Feedback Fundamentals

Karl Johan Åström

Automatic Control LTH
Lund University

January 13, 2020

Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Gangs of Four and Seven
4. The Sensitivity Functions
5. Fundamental Limitations
6. Summary

Theme: A closer look at feedback

The Magic of Feedback

Good properties:

- Attenuate effects of disturbances - process control, automotive
- Make good systems from bad components - feedback amplifier
- Follow command signals - robotics, automotive
- Stabilize and shape behavior - flight control

Bad properties:

- Feedback may cause instability
- Feedback feeds measurement noise into the system

Arthur C. Clarke: Any sufficiently advanced technology is indistinguishable from magic

The Magic of Integral Action

Consider a system (linear or nonlinear) controlled by a controller having integral action

$$u(t) = \int_0^t e(\tau) d\tau + \ldots.$$

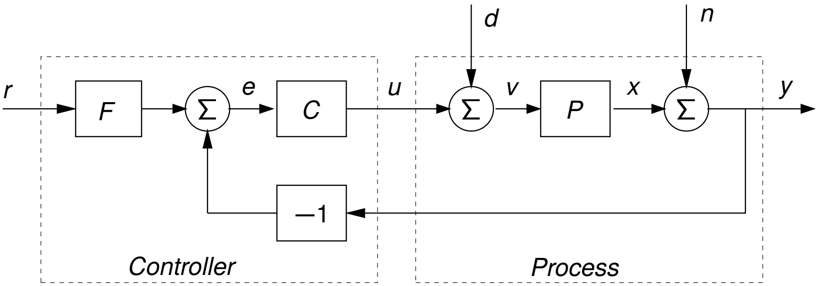
If the closed loop is stable then the equilibrium must be such that the error is zero!

Proof: Assume that the error converges to a nonzero value leads to a contratdction!

Important Issues

- Some important considerations
  - Load disturbances
  - Measurement noise
  - Process variations
  - Uncertainties in modeling
  - Command signal following
- Evaluation, testing and specifications of control systems
- Quantification of performance and robustness
- Measurement and testing of performance and robustness
- Limitations what can and cannot be achieved by feedback

A Basic Feedback Loop

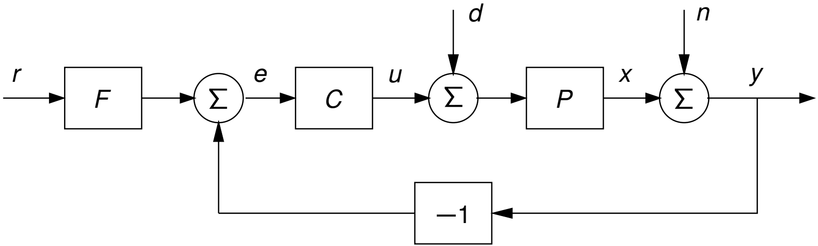


Ingredients:

- Controller: feedback C, feedforward F
- Load disturbance d : Drives the system from desired state - e.g. slope of road
- Process: transfer function P
- Measurement noise n : Corrupts information about x
- Process variable x should follow reference r

Quiz

Look at the block diagram

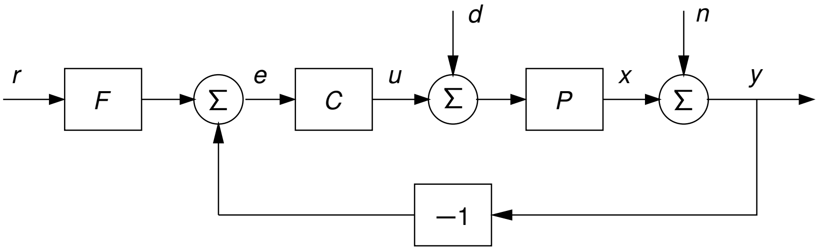


Find transfer function from r to u!

The Audience is Thinking ...

Quiz

Look at the block diagram



Find transfer function from r to u!

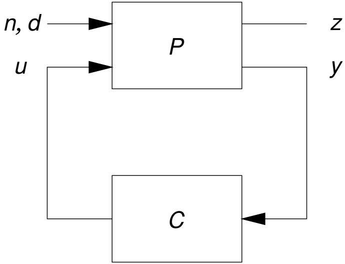
The Audience is Thinking ...

Simple rule: blocks between singnals and $1 + PC$ in denominator!!

$$G_{ur} = \frac{CF}{1 + PC}$$

A More General Setting

Load disturbances and measurement need not enter at the process input and measurement noise not at the output. A more general situation is.



$$w = (d, n, y_{sp}), \;\; z = (e, v), \text{ find } C \text{ to make } z \text{ small!}$$

These problems can be dealt with in the same way but we will stick to the simpler case. Always useful to understand disturbances, who they are and where they enter the system.

