Repetition

Automatic Control, Basic Course, Lecture 11

Fredrik Bagge Carlson

December 17, 2016

Lund University, Department of Automatic Control

Content

Lecture 1

Lecture 1 - Intro PID and State-space

Contents

- PID controller heuristically derived
- State-space models, relation to differential equations

PID controller

P Control signal proportional to the error. If the error is large, it seems reasonable to use more effort than if the error is small.

I Increase effort if error persist. Keep increasing the effort until the error vanishes.

D Modify effort based on prediction. If the error is rapidly decreasing, ease up on the effort to not overshoot.

State-space models

- Systems often modeled as differential equations.
- All ordinary differential equations can be put in state-space form.
- State-space models admit easy analysis of stability and control design.

Control design in time-domain

Control is about changing the solution to the differential equation governing the system!

$$\dot{y} = ay(t) + bu(t)$$

- We are free to chose $u(t)$ to alter the solution to the differential equation.
- $u(t) = -ay(t) \Rightarrow \dot{y} = 0$
- $u(t) = -ay(t) + f(t) \Rightarrow \dot{y} = f(t)$ We can assign arbitrary dynamics to this system.
- Control design in the time-domain becomes hard as the order of the system increases. Analysis in the frequency domain remains simple no matter the order of the system.

Lecture 2

Linearlization

- Nonlinear models are hard to analyze.
- Linearization gives a linear model, valid around the linearization point.

$$\dot{x} = f(x, u)$$
$$y = g(x, u)$$

$$\dot{\Delta x} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \dfrac{\partial f_1}{\partial x_2} \\ \dfrac{\partial f_2}{\partial x_1} & \dfrac{\partial f_2}{\partial x_2} \end{pmatrix} \Delta x + \begin{pmatrix} \dfrac{\partial f_1}{\partial u} \\ \dfrac{\partial f_2}{\partial u} \end{pmatrix} \Delta u$$

$$\Delta y = \begin{pmatrix} \dfrac{\partial g}{\partial x_1} & \dfrac{\partial g}{\partial x_2} \end{pmatrix} \Delta x$$

Transfer functions

- The transfer function is given by the Laplace transform of the system differential equation.
- Only linear differential equations have a transfer function.
- The transfer function relates output to input.
- The transfer function is the Laplace transform of the impulse response.
- A sinusoid $\sin(\omega t)$ on the input is changed to $|G(i\omega)|\sin(\omega t + \arg G(i\omega))$ at the output.

Example

$$\ddot{y} + a_1\dot{y} + a_2 y = b_1\dot{u} + b_2 u$$
$$s^2 Y(s) + a_1 s Y(s) + a_2 Y(s) = b_1 s U(s) + b_2 U(s)$$
$$(s^2 + a_1 s + a_2)Y(s) = (b_1 s + b_2)U(s)$$
$$Y(s) = \frac{b_1 s + b_2}{s^2 + a_1 s + a_2}U(s)$$

State-space → TF

$$\dot{x} = Ax + Bu$$
$$y = Cx$$

Laplace transformation yields

$$sX(s) = AX(s) + BU(s)$$
$$Y(s) = CX(s)$$

Solve for $X(s)$ in the first equation and subsequently eliminate $X(s)$ from the second one

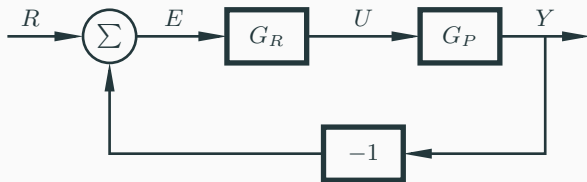$$X(s) = (sI - A)^{-1}BU(s)$$
$$Y(s) = C(sI - A)^{-1}BU(s)$$

The transfer function thus becomes

$$G(s) = C(sI - A)^{-1}B$$

- No unique solution.
- Three *canonical* solutions given in collection of formulae.

Block diagrams

Calculations in block diagrams are used to obtain transfer functions from one signal to another

Example:

The closed-loop transfer function from R to Y is given by

$$Y = G_R G_P (R - Y)$$
$$Y + G_R G_P Y = G_R G_P R$$
$$= \frac{G_R G_P}{1 + G_R G_P} Y$$

Step response

Why is the Laplace transform of both the integration operation and a step function the same? $1/s$

- The Laplace transform of integration $\int f(t) = \dfrac{1}{s} F(s)$

- If we integrate an impulse $\delta(t)$, the resulting function is the Heaviside step function $\int \delta(t) = \vartheta(t)$

- The Laplace transform of $\delta(t)$ is 1.

