Essentials of Control Systems

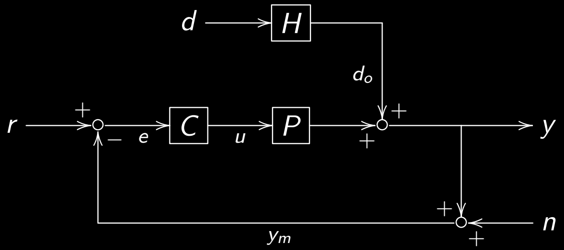
Principles of Feedback Design

Xu Chen

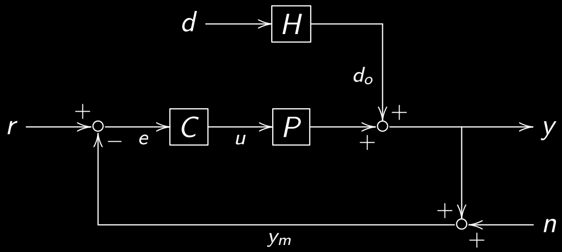
University of Washington

Outline

Goals of Feedback Design



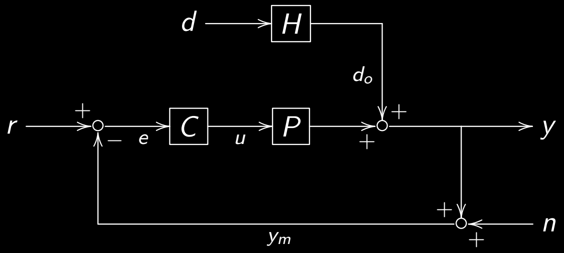
- The goal of feedback control is to make $y \approx r$ independent of d (d_o) and n, and make the behavior robust to modest/small changes in P.



- In matrix form, the closed-loop relationship from all of the inputs (r, d, n) to the outputs (y, u) is

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} G_{r\to y} & G_{d_o\to y} & G_{n\to y} \\ G_{r\to u} & G_{d_o\to u} & G_{n\to u} \end{bmatrix} \begin{bmatrix} r \\ d_o \\ n \end{bmatrix} = \begin{bmatrix} \frac{PC}{1+PC} & \frac{1}{1+PC} & \frac{-PC}{1+PC} \\ \frac{C}{1+PC} & \frac{-C}{1+PC} & \frac{-C}{1+PC} \end{bmatrix} \begin{bmatrix} r \\ d_0 \\ n \end{bmatrix}$$

Goals of Feedback Design



In mathematical terms, the goals are:

1. Make the magnitude of $G_{d_o \to y}$ small.
2. Make $G_{r \to y}$ approximately equal to 1, and insensitive to P.
3. Make the magnitude of $G_{n \to y}$ "small."

Goals of Feedback Design

$$\begin{bmatrix} G_{r\to y} & G_{d_o\to y} & G_{n\to y} \\ G_{r\to u} & G_{d_o\to u} & G_{n\to u} \end{bmatrix} = \begin{bmatrix} \frac{PC}{1+PC} & \frac{1}{1+PC} & \frac{-PC}{1+PC} \\ \frac{C}{1+PC} & \frac{-C}{1+PC} & \frac{-C}{1+PC} \end{bmatrix}$$

- Goal 1 (Make $|G_{d_o\to y}|$ small) implies

$$\left|\frac{1}{1+PC}\right| \ll 1 \Longleftarrow \text{design } |PC| \gg 1$$

- Goal 2 (Make $G_{r\to y} \approx 1$) implies

$$\frac{PC}{1+PC} \approx 1.$$

  We have already established that $|PC|$ is large (relative to 1), so

$$\frac{PC}{1+PC} \approx \frac{PC}{PC} = 1.$$

- Goal 3 (Make $|G_{n\to y}|$ "small") requires

$$\left|\frac{PC}{1+PC}\right| \ll 1 \Longleftarrow \text{design } |PC| \ll 1$$

  in conflict with Goal 1!