Typical Requirements

A controller should

A: Reduce effects of load disturbances

B: Do not inject too much measurement noise into the system

C: Make the closed loop insensitive to variations in the process

D: Make output follow command signals

Performance is expressed by

- Response to command signals
- Attenuation of load disturbances

Robustness is expressed by sensitivity to

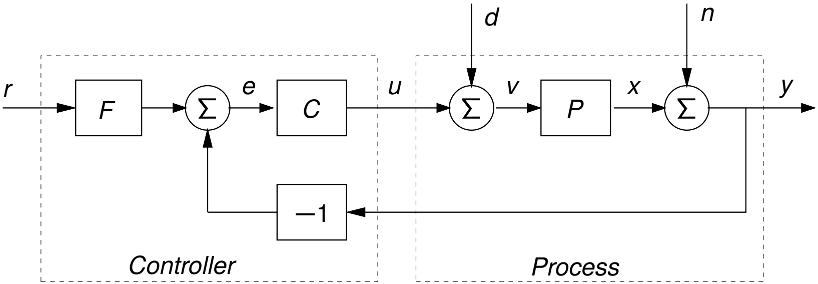
- Load disturbances
- Model uncertainty

Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Gangs of Four and Seven
4. The Sensitivity Functions
5. Fundamental Limitations
6. Summary

Theme: A closer look at feedback

Controller with Two Degrees of Freedom



- Load disturbance d : Drives the system from desired state
- Measurement noise n : Corrupts information about x

The controller has two degrees of freedom 2DOF because the signal transmissions from reference r to control u and from measurement y to control u are different. Horowitz 1963.

A Separation Principle for 2DOF Systems

Design the feedback C to achieve

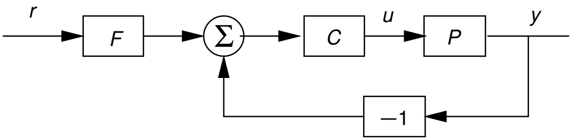
- Low sensitivity to load disturbances d
- Low injection of measurement noise n
- High robustness to process uncertainty and process variations

Design the feedforward F to achieve

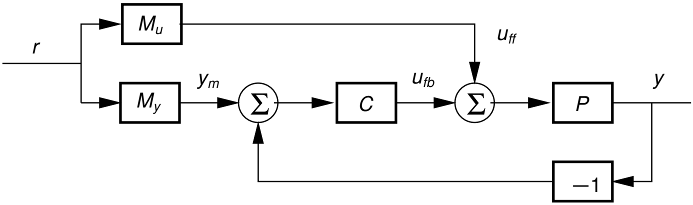
- Desired response to command signals r

Many Versions of 2DOF

Basic loop

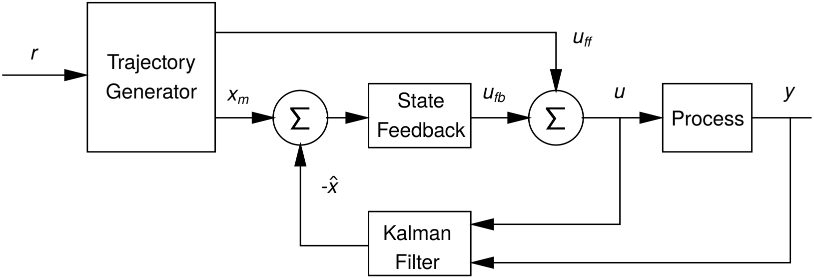


Explicit representation of ideal response y_m and control signal u_{ff}



For linear systems all 2DOF configurations have the same properties. For the systems above we have $CF = M_u + CM_y$

Many Versions of 2DOF - Kalman Filter Architecture

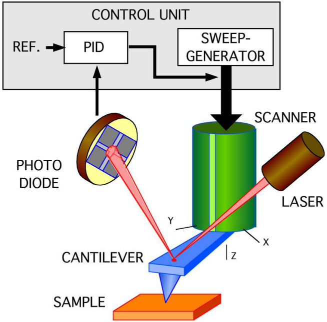


- A nice separation of the different functions
- The signals x_m and u_{ff} can be generated from r in real time or from stored tables

Some Systems only Allow Error Feedback

Disk drive

Atomic Force Microscope



Only error can be measured

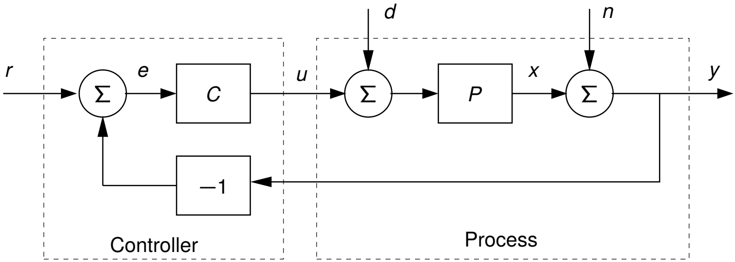
Design for command disturbance attenuation, robustness and command signal response can not be separated!

Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Sensitivity Functions
4. Fundamental Limitations
5. Summary

Theme: A closer look at feedback

System with Error Feedback



Three inputs r, d and n, four interesting signals e, u, x and y!

$$G_{xr} = \frac{PC}{1+PC} \quad G_{xd} = \frac{P}{1+PC} \quad G_{xn} = -\frac{PC}{1+PC}$$

$$G_{yr} = \frac{PC}{1+PC} \quad G_{yd} = \frac{P}{1+PC} \quad G_{yn} = \frac{1}{1+PC}$$

$$G_{ur} = \frac{C}{1+PC} \quad G_{ud} = -\frac{PC}{1+PC} \quad G_{un} = -\frac{C}{1+PC}$$

$$G_{er} = \frac{1}{1+PC} \quad G_{ed} = -\frac{P}{1+PC} \quad G_{en} = -\frac{1}{1+PC}$$

The Gang of Four GOF - controller with error feedback

$$G_{xr} = \frac{PC}{1+PC}, \quad G_{xd} = \frac{P}{1+PC} \quad G_{xn} = -\frac{PC}{1+PC},$$
$$G_{yr} = \frac{PC}{1+PC}, \quad G_{yd} = \frac{P}{1+PC}, \quad G_{yn} = \frac{1}{1+PC},$$
$$G_{ur} = \frac{C}{1+PC}, \quad G_{ud} = -\frac{PC}{1+PC}, \quad G_{un} = -\frac{C}{1+PC},$$
$$G_{er} = \frac{1}{1+PC}, \quad G_{ed} = -\frac{P}{1+PC}, \quad G_{en} = -\frac{1}{1+PC},$$

