})C(i\omega_{gc})\right) \approx n_{gc}\pi/2$ gives

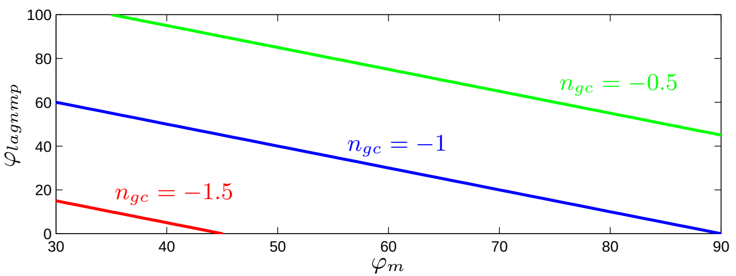
$$\arg P_{nmp}(i\omega_{gc}) \geq -\varphi_{lagnmp}$$

$$\varphi_{lagnmp} = \pi - \varphi_m + n_{gc}\frac{\pi}{2}$$

This inequality is called, the phase crossover inequality. Equality holds if $P_{mp}C$ is Bode's ideal loop transfer function, the expression is an approximation for other designs if n_{gc} is the slope of the gain curve at the crossover frequency.

Reasonable Values of φ_{nmplag}

Admissible phase lag of non-minimum phase factor P_{nmp} as a function of the phase margin φ_m and the slope n_{gc} (roll-off) at the gain crossover frequency



- $\varphi_m = \frac{\pi}{6}, n_{gc} = -\frac{1}{2}$ give $\varphi_{lagnmp} = \frac{7\pi}{12} = 1.83\ (105°)$
- $\varphi_m = \frac{\pi}{4}, n_{gc} = -\frac{1}{2}$ give $\varphi_{lagnmp} = \frac{\pi}{2}\ (90°)$
- $\varphi_m = \frac{\pi}{3}, n_{gc} = -1$ give $\varphi_{lagnmp} = \frac{\pi}{6} = 0.52\ (30°)$
- $\varphi_m = \frac{\pi}{4}, n_{gc} = -1.5$ give $\varphi_{lagnmp} = 0$

Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

System with RHP Zero

$$P_{nmp}(s) = \frac{z-s}{z+s}$$

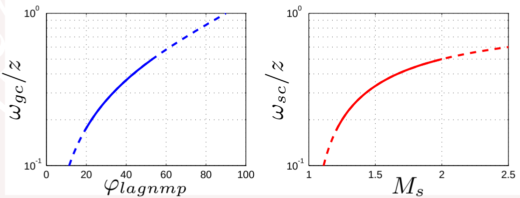
Cross over frequency inequality

$$\arg P_{nmp}(i\omega_{gc}) = -2\arctan\frac{\omega_{gc}}{z} \geq -\pi + \varphi_m - n_{gc}\frac{\pi}{2} = -\varphi_{lagnmp}$$

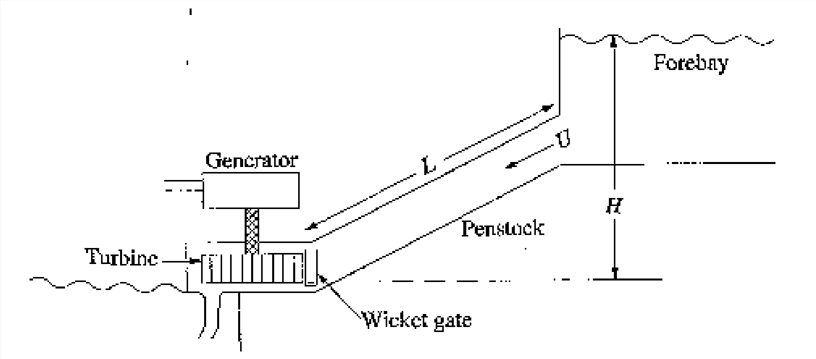
$$\frac{\omega_{gc}}{z} \leq \tan(\frac{\pi}{2} - \frac{\varphi_m}{2} + n_{gc}\frac{\pi}{4}) = \tan\frac{\varphi_{lagnmp}}{2}$$

