Challenges in Control Education

K. J. Åström

Department of Automatic Control, Lund University

The Field of Control

- Servomechanism theory 1945
  - Drivers: gun control, radar, ...
  - A holistic view: theory, simulation and implementation
  - Block diagrams, Transfer functions, analog computing
- The second phase 1965
  - Drivers: space race, digital control, mathematics
  - Subspecialities: linear, nonlinear, optimal, stochastic, ...
  - Design methods: state feedback, Kalman filter, LQG, H_∞-control
  - Computational tools emerged
  - Impressive theory development, holistic view was lost
- The third phase 2005?
  - Drivers: Embedded systems, control over/of communication networks, systems biology
  - Exploding applications - Control everywhere
  - Software and hardware platforms
  - Recover the holistic view?

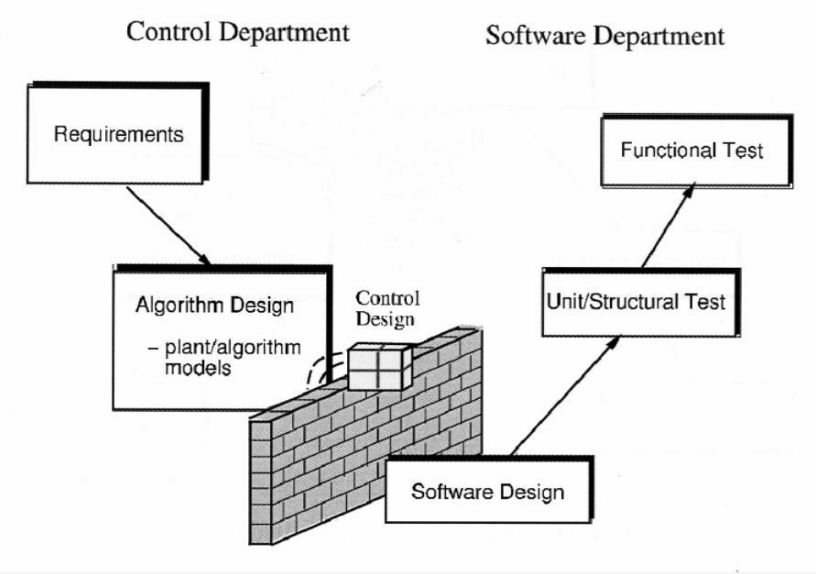
The Role of Computing

- Vannevar Bush 1927. *Engineering can proceed no faster than the mathematical analysis on which it is based. Formal mathematics is frequently inadequate for numerous problems, a mechanical solution offers the most promise.*
- Herman Goldstine 1962: *When things change by two orders of magnitude it is revolution not evolution.*
- Gordon Moore 1965: *The number of transistors per square inch on integrated circuits has doubled approximately every 18 months.*
- Moore+Goldstine: *A revolution every 15 year!*
- Unfortunately software does not evolve as fast as hardware

Computing and Control

- Implementation of controllers
- Design, analysis and simulation
- Controllers
  - Mechanical, pneumatical, electrical
  - Transistors, operational amplifiers
  - Computers
- Analog computing
  - Differential analyzer
  - Electron tubes, transistors, semiconductors
- Digital computers
  - General purpose, process control, DSP, FPGA
- Computer process control
- Embedded systems
  - Distributed networked systems

The Brick Wall - Loosing the Holistic View



Control Education

The dilemma of emerging fields

- Teach all that is known
- Develop more material
- Add more material
- Add more courses

Consolidation

- Sort, evaluate, select, and organize
- Focus on fundamentals, insight and practical relevance
- Exploit computing
- Don’t forget back-on-the-envelope calculations

The AFOSR Panel 2004

- Encourage the development of new courses and course materials that will significantly broaden the standard first introductory control course at the undergraduate level.
- Future Directions in Control in an Information-Rich World IEEE CSM 23 (2003) April pp 20–33.
- The panel believes that control principles are now a required part of any educated scientist's or engineers background ...
- Invest in new approaches to education and outreach for the dissemination of control concepts and tools to non-traditional audiences. The community must do a better job of educating a broader range of scientist and engineers on the principles of feedback and the use of control to alter the dynamics of systems and manage uncertainty

A Physicist View

The difficulties facing interested scientists to learn about control is illustrated by the following quote from John Bechhoefer. Feedback for Physicists. Rev. Mod. Phys 77 July 2005, 783-836.