Only four transfer functions!!! (Sensitivity functions - the Gang of Four)!

$$S = \frac{1}{1+PC}, \quad T = \frac{PC}{1+PC} = 1 - S, \quad PS = \frac{P}{1+PC}, \quad CS = \frac{C}{1+PC}$$

Interpretation of The Gang of Four

Response of output y to load disturbance d is characterized by

$$G_{yd} = \frac{P}{1+PC} = PS = \frac{T}{C}$$

Response of control signal u to measurement noise n is characterized by

$$G_{un} = -\frac{C}{1+PC} = -CS = -\frac{T}{P}$$

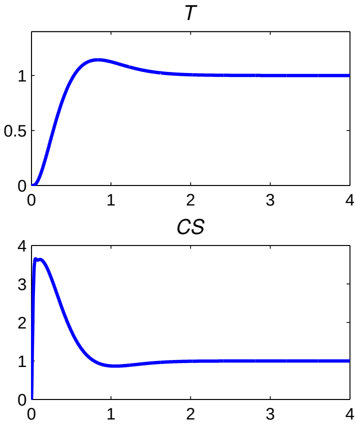
Responses of y and u to reference signal r are characterized by

$$G_{yr} = \frac{PC}{1+PC} = T, \quad G_{ur} = \frac{C}{1+PC} = CS$$

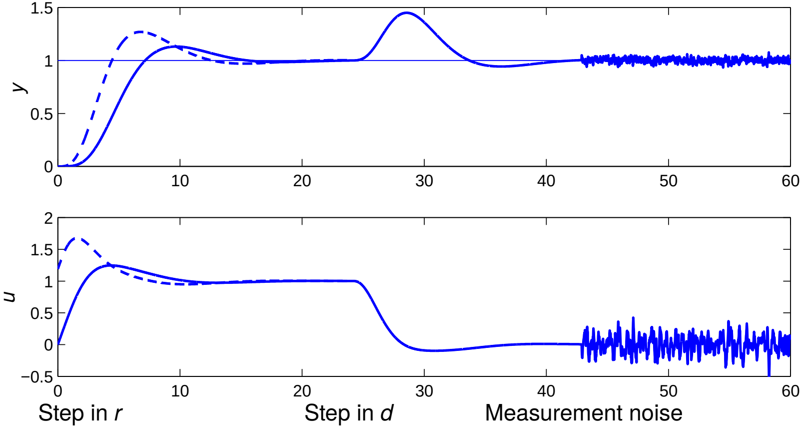
Robustness to process variations is characterized by

$$S = \frac{1}{1+PC}, \quad T = \frac{PC}{1+PC}$$

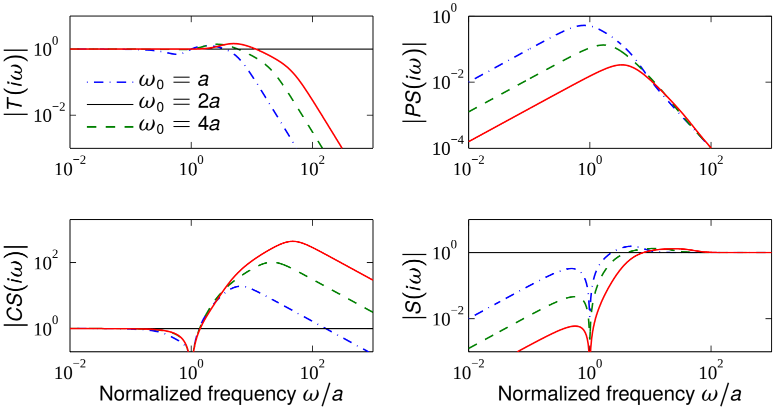
Visualizing the GOF - Time Responses



Another Way to Show Time Responses



Visualizing the GOF - Frequency Responses



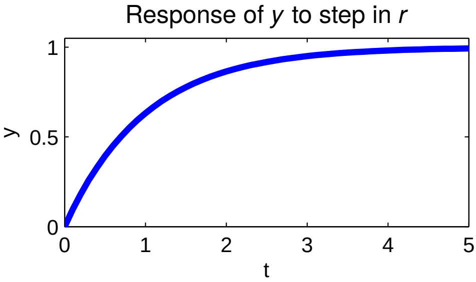
Gain curves for the GOF is a good way to get a quick overview of a feedback system. Curves represent three different controller are designed for a nano-positioner

Discuss!

A Warning!

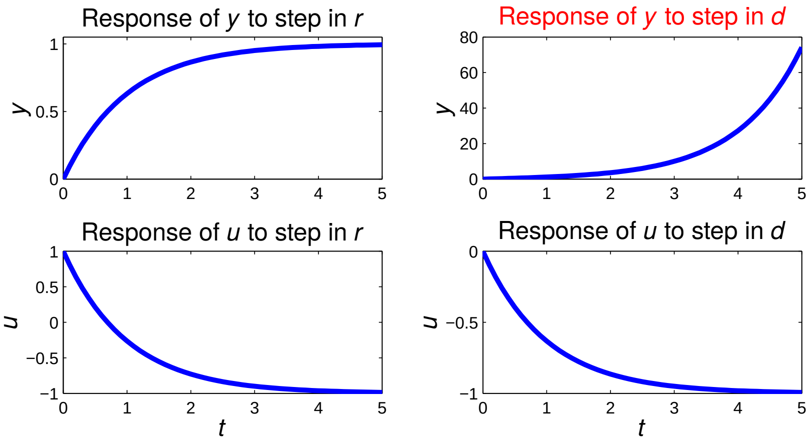
Remember to always look at all responses!