Compare with inequality for ω_{sc} in Requirements Lecture

$$\frac{\omega_{sc}}{z} < \frac{M_s - 1}{M_s}$$



Water Turbine

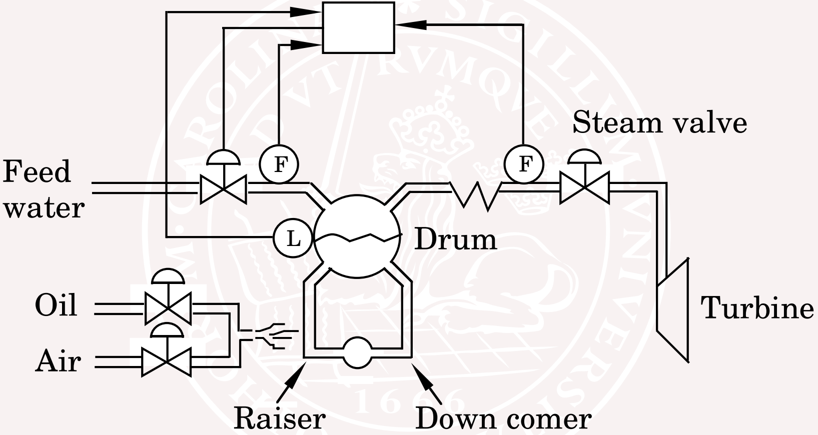


Transfer function from valve opening to power, (T = time for water to flow through penstock)

$$G_{PA} = \frac{P_0}{u_0} \frac{1 - 2u_0 sT}{1 + u_0 sT}$$

A first principles physics model is available in kjå Reglerteori 1968 sid 75-76

Drum Level Control



The shrink and swell effect: steam valve opening to drum level

System with Time Delay

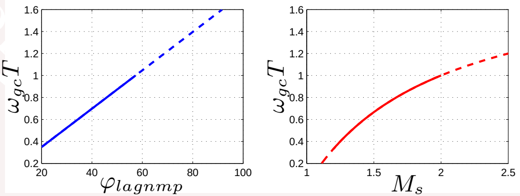
$$P_{nmp}(s) = e^{-sT} \approx \frac{1 - sT/2}{1 + sT/2}$$

Cross over frequency inequality

$$\omega_{gc}T \leq \pi - \varphi_m + n_{gc}\frac{\pi}{2} = \varphi_{lagnmp}$$

The simple rule of ($\varphi_{lagnmp} = \pi/4$) gives $\omega_{gc}T \leq \dfrac{\pi}{4} = 0.8$. Pade approximation gives the zero at $z = \frac{1}{2T}$ using the inequality for RHP zero gives similar result. Comp inequality in Requirements lecture

$$\omega_{sc}T < 2\frac{M_s - 1}{M_s}$$



System with RHP Pole

$$P_{nmp}(s) = \frac{s+p}{s-p}$$

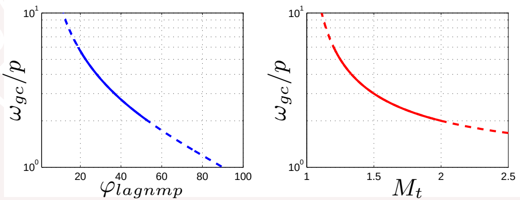
Cross over frequency inequality

$$-2\arctan\frac{p}{\omega_{gc}} \geq -\pi + \varphi_m - n_{gc}\frac{\pi}{2} = -\varphi_{lagnmp}$$

$$\frac{\omega_{gc}}{p} \geq \frac{1}{\tan\varphi_{lagnmp}/2}$$

Compare with inequality for ω_{tc} in Requirements lecture

$$\frac{\omega_{tc}}{p} \geq \frac{M_t}{M_t - 1}$$

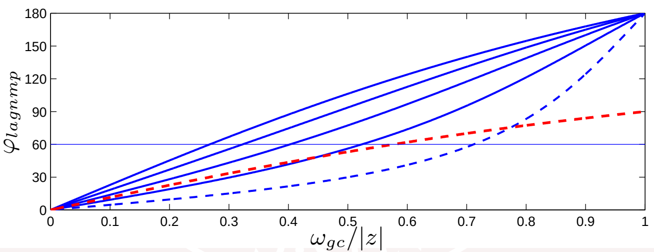


System with complex RHP Zero

$$P_{nmp} = \frac{(x + i\,y - s)(x - i\,y - s)}{(x + i\,y + s)(x - i\,y + s)}$$

$$\varphi_{lagnmp} = 2\arctan\frac{y+\omega}{x} - 2\arctan\frac{y-\omega}{x}$$

$$= 2\arctan\frac{2\omega x}{x^2+y^2-\omega^2} = 2\arctan\frac{2\omega|z|\zeta}{|z|^2-\omega^2}$$