The obvious places to learn about control theory - introductory engineering textbooks ... - are not very satisfactory places for a phycisist to start. They are long - 800 pages is typical - with the relevant information often scattered in different sections. Their examples are understandably geared more to the engineer than to the physicist. They often cloak concepts familiar to the physicist in unfamiliar language and notation. ... The main alternative, more mathematical texts, ..., are terse but assume that the reader already has an intuitive understanding of the subject.

Challenges for Education

Education of control engineerings

- Centralized or decentralized departments
- How to filter out the essence and compress it?
- Relations to computer science and communication,
- Experiments and the system aspects
- Exploit computers
- Experiences from Lund and UCSB

Education of other scientists: Mathematics, Physics, Biology, ...

- Why?
- Because feedback is fundamental and we know it best!
- Control for Scientists and Engineers - Caltech, Lund

Outline

- Introduction
- Introductory Course
- Feedback Fundamentals
- Advanced Courses
- The Systems View - Laboratories
- Conclusions

Introductory Course for Engineers

- An introductory course for engineers who only take one course
- Not the first of a long string of courses!
- UCSB Mechanical Engineering
- Lund Experience
- The Caltech course
- Lab development
- Interactive learning modules
- I am not sure we got it right yet

Goals for Introductory Course for Engineers

- Understand why control is useful - the Magic of Feedback
- Know the language, the key ideas and the concepts
- Know relevant mathematical theory
- Be able to solve simple control problems and to recognize difficult control problems
- Understand fundamental fundamental limitations
- Recognize when a system is easy or difficult to control
- Understand how to formulate and interprete specifications
- Have a working knowledge of the PID controller
- Practical hands-on experience of simple feedback loops
- Be able to use computational tools (Matlab)

The Magic of Feedback

Feedback has some amazing properties, it can

- make good systems from bad components,
- make a system insensitive to disturbances and component variations,
- stabilize an unstable system,
- create desired behavior, for example linear behavior from nonlinear components.

The major drawback is that

- Feedback can cause instabilities

What to Include?

- Transfer functions
- State feedback
- Observers
- Modeling
- Reachability and observability
- Stability theory
- Sensitivity functions
- Controller structure
- Laplace transforms?
- PID control
- Block diagrams
- Loop shaping?
- Fundamental limitations
- Root locus?
- Bode plots
- Nyquist plots
- Frequency response
- Design methods?
- Nonlinearities
- Simulation
- Linearization

The Language of Control

Concepts

- Feedback, feedforward, integral action
- Systems with two-degrees of freedom
- Minimum phase
- Observability, controllability

Standard models

- Block diagrams
- Transfer functions
- State models

Control algorithms

- PID
- State feedback
- Observers

Standard Models

- The role of standard models
- One way to deal with complexity
- Simple paths between problems and solutions
- Transfer functions
- State models

$$\frac{dx}{dt} = f(x,u) = Ax + Bu$$
$$y = g(x,u) = Cx + Du,$$

- Differential algebraic equations

$$F(z,\dot{z},u) = 0, \quad E\frac{dx}{dt} = Ax + Bu,$$

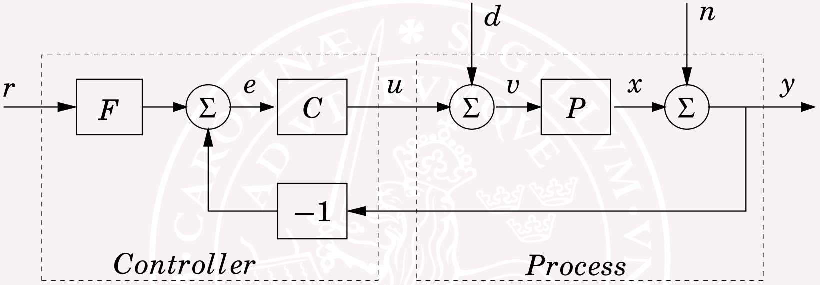
Outline

- Introduction
- Introductory Course
- Feedback Fundamentals
- Advanced Courses
- The Systems View - Laboratories
- Conclusions