- Hence, $\mathcal{L}\vartheta(t) = \mathcal{L} \int \delta(t) = \dfrac{1}{s}\mathcal{L}\delta(t) = \dfrac{1}{s}$

Step response

Why is the Laplace transform of both the integration operation and a
step function the same? $1/s$

• The Laplace transform of integration $\int f(t) = \frac{1}{s} F(s)$
• If we integrate an impulse $\delta(t)$, the resulting function is the
  Heaviside step function $\int \delta(t) = \theta(t)$
• The Laplace transform of $\delta(t)$ is 1.
• Hence, $\mathcal{L}\theta(t) = \mathcal{L}\int \delta(t) = \frac{1}{s}\mathcal{L}\delta(t) = \frac{1}{s}$

Why does an integrator have constant phase -90°?

The integration of a sinusoid $\sin(t)$ is $\int \sin(t) = -\cos(t)$, these differ by -90° in phase. The differentiation of a sinusoid $\sin(t)$ is $\dfrac{d\sin(t)}{dt} = \cos(t)$, these differ by 90° in phase.

Lecture 3

Relation between TF and step response

- Initial and final value theorems can be used to provide

$$\lim_{t \to 0} f(t) = \lim_{s \to \infty} sF(s)$$

$$\lim_{t \to \infty} f(t) = \lim_{s \to 0} sF(s)$$

- Can only be used when $sF(s)$ has all poles strictly in the left half plane.
- Can be used to study the response to an arbitrary signal with known Laplace transform.

Examples

$$F(s) = G(s)\frac{1}{s} \quad \text{Step response}$$

$$F(s) = G(s)\frac{1}{s^2} \quad \text{Ramp response}$$

Relation between TF and step response

- Initial and final value theorems can be used to provide
$$\lim_{t \to 0} f(t) = \lim_{s \to \infty} sF(s)$$
$$\lim_{t \to \infty} f(t) = \lim_{s \to 0} sF(s)$$
- Can only be used when $sF(s)$ has all poles strictly in the left half plane.
- Can be used to study the response to an arbitrary signal with known Laplace transform.

Examples

$$F(s) = G(s)\frac{1}{s} \quad \text{Step response}$$
$$F(s) = G(s)\frac{1}{s^2} \quad \text{Ramp response}$$

When using these theorems, it is often of interest to study the response of steps or ramps through the transfer function from reference to error, from reference to output or from disturbance to error.

Interactive software (link)

Lecture 4

Stability definitions

Asymptotic stability
: All poles are strictly in the left half plane. The impulse response eventually goes to zero.

Stability
: One or more distinct poles on the imaginary axis. Impulse response is bounded.

Instability
: One or more poles in the right half plane, or multiple overlaying poles on the imaginary axis. impulse response is unbounded.

For state-space realization of the system, the poles are the same as the eigenvalues of the A-matrix.

The location of the poles are related to the solution of the underlying differential equation. For an unstable system, the solution blows up. For a stable system, it remains bounded, for an as. stable system, it goes to zero.

All asymptotically stable systems are also stable. If there are two or more poles on the same location on the imaginary axis, the system is unstable even tough no poles might exist in the right half plane.

Characteristic polynomial - stability

First order system Stable if all coeffs > 0

Second order system Stable if all coeffs > 0

Third order system Stability criteria in collection of formulae.

Lecture 5

Nyquist criterion

- The stability of the closed-loop system under static feedback is determined by the behavior of the open-loop system at the frequency ω_0 where the phase of the open-loop system is -180°.
- This frequency is where the Nyquist curve crosses the negative real line.
- Gain must be small for this frequency
- Formally, the Nyquist curve may not encircle the point -1
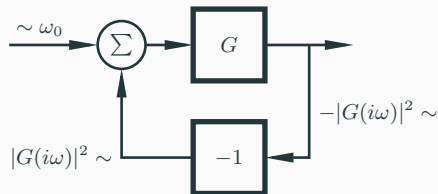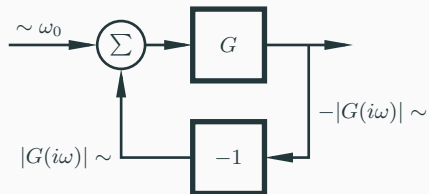
Blackboard

Sensitivity function and disturbance rejection

The sensitivity function

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

is a measure of disturbance rejection.