Damping ratio $\zeta = 0.2$ (dashed), 0.4, 0.6. 0.8 and 1.0, red dashed curve single RHP zero. Small ζ easier to control.

System with RHP Pole and Zero Pair

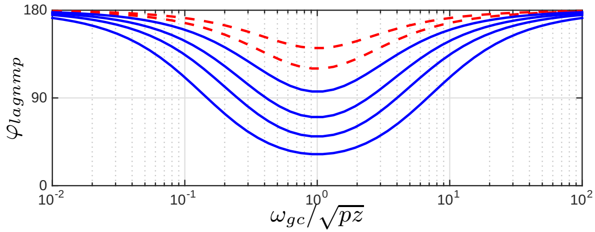
$$P_{nmp}(s) = \frac{(z-s)(s+p)}{(z+s)(s-p)}, \qquad M_s > \frac{z+p}{z-p}$$

Cross over frequency inequality for $z > p$

$$-2\arctan\frac{\omega_{gc}}{z} - 2\arctan\frac{p}{\omega_{gc}} \geq -\varphi_{lagamp}, \quad \frac{\omega_{gc}}{z} + \frac{p}{\omega_{gc}} \leq \left(1 - \frac{p}{z}\right)\tan\frac{\varphi_{lagamp}}{2}$$

The smallest value of the left hand side is $2\sqrt{p/z}$, which is achieved for $\omega_{gc} = \sqrt{pz}$, hence $\varphi_{lagnmp} = 2\arctan\left(2\sqrt{pz}/(z-p)\right)$

Plot of φ_{lagnmp} for $\frac{z}{p}$ =2, 3, 5, 10, 20, 50 and M_s =3, 2, 1.5, 1.2, 1.1, 1.05



An Example

From Doyle, Francis Tannenbaum: Feedback Control Theory 1992.

$$P(s) = \frac{s-1}{s^2+0.5s-0.5}, \qquad P_{nmp} = \frac{(1-s)(s+0.5)}{(1+s)(s-0.5)}$$

Keel and Bhattacharyya Robust, Fragile or Optimal AC-42(1997) 1098-1105: In this paper we show by examples that optimum and robust controllers, designed by the H_2, H_∞, L_1 and μ formulations, can produce extremely fragile controllers, in the sense that vanishingly small perturbations of the coefficients of the designed controller destabilize the closed loop system. The examples show that this fragility usually manifests itself as extremely poor gain and phase margins of the closed loop system.

- Pole at $s = 0.5$, zero at $s = 1$, $\varphi_{lagnmp} = 2.46\,(141^\circ)$,
  $M_s > (z+p)/(z-p) = 3$,
  $\varphi_m \approx 2\arcsin(1/(2M_s)) = 0.33\,(19^\circ)$
- Hopeless to control robustly
- You don't need any more calculations

Example - The X-29

Advanced experimental aircraft. Many design efforts with many methods and high cost.

Requirements $\varphi_m = 45^\circ$ could not be met. Here is why! Process has RHP pole $p = 6$ and RHP zero $z = 26$. Non-minimum phase factor of transfer function

$$P_{nmp}(s) = \frac{(s+26)(6-s)}{(s-26)(6+s)}$$

The smallest phaselag $\varphi_{lagnmp} = 2.46(141^\circ)$ of P_{nmp} is too large. The zero pole ratio is $z/p = 4.33$ gives $M_s > \frac{z+p}{z-p} = 1.6$ $\varphi_m \approx 2\arcsin(\frac{1}{2Ms}) = 0.64(36^\circ)$. Not possible to get a phase margin of 45°!