Understanding the Basic Feedback Loop

- Effects of
  - Load disturbances
  - Measurement noise
  - Process variations
  - Command signals
- Assessment of the properties of a control system
- A basis for analysis, specification and design
- Concepts and insights
  - Systems with two degrees of freedom
  - Sensitivity functions
- Fundamental limitations
  - Bodes integral - the waterbed effect
  - Bodes relations - non-minimum phase

A Basic Control System with 2DOF

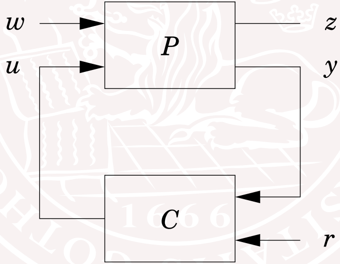


Ingredients:

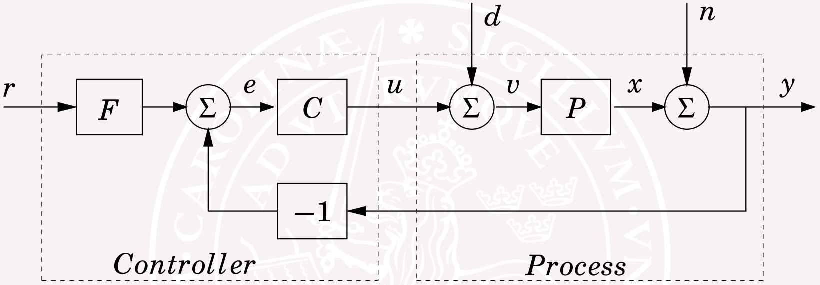
- Controller: feedback C, feedforward F
- Load disturbance d : Drives the system from desired state
- Measurement noise n : Corrupts information about x
- Process variable x should follow reference r

A Remark on Load Disturbances

Load disturbances are assumed to enter at the process input and measurement noise at the process output. The same idea can be applied to other configurations. A general structure is given below.



Criteria for Control Design



Ingredients

- Attenuate effects of load disturbance d
- Do not feed in too much measurement noise n
- Make the system insensitive to process variations
- Make state x follow command r

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 variations

Then design the feedforward F to achieve the desired response to command signals r.

Notice

- Many books and papers show only the set point response
- In process control the load disturbance response is much more important than the set point response.
- The set point response is more important in motion control.

Process Control

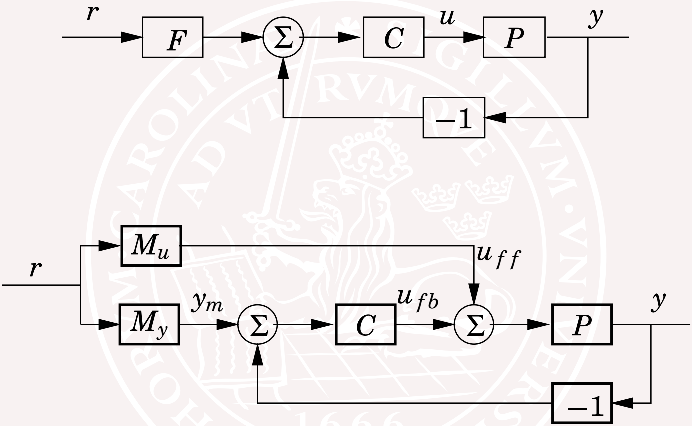
The tuning debate: Should controllers be tuned for set-point response or for load disturbance response?