The step response below looks fine



BUT ...

All Four Responses



The system is unstable!

What is going on?

The Gang of Four

Process: $P(s) = \dfrac{1}{s-1}$, Controller: $C(s) = \dfrac{s-1}{s}$

Loop transfer function: $L = PC = \dfrac{1}{s-1} \times \dfrac{s-1}{s} = \dfrac{1}{s}$

Notice cancellation of the factor $s-1$! The Gang of Four

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

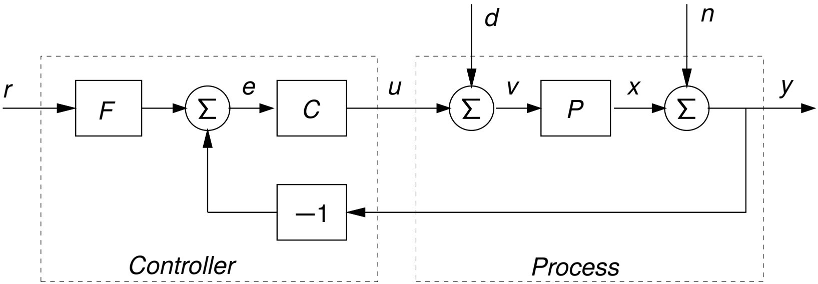
$$\frac{C}{1+PC} = \frac{s-1}{s+1} \qquad \frac{1}{1+PC} = \frac{1}{s+1}$$

Response of y to step in load disturbance d

$$G_{yd}(s) = \frac{P}{1+PC} = \frac{s}{(s+1)(s-1)}$$

This transfer function represents an unstable system

2DOF System General – The Gang of Seven



The Gang of Four and transfer functions from reference r to e, x, y and u

$$G_{er} = \frac{F}{1+PC} \quad G_{xr} = \frac{PCF}{1+PC} \quad G_{yr} = \frac{PCF}{1+PC} \quad G_{ur} = \frac{CF}{1+PC}$$

Some Observations

- To fully understand a system it is necessary to look at all transfer functions
- A system based on error feedback is characterized by *four* transfer functions *The Gang of Four*
- The system with a controller having two degrees of freedom is characterized by *seven* transfer function *The Gang of Seven*
- It may be strongly misleading to only show properties of a few systems for example the response of the output to command signals. A common omission in many papers and books.
- The properties of the different transfer functions can be illustrated by their transient or frequency responses.

Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Gangs of Four and Seven
4. The Sensitivity Functions
5. Fundamental Limitations
6. Summary

Theme: A closer look at feedback

The Sensitivity Functions

- Sensitivity function $S = \dfrac{1}{1+PC} = \dfrac{1}{1+L}$
- Complementary sensitivity function $T = 1 - S = \dfrac{PC}{1+PC} = \dfrac{L}{1+L}$
- Input sensitivity function $G_{xl} = \dfrac{P}{1+PC} = PS$
- Output sensitivity function $G_{un} = \dfrac{C}{1+PC} = CS$

are called sensitivity functions. They have interesting properties and useful physical interpretations. We have

- The functions S and T only depend on the loop transfer function L
- $S + T = 1$
- Typically $S(0)$ small and $S(\infty) = 1$ and consequently $T(0) = 1$ and $T(\infty)$ small

Poles Zeros and Sensitivity Functions

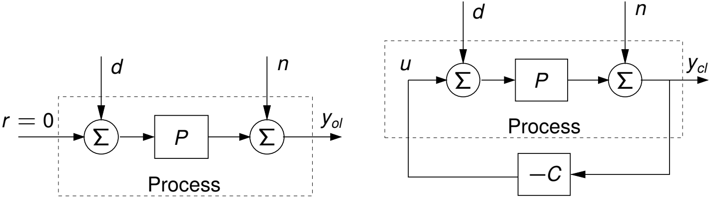
The sensitivity functions depend only on the loop transfer function $L = PC$

$$S = \frac{1}{1+L} \qquad T = \frac{L}{1+L}$$

Notice that

- The sensitivity function S is zero and the complementary sensitivity function is one at the poles of L
- The sensitivity function S is one and complementary sensitivity function T is zero at the zeros of L

Disturbance Attenuation



Output without control $Y = Y_{ol}(s) = N(s) + P(s)D(s)$
Output with feedback control

$$Y_{cl} = \frac{1}{1+PC}\big(N+PD\big) = \frac{1}{1+PC}Y_{ol} = SY_{ol}$$

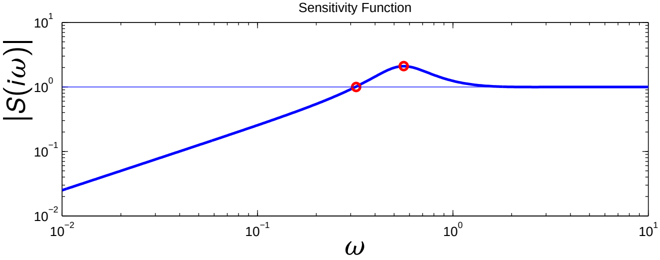
The effect of feedback is thus like sending the open loop output through a system with the transfer function $S = 1/(1+PC)$. Disturbances with frequencies such that $|S(i\omega)| < 1$ are reduced by feedback, disturbances with frequencies such that $|S(i\omega)| > 1$ are amplified by feedback.