Bicycle with Rear Wheel Steering

Richard Klein at UIUC has built several UnRidable Bicycles (URBs). There are versions in Lund and UCSB.

Transfer function

$$P(s) = \frac{am\ell V_0}{bJ} \frac{-s + \dfrac{V_0}{a}}{s^2 - \dfrac{mg\ell}{J}}$$

Pole at $p = \sqrt{\dfrac{mg\ell}{J}} \approx 3 \ \mathrm{rad/s}$

RHP zero at $z = \dfrac{V_0}{a}$

Pole independent of velocity but zero proportional to velocity. There is a velocity such that $z = p$ and the system is uncontrollable. The system is difficult to control robustly if z/p is in the range of 0.25 to 4.

RHP Pole and Time Delay

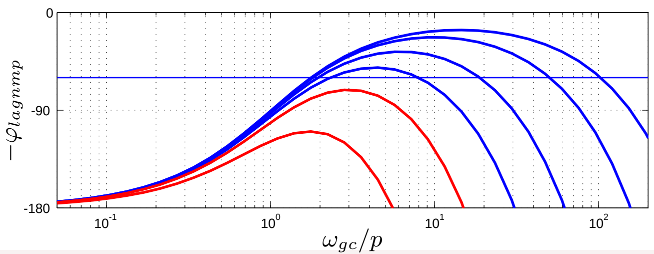
NMP part of process transfer function

$$P_{nmp}(s) = \frac{s+p}{s-p} e^{-sL}, \qquad M_s > e^{pL} \qquad pL < 2$$

$$\arg P_{nmp}(i\omega_{gc}) = -2 \arctan \frac{p}{\omega_{gc}} - \omega_{gc} L > -\varphi_{lagnmp}$$

$$\varphi_{lagnmp} = \pi - \varphi_m + n_{gc} \frac{\pi}{2}$$

Plot of φ_{lagnmp} for $pL = 0.01, 0.02, 0.05, 0.1, 0.2, 0.7$

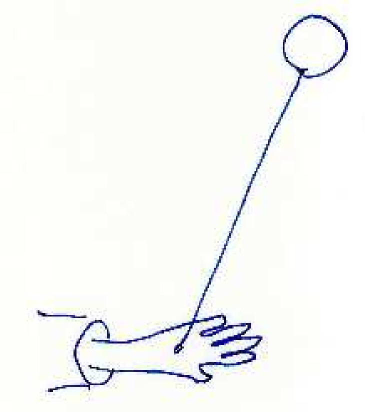


Stabilizing an Inverted Pendulum with Delay

Right half plane pole at

$$p = \sqrt{g/\ell}$$

With a neural lag of 0.07 s and the robustness condition $pL < 0.3$ we find $\ell > 0.5$.



A vision based system with sampling rate of 50 Hz (a time delay of 0.02 s) and $pL < 0.3$ shows that the pendulum can be robustly stabilized if $\ell > 0.04$ m.

Dynamics Limitations for NMP Systems - Part 1

For controllers with no poles in the RHP we have

- A RHP zero z gives an upper bound on the bandwidth:

$$\frac{\omega_{gc}}{a} < \tan\frac{\varphi_{lagnmp}}{2}, \qquad \frac{\omega_{sc}}{a} < \frac{M_s - 1}{M_s}$$

$$\varphi_{lagnmp} = \pi - \varphi_m + n_{gc}\frac{\pi}{2}$$

- A time delay L gives a upper bound on the bandwidth:

$$\omega_{gc}L < \varphi_{lagnmp}, \qquad \omega_{sc}L < 2\frac{M_s - 1}{M_s}$$

- A RHP pole p gives a lower bound on the bandwidth:

$$\frac{\omega_{gc}}{p} > \frac{1}{\tan\frac{\varphi_{lagnmp}}{2}}, \qquad \frac{\omega_{tc}}{p} > \frac{M_t}{M_t - 1}$$

Dynamics Limitations for NMP Systems - Part 2

For controllers with no poles in the RHP we have

- RHP poles and zeros must be sufficiently separated with $z > p$

$$M_s > \frac{z+p}{z-p}, \qquad \varphi_{lagnmp} > \frac{\pi}{3}\,(60^\circ)$$

- A process with a RHP poles zero pair with $p > z$ cannot be controlled robustly with a controller having no poles in the RHP
- The product of a RHP pole and a time delay cannot be too large

$$M_s > e^{pL}, \qquad \varphi_{lagnmp} < \frac{\pi}{3}\,(60^\circ)$$

What about a controller with RHP poles?