- Different tuning rules for PID controllers
- Shinskey: Set-point disturbances are less common than load changes.
- Resolved by set-point weighting (poor mans 2DOF)

$$u(t) = k\big(\beta r(t) - y(t)\big) + k_i \int_0^t \big(r(\tau) - y(\tau)\big) d\tau + k_d\Big(\gamma \frac{dr}{dt} - \frac{dy_f}{dt}\Big)$$

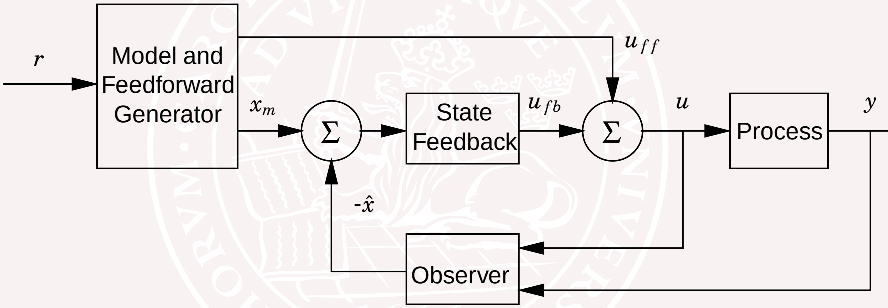
- Tune k, k_i, and k_d for load disturbances, filtering for measurement noise and β, and γ for set-points

Many Versions of 2DOF



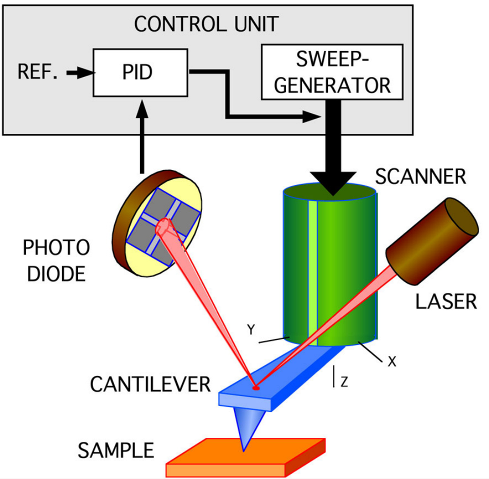
For linear systems all 2DOF configurations have the same properties. For the systems above we have $CF = M_u + CM_y$

A More General Structure

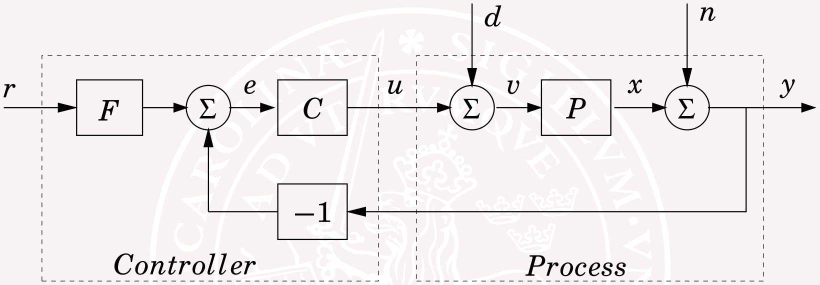


Some Systems only Allow Error Feedback

There are systems where only the error is measured, and the controller then has to be restricted to error feedback.



The Gangs of Four and Six



$$
\begin{aligned}
X &= \frac{P}{1+PC}D - \frac{PC}{1+PC}N + \frac{PCF}{1+PC}R\\
Y &= \frac{P}{1+PC}D + \frac{1}{1+PC}N + \frac{PCF}{1+PC}R\\
U &= -\frac{PC}{1+PC}D - \frac{C}{1+PC}N + \frac{CF}{1+PC}R
\end{aligned}
$$

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 *six* transfer function *The Gang of Six*
- It may be strongly misleading to only show properties of a few systems for example the response of the output to command signals. This is a common omission in papers and books.
- The properties of the different transfer functions can be illustrated by their transient or frequency responses.