Assessment of Disturbance Reduction - Bode

We have

$$\frac{Y_{cl}(s)}{Y_{cl}(s)} = S(s) = \frac{1}{1 + P(s)C(s)}$$

Feedback attenuates disturbances of frequencies ω such that $|S(i\omega)| <$.
It amplifies disturbances of frequencies such that $|S(i\omega)| > 1$

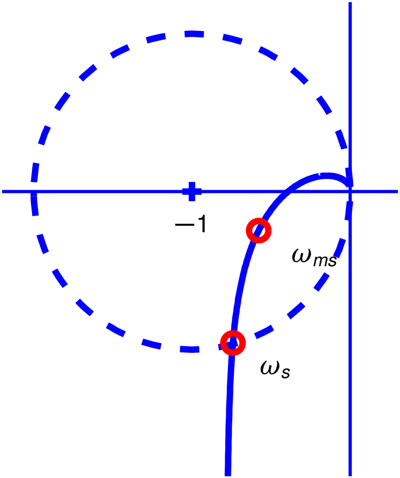


Assessment of Disturbance Reduction - Nyquist

$$\frac{Y_{cl}}{Y_{ol}} = \frac{1}{1 + PC} = S$$

Geometric interpretation: Disturbances with frequencies inside the circle are amplified by feedback. Disturbances with frequencies outside the circle are reduced. Disturbancds with frequencies inside the circle are amplified. Worst amplifiaticon for frequencies closest to the critical point

Disturbances with frequencies less than ω_s are reduced by feedback, those with higher frequencies afre amplified.



Properties of the Sensitivity Function

- Can the sensitivity be small for all frequencies?

  No we have $S(\infty) = 1$!
- Can we have $|S(i\omega)| \leq 1$?

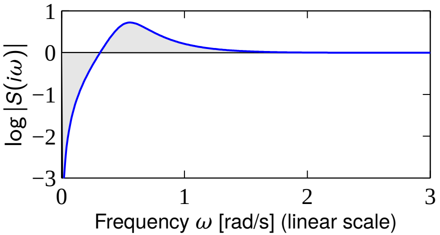
  If the Nyquist curve of $L = PC$ is in the first and third quadrant!

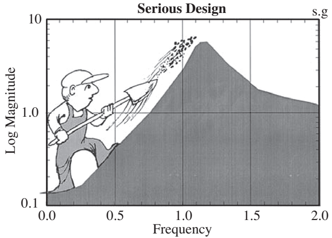
  Passive systems!
- Bode's integral, p_k RHP poles of $L(s)$, z_k RHP zeros of $L(s)$

$$\int_0^\infty \log |S(i\omega)| d\omega = \pi \sum \mathrm{Re}\, p_k - \frac{\pi}{2} \lim_{s\to\infty} sL(s)$$

- Fast poles (and slow zeros) in the RHP are bad!
- Useful to let the loop transfer function go to zero rapidly for high frequencies (*high-frequency roll-off*) because the last term vanishes
- The "water-bed effect". Push the curve down at one frequency and it pops up at another! Design is a compromise!

The Water Bed Effect - Bode's Integral





$$\int_0^\infty \log |S(i\omega)| d\omega = \pi \sum \mathrm{Re}\, p_k - \frac{\pi}{2} \lim_{s\to\infty} sL(s)$$

The sensitivity can be decreased at one frequency at the cost of increasing it at another frequency.

Feedback design is a trade-off

Robustness

Effect of small process changes dP on closed loop response $T = PC/(1 + PC)$

$$\frac{dT}{dP} = \frac{C}{(1 + PC)^2} = \frac{ST}{P}, \qquad \frac{dT}{T} = S\frac{dP}{P}$$

Effect of large process canges: how much ΔP can the process change without making the closed loop unstable?

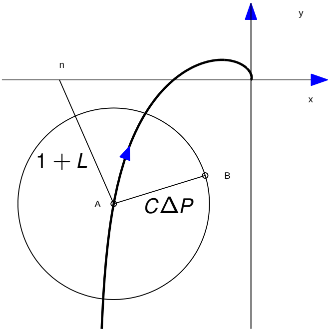
$$|C\Delta P| < |1 + PC|$$

or

$$\frac{|\Delta P|}{|P|} < \frac{1}{|T|}$$

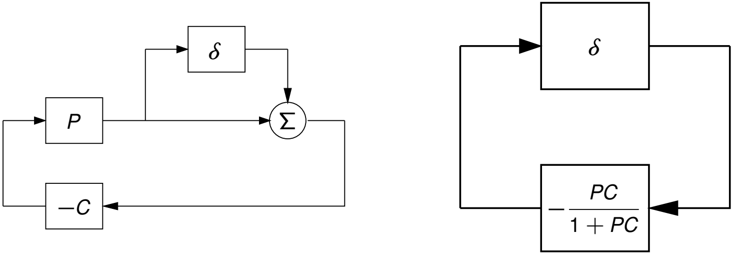
ΔP must be stable

Nyquist plot



Another View of Robustness

A feedback system where the process has multiplicative uncertainty, i.e. $P + \Delta P = P(1 + \delta)$, where $\delta = \Delta P/P$ is the relative error, can be represented with the following block diagrams



The small gain theorem gives the stability condition

$$|\delta| = \frac{|\Delta P|}{|P|} < \left|\frac{1 + PC}{PC}\right| = \frac{1}{|T|}$$

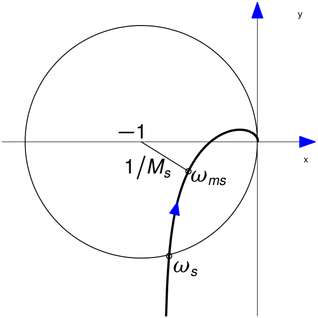
Same result as obtained before!