Dynamics Limitations - Ball Park Numbers

- A RHP zero z: gives an upper bound to bandwidth: $\frac{\omega_{gc}}{z} < 0.5$
- A double RHP zero: $\frac{\omega_{gc}}{z} < 0.25$
- A time delay L gives an upper bound to bandwidth: $\omega_{gc}L < 1$
- A RHP pole p gives a lower bound to bandwidth: $\frac{\omega_{gc}}{p} > 2$
- A double RHP pole: $\frac{\omega_{gc}}{p} > 4$
- A RHP pole zero pair requires: $\frac{z}{p} > 4$

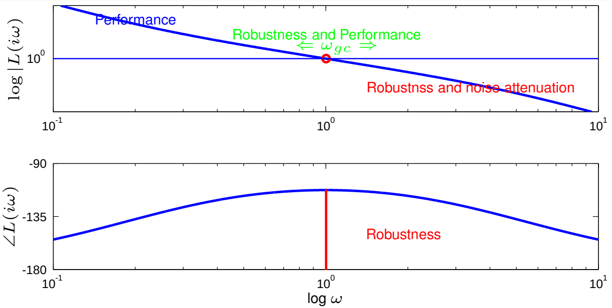
These rules, which are easy to remember, give sensitivities M_s and M_t around 2 and phase lags φ_{lagnmp} of the nonminimum phase factor around $90°$.

Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Back to Bode



- Pick ω_{gc} to achieve desired performance, subject to constraints due to measurement noise and non-minimum phase dynamics
- Add effects of modeling uncertainty (QFT)
- Increase low frequency gain if necessary for tracking and add high frequency roll-off for noise and robustness
- Tweak behavior around crossover to obtain robustness ($\mathcal{H}_\infty$ loopshaping)

The Assessment Plot - Picking ω_{gc}

The *assessment plot* is an attempt to give a gross overview of the properties of a controller and to guide the selection of a suitable gain crossover frequency. It has a gain curve $K_c(\omega_{gc})$ and two phase curves $\arg P(i\omega)$ and $\arg P_{nmp}(i\omega)$

- Attenuation of disturbance captured by ω_{gc}
- Injection of measurement noise captured by the high frequency gain of the controller $K_c(\omega_{gc})$

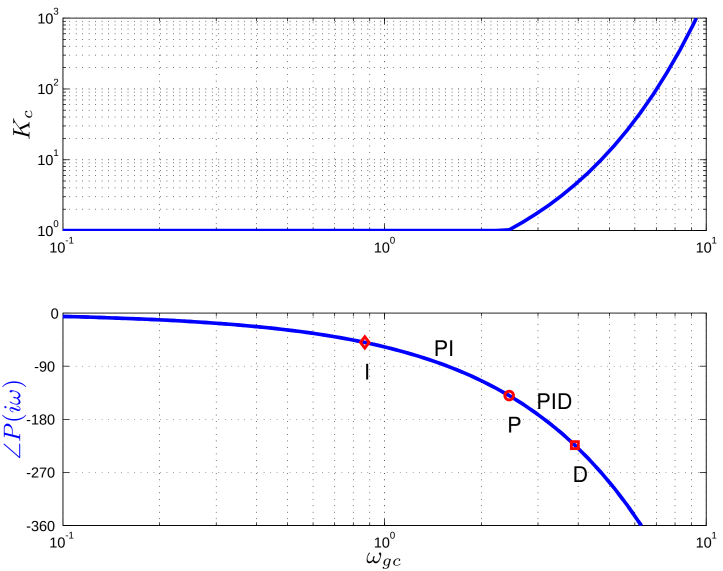
$$K_c(\omega) = \max_{\omega \geq \omega} |C(i\omega)| = \frac{\max(1, e^{\gamma(-\pi+\varphi_m - \arg P(i\omega))})}{|P(i\omega)|}$$