Outline: Essentials of Control Systems

1. Goals of Feedback Design
2. Tradeoffs
3. The Concept of High-gain Control
4. Fundamental Limitations of LTI Systems

Tradeoffs

$$\begin{bmatrix} y \\ u \end{bmatrix} = \begin{bmatrix} G_{r\to y} & G_{d_o\to y} & G_{n\to y} \\ G_{r\to u} & G_{d_o\to u} & G_{n\to u} \end{bmatrix} \begin{bmatrix} r \\ d_o \\ n \end{bmatrix} = \begin{bmatrix} \frac{PC}{1+PC} & \frac{1}{1+PC} & \frac{-PC}{1+PC} \\ \frac{C}{1+PC} & \frac{-C}{1+PC} & \frac{-C}{1+PC} \end{bmatrix} \begin{bmatrix} r \\ d_0 \\ n \end{bmatrix}$$

- Two key closed-loop properties:

$$\text{sensitivity function: } S \triangleq (1+PC)^{-1}$$

$$\text{complementary sensitivity function: } T \triangleq PC(1+PC)^{-1}$$

- S: measures the closed-loop disturbance-attenuation property;
- T : defines how the system responds to the reference input and the sensor noise.

Tradeoffs

$$\text{sensitivity function: } S \triangleq (1 + PC)^{-1}$$

$$\text{complementary sensitivity function: } T \triangleq PC(1 + PC)^{-1}$$

- Fundamental constraint:

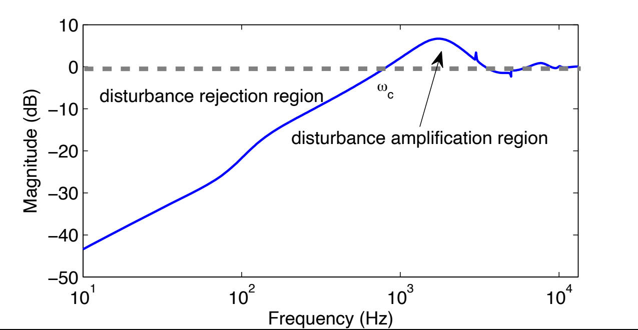
$$S + T = 1$$

  If one of the sensitivity measures is very small, then the other sensitivity measure will be approximately 1!

- e.g., wherever the magnitude of S is small, T will be close to unity. Hence, at frequencies where we achieve good disturbance rejection, we also obtain improved reference tracking but amplified sensor noise.

Tradeoffs

A typical magnitude response of S for single-input single-output (SISO) systems (x axis is in log scale):



How will T look like?

Outline: Essentials of Control Systems

The Concept of High-gain Control

- High-gain feedback control: To make $S = (1 + PC)^{-1}$ small at certain frequencies, $L \triangleq PC$ needs to be large in the same region.
- *Loop shaping*: designing the magnitude response of L to stabilize the loop and satisfy the desired performance metrics.

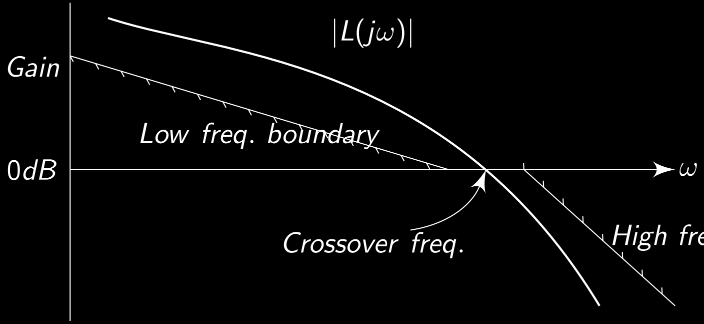


Figure: A typical loop shape in feedback servo design: Below the crossover frequency, $|L(j\omega)|$ is lower bounded to achieve disturbance rejection and reference tracking; at high frequencies, the loop shape is upper bounded for robustness.

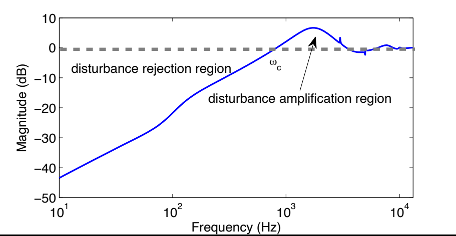
Outline: Essentials of Control Systems

1. Goals of Feedback Design
2. Tradeoffs
3. The Concept of High-gain Control
4. Fundamental Limitations of LTI Systems

Fundamental Limitations of LTI Systems

Mathematically, $G(s)$ / $G(z)$ is a rational function of complex variables. Complex analysis provides several guidance and fundamental limitations about the achievable performance in LTI systems. We review a few that are most relevant to this course.