Robustness and Sensitivity

Gain and phase margins

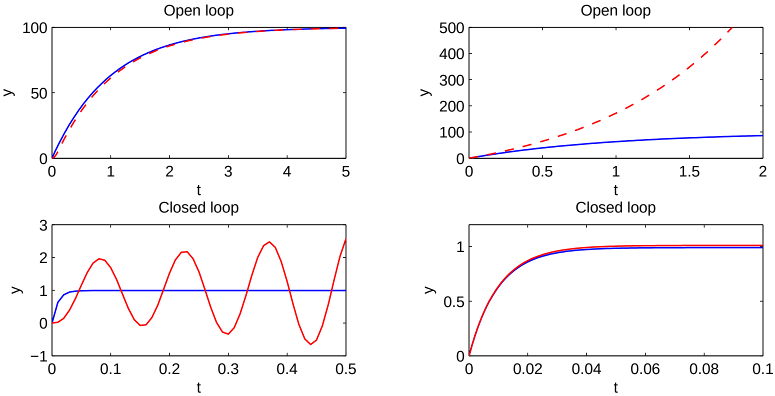
$$g_m \geq \frac{M_s}{M_s - 1}, \quad \varphi_m \geq 2 \arcsin \frac{1}{2M_s}$$

Constraints on both gain and phase margins can be replaced by one constraint on maximum sensitivity M_s.



- $M_s = 2$ guarantees $g_m \geq 2$ and $\varphi_m \geq 30°$
- $M_s = 1.6$ guarantees $g_m \geq 2.7$ and $\varphi_m \geq 36°$
- $M_s = 1.4$ guarantees $g_m \geq 3.5$ and $\varphi_m \geq 42°$
- $M_s = 1$ guarantees $g_m = \infty$ and $\varphi_m \geq 60°$

When are Two Systems Close?

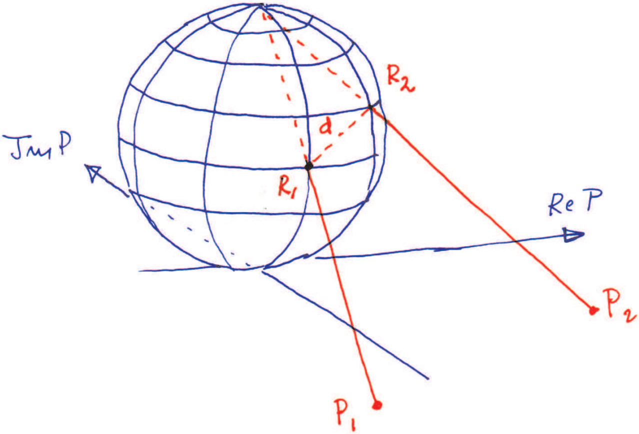


Comparing open loop step responses can be misleading!

Open loop frequency responses are slightly better

Much better to compare closed loop responses

Vinnicombe's Metric



Summary of the Sensitivity Functions

$$S = \frac{1}{1+L}, \quad T = \frac{L}{1+L}, \quad M_s = \max|S(i\omega)|, \quad M_t = \max|T(i\omega)|$$

The value $1/M_s$ is the shortest distance from the Nyquist curve of the loop transfer function $L(i\omega)$ to the critical point -1.

$$S = \frac{\partial \log T}{\partial \log P} = \frac{Y_{cl}(s)}{Y_{ol}(s)}$$

How much can the process be changed with stable ΔP without making the closed loop system unstable?

$$\frac{|\Delta P|}{|P|} < \frac{1}{|T|}$$

Bode's integral the water bed effect.

$$\int_0^\infty \log|S(i\omega)|d\omega = \pi \sum \mathrm{Re}\, p_k - \frac{\pi}{2} \lim_{s\to\infty} sL(s)$$

Requirements and Sensitivity Functions

Disturbances

- Effect of feedback: $y_{cl} = Sy_{ol}$
- Load disturbances: $G_{yd} = PS$
- Measurement noise: $G_{un} = -CS$

Process uncertainty

- Small variations: $\delta T/T = S\delta P/P$
- Large variations: $|\Delta P|/|P| \leq 1/|T|$, stable ΔP
- Gain and phase and sensitivity margins: g_m, φ_m, $s_m = \frac{1}{M_s}$

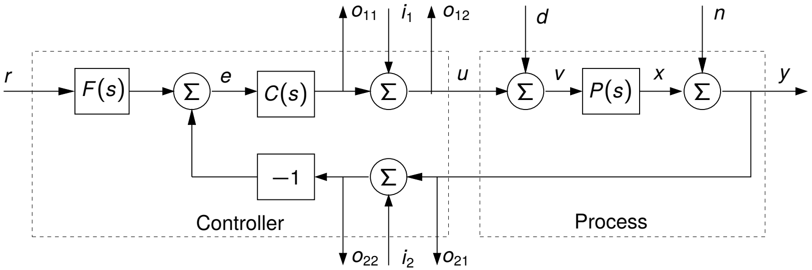
Command signal following

- Error feedback: $G_{yr} = T$, $G_{ur} = CS$
- 2DOF: $G_{yr} = TF$, $G_{ur} = CSF$

Testing Requirements

Introduce test points in the control system!

Use test signals in design phase and on the real system!



Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Gangs of Four and Seven
4. The Sensitivity Functions
5. Fundamental Limitations
6. Summary

Theme: A closer look at feedback

Fundamental Limitations

Large signal behavior

- Limits on actuation: size and rate
- Limits due to equipment safety

Small signal behavior

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

Dynamics

- Nonminimum phase dynamics

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

Voice coil drive for a hard disk drive

$$J\frac{d^2\varphi}{dt^2} = T = k_t I$$

$$m\frac{d^2x}{dt^2} = F = k_t I$$

$$r = 0.05\, m$$

$$J = 5 \times 10^{-6}\, kg\, m^2$$

$$m = 2 \times 10^{-3}\, kg$$

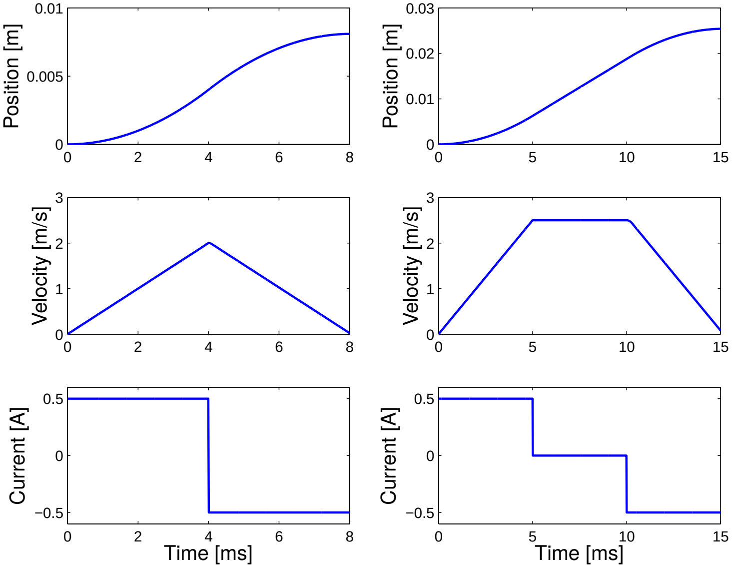
$$k_t = 2\, N/A$$

$$I_{max} = 0.5A$$

$$V_{max} = 5\, V$$

Maximum acceleration $a_{max} = \frac{k_t I_{max}}{m} = 500\, m/s^2$ (50g) Maximum velocity $v_{max} = \frac{V_{max}}{k_t} = 2.5$ m/s

Minimum Time Transitions



Limitations due to NMP Dynamics

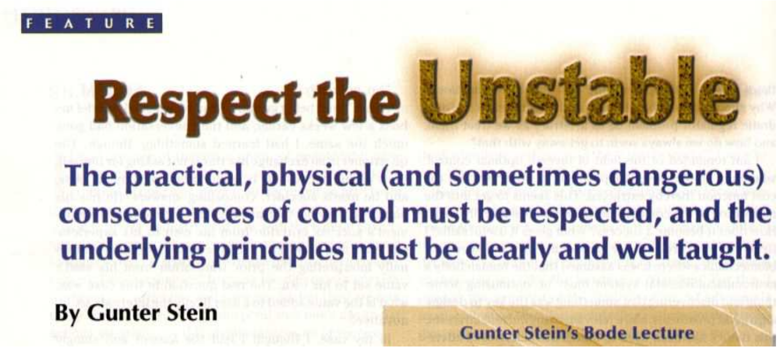
Process dynamics can impose severe limitations on what can be achieved.

- An important part of recognizing the difficult problems
- Time delays and RHP zeros limit the achievable bandwidth
- Poles in the RHP requires high bandwidth
- Systems with poles and zeros in the right half plane can be very difficult or even impossible to control robustly. Think about the bicycle with rear wheel steering!

Remedies:

- Add sensors and actuators (changes and removes zeros) or redesign the process

The First IEEE Bode Lecture 1989



Video by IEEE

`http://www.ieeecss-oll.org/video/respect-unstable`

Published in IEEE Control Systems Magazine August 2003

Summary of Dynamics Limitations

- A RHP zero z limits the achievable gain crossover frequency

$$\omega_{gc} \leq z \sqrt{\frac{M_s - 1}{M_s + 1}}$$

- A RHP pole p requires a high gain crossover frequency

$$\omega_{gc} \geq p \sqrt{\frac{M_s + 1}{M_s - 1}}$$

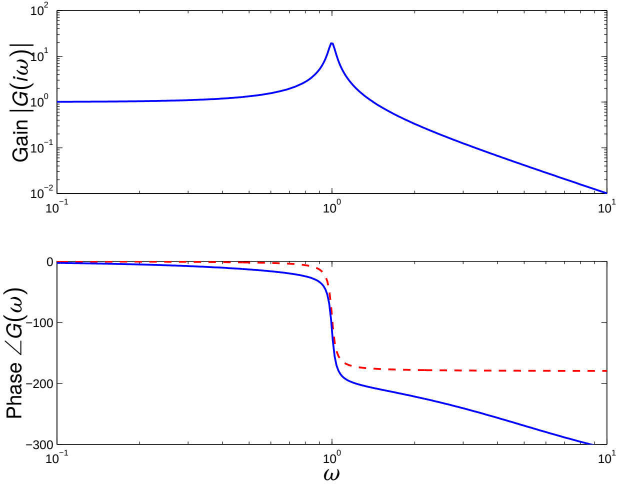
- A RHP pole-zero pair gives a lower bound to the maximum sensitivities

$$M_s \geq \left|\frac{p+z}{p-z}\right|, \qquad M_t \geq \left|\frac{p+z}{p-z}\right|$$