- Robustness limitations due to time delays and RHP poles and zeros captured by conditions on the admissible phaselag of the nonminimum phase factor $0.5 < \varphi_{lagnmp} < 1.5$

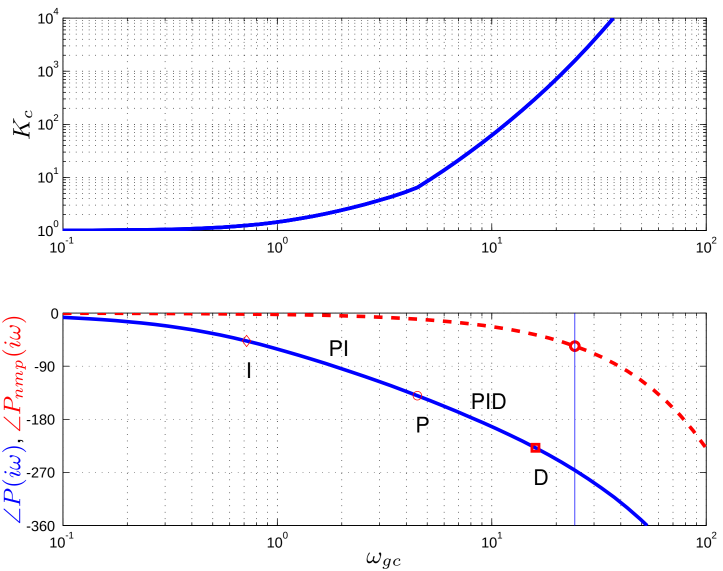
$$\varphi_{lagnmp}(\omega) = -\arg P_{nmp}(i\omega) = \pi - \varphi_m + n_{gc}\frac{\pi}{2}$$

- Controller complexity is captured by $\arg P(i\omega_{gc})$

Assessment Plot for $e^{-\sqrt{s}}$

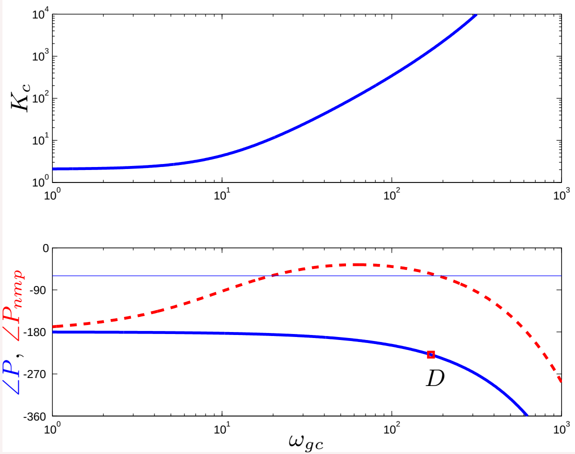


Assessesement Plot - Delay and Spread Lags



Assessment Plot for $P(s) = e^{-0.01s}/(s^2 - 100)$

$$P(s) = \frac{1}{(s+10)^2}\,\frac{s+10}{s-10}e^{-0.01s}, \qquad P_{nmp}(s) = \frac{s+10}{s-10}e^{-0.01s}$$



Loop Shaping

1. Introduction
2. Loop shaping design
3. Bode's ideal loop transfer funtion
4. Minimum phase systems
5. Non-minimum phase systems
6. Fundamental Limitations
7. Performance Assessment
8. Summary

Theme: Shaping Nyquist and Bode Plots

Summary

- A classic design method with focus on the Bode plot
- The concepts of minimum and non-minimum phase
- Fundamental limitations
  - Phase lag φ_{lagnmp} of non-minimum phase factor P_{nmp} cannot be too large ($20^\circ - 60^\circ$)
  - Maximum modulus theorem for S and T
  - The assumption that the controller has no RHP
  - The gain crossover frequency inequality
- Rules of thumb based on approximate expressions
- Assessment plots
- Extensions
  - What replaced the Bode plot for multivariable systems?
  - The idea of zero directions
  - More complicated systems - oscillatory dynamics
  - Process variations QFT