Bandwidth Limitation



- The bandwidth ω_c cannot be pushed to be arbitrarily large.
- *Waterbed effect*: Under mild conditions, it is inevitable to have $|S(\omega')| > 1$ at certain frequencies if $|S(\omega)| < 1$ holds over some frequency interval, namely, when certain disturbance components are attenuated in the feedback system, some other disturbance components will be amplified. This is mathematically elaborated in the Bode's Integral Theorem.

Theorem (Bode's Integral Theorem)

Let $S(s) = 1/(1+L(s))$*. If* $L(s)$ *and* $S(s)$ *are both rational and stable, then*

$$\frac{1}{\pi}\int_0^{\infty} \ln\left|S\left(j\omega\right)\right| d\omega = \frac{-1}{2} k_s$$

$$k_s = \lim_{s\to\infty} sL\left(s\right)$$

Special case: If the relative degree of $L\left(s\right)$ *is larger than or equal to 2, then*

$$\frac{1}{\pi}\int_0^{\infty} \ln\left|S\left(j\omega\right)\right| d\omega = 0$$

A Closer Look at the Waterbed Effect: The S Function

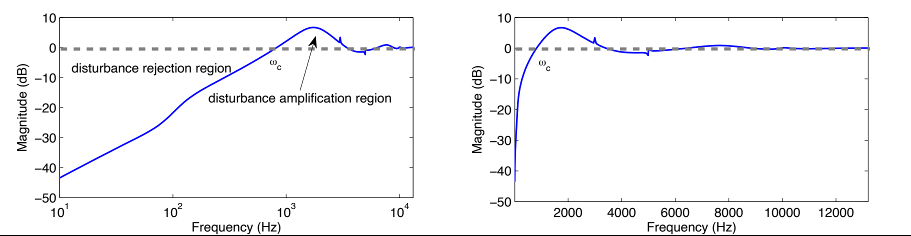


Figure: A typical magnitude response of the sensitivity function: left - x axis is in log scale, right - x axis is in linear scale.

A Closer Look at the Waterbed Effect: Conditions

Theorem (Bode's Integral Theorem)

Let $S(s) = 1/(1 + L(s))$. If $L(s)$ and $S(s)$ are both rational and stable, then

$$\frac{1}{\pi}\int_0^{\infty} \ln|S(j\omega)|d\omega = \frac{-1}{2}k_s$$

$$k_s = \lim_{s\to\infty} sL(s)$$

Special case: If the relative degree of $L(s)$ is larger than or equal to 2, then

$$\frac{1}{\pi}\int_0^{\infty} \ln|S(j\omega)|d\omega = 0$$

- Practical mechanical systems have small gains at high frequencies. The rate of high-frequency rolloff is usually higher than (or at least as fast as) $1/s$. (e.g., motors commonly take force/torque as the input, and generate (angular) position or velocity as the output.)

A Closer Look at the Waterbed Effect: Conditions

The waterbed effect gets worse if the open-loop system has unstable poles.

Theorem (Bode's integral formula for continuous-time SISO systems)

Let $L(s)$ be a proper, scalar rational transfer function, of relative degree larger than 1. Let $S(s) = 1/(1 + L(s))$ and assume that $S(s)$ has no poles in the right half plane, and $L(s)$ has $q \geq 0$ poles in the closed right half plane, at locations p_1, p_2, ..., p_q. Then

$$\frac{1}{\pi}\int_0^\infty \ln |S(j\omega)| d\omega = \sum_{k=1}^{q} p_k$$

Implication: When open-loop unstable poles (which becomes zeros of $S(s)$ in the closed right half plane) exist, disturbance amplification will always happen regardless of the presence of any disturbance attenuation.