- A RHP pole p and a time delay τ gives a lower bound to the maximum sensitivities

$$M_s \geq e^{p\tau}, \qquad M_t \geq e^{p\tau}$$

Bode Plots – Nonminimum Phase Systems



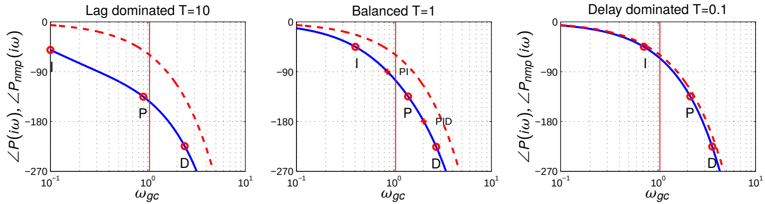
Blue curves Bode gain and phase curves of $G = G_{mp}G_{ap}$

Notice that $\left|G_{ap}(i\omega)\right| = 1$

Red dashed phase of minimum phase component G_{mp}

Assessment Plots – Achievable Gain Crossover Frequencies

Process transfer function $P(s) = \frac{1}{1+sT}e^{-s}$



- Blue line: phase curve for process transfer function $G = G_{mp}G_{ap}$
- Red dashed line: phase curve for allpass factor of process transfer function ($G_{ap} = e^{-s}$).
- Red full vertical line: phaselag of allpass component of process transfer function $G_{ap}(s)$ is 60°
- Designs for 45° phase margin for pure I: phase −45°, P: phase −135° and D: phase −225° controllers

Consequences for Design

- Poles are intrinsic properties of the system
- Changing poles requires redesign of the system
- Zeros depend on how inputs and outputs are coupled to the states
- Zeros and time delays can be changed by moving or adding sensors (and actuators)

A nice property of the FOTD $P(s) = \frac{K}{1 + sT} e^{-sL}$ and SOTD $P(s) = \frac{K}{(1 + sT_1)(1 + sT_2)} e^{-sL}$ are that they capture the time delay which limits the achievable performance. Time delay is influenced by sensor positions!

Stabilizing an Inverted Pendulum with Delay

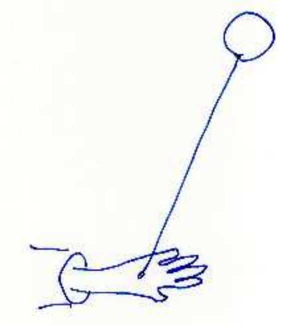
Right half plane pole at

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

Requiring $p\tau < 0.33$ gives $\tau\sqrt{g/\ell} < 0.33$
or

$$\ell > \frac{g\tau^2}{0.3^3} \approx 100\tau^2$$

A neural lag of 0.07 s gives $\ell > 0.5$ m
Make a demo!



A vision based system with sampling rate of 50 Hz (a time delay of 0.02 s) can robustly stabilize the pendulum if $\ell > 0.04$ m.

Example - The X-29

Advanced experimental aircraft. Much design effort was done with many methods and much cost. Specifications $\varphi_m = 45^\circ$ could not be reached.

Nonminimum phase factor

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

The zero-pole ratio is $\dfrac{z}{p} = \dfrac{13}{3}$

$$M_s \geq \frac{13+3}{13-3} = 1.6$$

$$\varphi_m \leq 2\arcsin\frac{1}{2M_s} = 36^\circ$$

The simple calculation shows

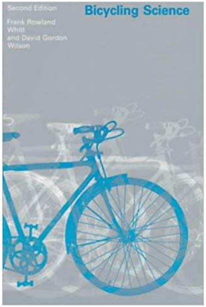
that it is impossible to obtain a phase margin of 45°!

Millions of dollars could have been saved!!

Bicycles with Rear Wheel Steering

Whitt and Wilson Bicycling Science 1982

Many people have seen theoretical advantages in the fact that front-drive, rear-steered recumbent bicycles would have simpler transmissions than rear-driven recumbents and could have the center of mass nearer the front wheel than the rear. The U.S. Department of Transportation commissioned the construction of a safe motorcycle with this configuration. It turned out to be safe in an unexpected way: No one could ride it.



The difficulties are caused by the dynamical properties of rear wheel steering. One reason for learning control is to find such difficulties at an early stage of the design.

Richard Klein University of Illinois

Using bicycles in control education for Mechanical Engineers

Bicycle with Rear Wheel Steering

Klein's bikes

Transfer function regular bike

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

Transfer function rear wheel steering (change sign of V_0)

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

RHP pole at $\sqrt{mg\ell/J}$

RHP zero at $-V_0/a$

The Lund Unridable Rear-Steered Bike

The UCSB Ridable Rear-Steered Bike

Feedback Fundamentals

1. Introduction
2. Controllers with Two Degrees of Freedom
3. The Gangs of Four and Seven
4. The Sensitivity Functions
5. Fundamental Limitations
6. Summary

Theme: A closer look at feedback

Summary

- Error feedback and systems with two degrees of freedom 2DOF

  2DOF allows separation of command signal response from the other requirements
- A system with error feedback is characterized by four transfer functions (Gang of Four GOF) S, T, PS, CS
- A system with two degrees of freedom is characterized by seven transfer functions (Gang of Seven = GOF, FS, FT and FCS)
- Several transfer functions are required to understand a feedback system. Analysis and specifications should cover all transfer functions!
- There are fundamental limiations caused by nonlinear as well as linear non-minimum phase dynamics