- $S(s)$ is the transfer function from measurement noise to output.
- $S(s)G_P(s)$ is the transfer function from load disturbance to output.
- Disturbances for which $|(S(i\omega)|$ is small are attenuated by the feedback.
- Always plot S after designing your controller! Especially after pole placement.

Sensitivity function and disturbance rejection

The sensitivity function

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

is a measure of disturbance rejection.

- $S(s)$ is the transfer function from measurement noise to output.
- $S(s)G_P(s)$ is the transfer function from load disturbance to output.
- Disturbances for which $|S(i\omega)|$ is small are attenuated by the feedback.
- Always plot S after designing your controller! Especially after pole placement.

When doing loop-shaping, we typically design our controller with robustness in mind. When doing pole placement, however, it's harder to have a feeling for the robustness properties of the closed-loop system. It is thus extra important to check the sensitivity function after performing pole placement to verify satisfactory robustness (small maximum value of the sensitivity function).

Lecture 6

State feedback

- The PID controller solves most ($> 90\%$?) of all simple control problems
- For strongly resonant systems, the PID controller performs poorly.
- The PID controller can achieve arbitrary pole placement for a second order system.
- Higher order systems require more sophistication for arbitrary pole placement.

Enter state feedback!

State feedback

If the system is *controllable*, one can achieve *arbitrary* pole placement using the control law

$$u = -Lx$$

State feedback

If the system is controllable, one can achieve arbitrary pole placement using the control law

$$u = -Lx$$

Determining controllability by checking the rank condition on the controllability matrix is actually a poor way of determining practical controllability. Systems which are *almost* uncontrollable still have a full rank controllability matrix. A better way is checking the singular values of the controllability *Gramian*, a concept which is covered in the course FRTN10 Multivariable control (link)

State feedback

With the control law

$$u = -Lx$$

The closed-loop system becomes

$$\dot{x} = Ax + B(-Lx) = (A - BL)x$$

hence, we have changed the dynamics of the system from A to $A - BL$

The trick is to choose L such that the eigenvalues of $A - BL$ are placed where we want them.

State feedback - merits and drawbacks

- We can place the poles *arbitrarily* for controllable systems.
- We have less feeling for the frequency properties of our controller.
- We have less feeling for the robustness properties of our closed-loop system.

Conclusion

Powerful method that gives limited insight.

Always check sensitivity function after pole placement!

How to place the poles

Roughly speaking: The further we move the poles of the open-loop system

- The more effort is required by the controller.
- The more sensitive to modeling errors we become.
- Well *damped* poles does **not** imply robustness!

Design rules of thumb

- Avoid making fast poles slower, increase damping if necessary.
- Speed of closed-loop system mostly determined by the slowest poles, use these poles for tuning.
- Low frequency process zeros must be matched by corresponding closed loop poles[1].

---

[1]Ref: note on next slide.

See Pole placement lecture by Bo Bernhardsson and Karl-Johan Åström (link to pdf from PhD course in control) for additional details, such as

- Design rule: Pick closed loop poles close to slow stable process zeros and fast stable process poles. Picking closed loop poles and zeros identical to slow stable zeros and fast stable poles give cancellations and simple calculations.

- Unstable poles and zeros cannot be canceled, slow unstable zeros and fast unstable poles therefore give *fundamental limitations*.

Lecture 7

Closely related to controllability.

If a system is observable, we can estimate all states using an *observer*.

An observer[2] makes use of both a *model of the system*, and the measured *output* to estimate the state of the system.

$$\dot{\hat{x}} = A\hat{x} + Bu + K(y - \hat{y})$$
$$\hat{y} = C\hat{x}$$

The estimate $\hat{x}$ evolves according to the model, with an added correction term $K(y - \hat{y})$ that compensates for the *estimation error*.

---

[2]A Kalman filter is an observer where the poles are placed based on the statistical properties of the system and measurement noise.

Observer - Pole placement

The observer can be designed using pole placement.

- Fast poles imply fast convergence of the estimation error.
- Fast poles imply high controller gain for high frequencies, i.e., high noise amplification.
- Design is a trade off between speed and noise sensitivity.
- Rule of thumb: make observer poles 1.5-2 times as fast as corresponding closed-loop poles.

Since the observer makes use of a model of the system, we can not expect too high performance if the model is not accurate.