More General Integral Constraints

If the relative degree is not larger than 1, the situation is easier:

Theorem (Extended integral equality for continuous-time SISO systems of unity relative degrees)

Let $L(s)$ be a proper, scalar rational transfer function, of relative degree 1. Let $S(s) = 1/(1 + L(s))$ and assume that $S(s)$ has no poles in the right half plane, and $L(s)$ has $q \geq 0$ poles in the closed right half plane, at locations p_1, p_2, ..., p_q. Then

$$\int_0^\infty \ln |S(j\omega)| \, d\omega = \pi \sum_{k=1}^{q} p_k - \frac{1}{2}\pi \lim_{s\to\infty} sL(s). \tag{1}$$

More General Integral Constraints

Theorem (Extended integral equality for continuous-time SISO systems with zero relative degrees)

Let $L(s)$ be a proper, scalar rational transfer function:

$$L(s) = \frac{K \prod_{i=1}^{n} (s - z_i)}{\prod_{i=1}^{n} (s - p_i)}, \; K \neq 0.$$

Let $S(s) = 1/(1 + L(s))$ and assume that $S(s)$ has no poles in the right half plane, and $L(s)$ has $q \geq 0$ stable poles in the open left half plane, at locations p_{s_1}, p_{s_2}, ..., p_{s_q}. Then

$$\int_0^{\infty} \ln |S(j\omega)| \, d\omega = \begin{cases} +\infty & \text{if } -2 < K < 0 \\ -\infty & \text{if } K > 0 \text{ or } K < -2 \\ \pi \left(\sum_k z_k - \sum_{k=1}^{q} p_{s_k} \right) & \text{if } K = -2 \end{cases} . \tag{2}$$

Waterbed Effect: Discrete-Time Case

For continuous-time systems, the waterbed effect holds if the relative degree of the loop transfer function $L = PC$ is no less than two. In the discrete-time case, the waterbed effect is more inevitable:

Theorem (Bode's integral formula for discrete-time systems)

For all well-posed, closed-loop stable discrete-time feedback systems under ZOH, the sensitivity function $S(z) = 1/(1 + L(z))$ satisfies

$$\int_0^{\pi} \ln \left| S\left(e^{j\omega}\right) \right| d\omega = \pi \left(\sum_{k=1}^{q} \ln |p_k| - \ln |\gamma + 1| \right) \tag{3}$$

where $\{p_k\}_{k=1}^{q}$ are the open-loop unstable (i.e. outside the closed-unit disk) poles of $L(z)$, and $\gamma \triangleq \lim_{z\to\infty} L(z)$.
$\gamma = 0$ if L is strictly proper. One intuition for the drop of the relative degree requirement here is that the zero order hold has a low-pass type dynamics and introduces high-frequency rolloffs in the magnitude response.

Limitations From Nonminimum-phase Zeros

Interesting effect of nonminimum-phase zeros:

- unbounded input can produce zero output at the steady state of a nonminimum-phase system. Consider, for instance, $u(t) = e^{2t}$ and $G(s) = (s-2)/(s+1)^2$. The associated ordinary differencial equation is

$$\ddot{y} + 2\dot{y} + y = \dot{u} - 2u = \frac{d}{dt}e^{2t} - 2e^{2t} = 0$$

  As $G(s)$ is stable, $y(t)$ thus will be zero at steady state.

But more often, nonminimum-phase zeros will place fundamental limitations in feedback design, as we show next.

Limitations From Nonminimum-phase Zeros

An important theorem from complex analysis will be immediately useful.

Theorem (Maximum modulus theorem)

If a complex function $S(\sigma)$ is defined and continuous on a closed bounded set Ω, and it is analytic on the interior of Ω, then $|S(\sigma)|$ can not attain the maximum in the interior of Ω unless it is a constant.

- analytic functions on the right-half plane: cannot have unstable poles, e.g.
$$G(s) = \frac{1}{s+1},\ \frac{2+s}{s+3},\ \frac{s-2}{s+2}$$

Limitations From Nonminimum-phase Zeros

- the stable sensitivity function will always have magnitudes larger than one

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

  - if σ_o is a nonminimum-phase zero of $P(s)$, then $P(\sigma_o) = 0$ and $S(\sigma_o) = 1/(1 + 0 \times C(\sigma_o)) = 1$, regardless of the design of a stabilizing $C(s)$
  - if S is stable—hence analytic in the right-half plane—then the maximum modulus theorem indicates that the maximum of $|S(s)|$ is achieved on the imaginary axis
  - hence $\max_\omega |S(j\omega)| \geq |S(\sigma_o)| = 1$
  - there will thus always be a frequency region where we cannot achieve high servo performance, particularly around the frequencies of the nonminimum-phase zeros

Limitations From Nonminimum-phase Zeros (Cont'd)

- step responses can have initial undershoot. Moreover, zero crossovers occur for step responses. Consider

$$Y(s) = \frac{G(s)}{s} = \int_0^\infty y(t)\, e^{-st} dt$$

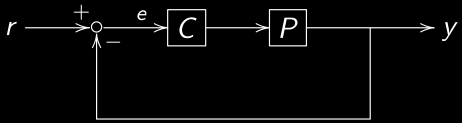
If $G(\sigma_o) = 0$ where $\sigma_o > 0$ is the unstable zero, then

$$Y(\sigma_o) = \int_0^\infty y(t)\, e^{-\sigma_o t} dt = 0$$

As $e^{-\sigma_0 t}$ is positive and decreasing, $y(t)$ must change signs for $t \in [0, \infty)$.

Limitations From Nonminimum-phase Pole-Zero Pairs

Consider a stable negative feedback system with $P(s)$ and $C(s)$.



- Assume that the plant $L(s) := P(s)C(s)$ has a right-half-plane (RHP) zero at z and a RHP pole at p.
- Either the RHP pole or the RHP zero alone will result in $\max_\omega |S(j\omega)|$ to be larger than one.
- We show next a tighter bound when both the RHP pole and zero exist.

Limitations From Nonminimum-phase Pole-Zero Pairs

- The RHP pole of $L(s)$ implies:

$$S(s) = \frac{1}{1 + L(s)}$$

has a right-half-plane zero at p. As $S(s)$ itself is stable, we can factorize out the unstable zero such that

$$S(s) = \frac{s - p}{s + p}\tilde{S}(s)$$

where all zeros and poles of $\tilde{S}(s)$ are minimum-phase. Notice that the magnitude response of $(s - p)/(s + p)$ is strictly one![1] Hence

$$\max_{\omega} |S(j\omega)| = \max_{\omega} \left| \frac{j\omega - p}{j\omega + p}\tilde{S}(j\omega) \right| = \max_{\omega} \left| \tilde{S}(j\omega) \right| \tag{4}$$

[1] The factorization is known as an all-pass factorization.

Limitations From Nonminimum-phase Pole-Zero Pairs (Cont'd)

- The RHP zero of $L(s)$ implies:

$$S(z) = 1$$

$\tilde{S}(s)$ is stable and anlytic on the RHP. Based on maximum modulus theorem

$$\max_{\omega} |\tilde{S}(j\omega)| \geq |\tilde{S}(z)| = \left| \cancelto{1}{S(z)} \frac{z+p}{z-p} \right| = \left| \frac{z+p}{z-p} \right| > 1$$

But $\max_{\omega} |S(j\omega)| = \max_{\omega} \left|\tilde{S}(j\omega)\right|$ from (4). Hence

$$\max_{\omega} |S(j\omega)| \geq \left| \frac{z+p}{z-p} \right| > 1$$

- **Implication**: If z and p are very close to each other, the maximum magnitude of $S(s)$ will be significantly larger than one.

Additional Reading

Many of the ideas about loop shaping and performance limits are from [1, 2]. [3] is a richer reference containing also MIMO control theory. In particular, Gunter Stein's "Respect the Unstable" as the 1989 Bode Lecture is an excellent read for every control engineer and researcher.

- J. C. Doyle, B. A. Francis, and A. Tannenbaum, *Feedback control theory*. Macmillan, 1992, vol. 134.
- G. Stein, "Respect the unstable," *IEEE Control Systems Magazine*, vol. 23, no. 4, pp. 12 – 25, Aug. 2003.
- S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*, 2nd ed. Wiley Chichester, UK, 2005.