For load disturbance rejection, the system model can be augmented with a model of the disturbance. The observer can then estimate the disturbance and compensate for it. Integral action can be achieved in this way by augmenting the system model with the model of a constant disturbance acting on the input.

Lecture 8

Lead-lag compensation

Methods to improve upon a nominal control design.

Lag compensation

- Affects low frequencies
- Decreases phase
- Used to reduce stationary errors

Lead compensation

- Affects high frequencies
- Increases phase
- Used to make system faster
- Used to increase the phase margin

Design in the frequency domain gives lots of insight!

Lead-lag compensation feels weird and abstract – help me!

- We would like to shape the behavior of our system, using feedback.
- The behavior is hard to analyze in the time domain.
- Reasoning in the frequency domain is much easier, but might feel abstract.
- Lead and lag links *shape* the frequency response of the open-loop system (loop shaping).
- Lead and lag links modify the differential equation governing the system, but it's hard to get a felling for how in the time domain.
- We have good measures of robustness[3] and speed in the frequency domain, we use those to *shape* the transfer function such that it becomes fast and robust.
- Get a feeling for how the links affect the Bode and Nyquist plots and you will master the art of loop-shaping.

[3] Phase and amplitude margins, φ_m, A_m

Lead-lag compensation feels weird and abstract – help me!

- We would like to shape the behavior of our system, using feedback.
- The behavior is hard to analyze in the time domain.
- Reasoning in the frequency domain is much easier, but might feel abstract.
- Lead and lag links shape the frequency response of the open-loop system (loop shaping).
- Lead and lag links modify the differential equation governing the system, but it's hard to get a feeling for how in the time domain.
- We have good measures of robustness[a] and speed in the frequency domain, we use those to shape the transfer function such that it becomes fast and robust.
- Get a feeling for how the links affect the Bode and Nyquist plots and you will master the art of loop-shaping.

[a]Phase and amplitude margins, φ_m, A_m

For hardcore loop-shaping a third useful plot exists: the Nichols Plot
https://cn.mathworks.com/help/control/ref/nichols.html

Loop shaping

- We shape the open-loop transfer function but are interested in the closed-loop performance.
- Nyquist stability criteria on the open-loop transfer function.
- It's easy to see the effect of a compensation link in the Bode plot of the open-loop system.

$$G_0^{new}(s) = G_K(s)G_0(s)$$
$$\log|G_0^{new}(i\omega)| = \log|G_K(i\omega)| + \log|G_0(i\omega)|$$
$$\arg G_0^{new}(i\omega) = \arg G_K(i\omega) + \arg G_0(i\omega)$$

Lag compensation

- Two parameters to choose.
- Gain M is chosen to get desired reduction of stationary errors[4].
- Bandwidth is chosen to not ruin the phase margin too much $(a \approx 0.1\omega_c)$.

---

[4]A lag link with infinite gain corresponds exactly to a PI controller, which we know eliminates stationary errors completely.

Lead compensation

The lead-link is used if we want to

- Make the system faster (increase ω_c).
- Make the system more robust.

In both cases we will need to lift the phase.

- Three parameters to choose
- N is determined by how much we need to lift the phase.
- b is chosen such that the maximum phase lift occurs at the desired crossover frequency.
- K_K is chosen to get the desired crossover frequency ω_c.

Lead compensation

If we increase the crossover frequency, we typically reduce the phase margin

1. Specify desired crossover frequency ω_c^d.
2. Determine required phase lift as the difference between system phase at ω_c^d and desired phase (given by desired phase margin φ_m).
3. Required phase lift gives required value of N from plot in collection of formulae.
4. Choose b to make phase top occur at ω_c^d (where it's needed the most).
5. Choose K_K to actually make ω_c^d the crossover frequency.

We could in principle choose a very high value of N and get very good robustness. Higher value of N means higher gain for high frequencies and high amplification of noise.

Lecture 9

Zeros

The zeros of a system affect the behavior.

- The effect is larger the slower (closer to the origin) the zero is.
- Zeros in the left half plane boosts the response, can cause overshoot if too severe.
- Zeros in the right half plane boosts the response in the wrong direction. These systems are hard to control.

P Setpoint handling

I Anti-windup

D Noise filtering

Setpoint handling

$$P = K(r - y)$$

- If the reference takes a step, u will take a step.
- Mitigated by $P = (br - y),\ b \in [0, 1]$
- The integrator will take care of the static gain.

Anti-windup

- If the control signal saturates due to an increase in the reference, we can not apply more control signal.

- Further integration of the error can not help.

- Allowing the integrator to continue integrating can cause severe harm! This is called integrator windup.

- It takes time to get rid of a large integrated error once the reference is lowered again and the control signal will stay saturated long after the reference is lowered.

Modifications of the D-part

- If r takes a step, the D-part becomes infinite $\rightarrow$ get rid of r in the D-part if r takes steps.

- The derivative causes very high amplification of noise $\rightarrow$ augment the derivative with a lowpass filter.
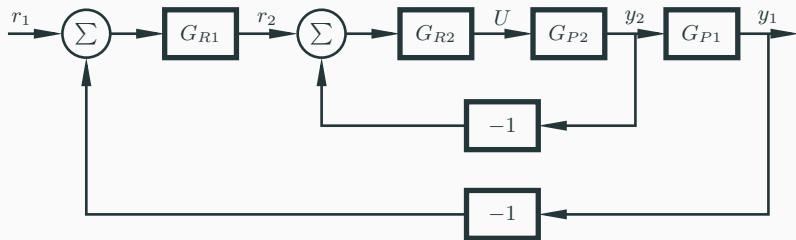
$$KT_d s \quad \longrightarrow \quad \frac{KT_d s}{1 + sT_d/N}$$

- N usually in $[5, 10]$
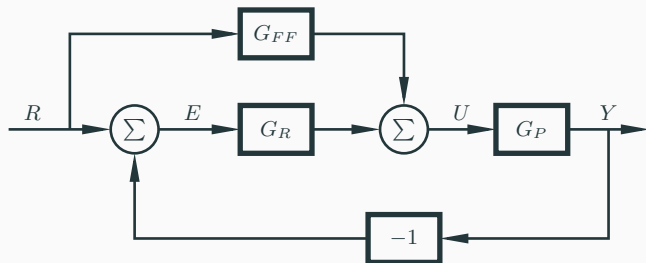
Cascade controllers

If an intermediate measurement y_2 is available, use it to close an inner loop and simplify the design.

- Easy method to design controllers that make use of more than one measurement signal.
- Can improve performance significantly if outer dynamics G_{P1} is slow.
- Disturbances in the inner loop can be taken care of before they become visible in y_1.

Feedforward

- Feedback requires something to happen before a reaction occurs, i.e., an error has to arise.
- If a known event is about to occur, and this event is known to create an error if not cared for, we may use feedforward to negate the effect of the event.
- Can be used for both measurable disturbances and setpoint changes.

$G_{FF} = G_P^{-1}$ would give perfect reference following but is often hard to realize in practice.

Feedforward

- Feedback requires something to happen before a reaction occurs, i.e., an error has to arise.
- If a known event is about to occur, and this event is known to create an error if not cared for, we may use feedforward to negate the effect of the event.
- Can be used for both measurable disturbances and setpoint changes.

$G_{FF} = G_P^{-1}$ would give perfect reference following but is often hard to realize in practice.

Feedforward obviously requires accurate knowledge of the disturbance we want to account for. If we see changes in the reference value as a disturbance, this is fulfilled. In this case we only need an accurate dynamics (inverse) model to eliminate the effect of the disturbance. If the disturbance is an auxiliary signal, we need an accurate measurement of this.

A typical linear model can be written like $Y(s) = G(s)U(s)$. For feedforward, we require the **inverse model** $U(s) = G^{-1}(s)Y(s)$, i.e., a model of the input, given the output. If we know the output Y we want, the inverse model tells us what input, or control signal U, to apply to get the desired output.

Feedforward from disturbance
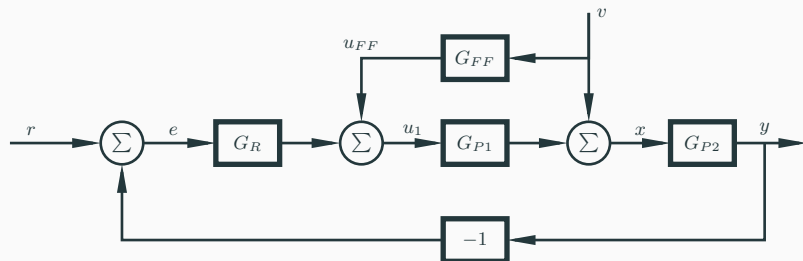
- Choose G_{FF} such that the effect of v at x is canceled.
- $G_{FF} = \dfrac{-1}{G_{P1}}$ does this.