A Possible Choice

Six transfer functions are required to show the properties of a basic feedback loop. Four characterize the response to load disturbances and measurement noise, compare $\mathcal{H}_\infty$-theory.

$$
\begin{array}{cc}
\dfrac{PC}{1+PC} & \dfrac{P}{1+PC} \\
\dfrac{C}{1+PC} & \dfrac{1}{1+PC}
\end{array}
$$

The ones in blue also capture robustness. Two more are required to describe the response to set point changes.

$$
\frac{PCF}{1+PC} \qquad \frac{CF}{1+PC}
$$

Design for Performance and Robustness

- Important to consider both performance and robustness
- Many design methods focus on performance but do not include robustness
  - Pole placement
  - LQG
  - Model predictive control
- Bad design choices can easily lead to non-robust systems
- A major advantage of $\mathcal{H}_2$ design is that robustness is explicitely taken into account
- Next we will give a simple illustration

A Simple Pole Placement Design

Consider a stable first order system

$$Y(s) = \frac{b}{s+a} U(s),$$

PI controller with set point weighting

$$U(s) = -k\beta Y(s) + k_i(R(s) - Y(s))$$

The transfer function from reference to output is

$$G_{yr}(s) = \frac{k\beta s + bk_i}{s^2 + (a+bk)s + bk_i}$$

Desired closed loop characteristic polynomial

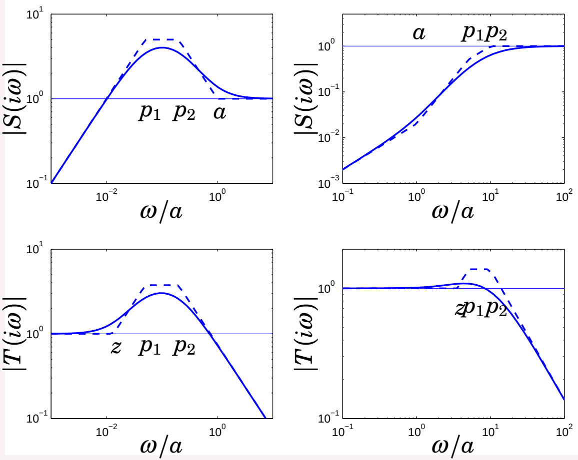
$$(s+p_1)(s+p_2),$$

Controller parameters

$$k = \frac{p_1 + p_2 - a}{b} \qquad k_i = \frac{p_1 p_2}{b}$$

Sensitivity Functions

$$S(s) = \frac{s(s+a)}{(s+p_1)(s+p_2)} \qquad T(s) = \frac{(p_1+p_2-a)s+p_1p_2}{(s+p_1)(s+p_2)}$$



Design Rules

The following rules give designs with low sensitivities

- Determine desired closed loop bandwidth
- Cancel fast stable process poles by controller zeros
- Approximate cancellation is obtained by eliminating poles in model before design
- Cancel slow stable process zeros by controller poles
- Unstable poles and zeros cannot be cancelled and they give rise to fundamental limitations

Limitations due to NMP Dynamics

Process dynamics can impose severe limitations on what can be achieved. Notice that dynamic phenomena do not show up in a traditional static analysis.

- 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:

- Zeros may be changed or removed by moving or adding sensors and actuators. Poles can be influenced only by process redesign.

Summary of Limitations - Part 1

- A RHP zero z gives an upper bound to bandwidth

$$\frac{\omega_{gc}}{z} \le \begin{cases} 0.5 & \text{for } M_s,\, M_t < 2 \\ 0.2 & \text{for } M_s,\, M_t < 1.4. \end{cases}$$

- A time delay T gives an upper bound to bandwidth

$$\omega_{gc} T \le \begin{cases} 0.7 & \text{for } M_s,\, M_t < 2 \\ 0.4 & \text{for } M_s,\, M_t < 1.4. \end{cases}$$

- A RHP pole p gives a lower bound to bandwidth

$$\frac{\omega_{gc}}{p} \ge \begin{cases} 2 & \text{for } M_s,\, M_t < 2 \\ 5 & \text{for } M_s,\, M_t < 1.4. \end{cases}$$

S

Summary of Limitations - Part 2

- RHP poles and zeros must be sufficiently separated

$$\frac{z}{p} \geq \begin{cases} 7 & \text{for } M_s,\, M_t < 2 \\ 14 & \text{for } M_s,\, M_t < 1.4. \end{cases}$$

- RHP poles and zeros must be sufficiently separated

$$\frac{p}{z} \geq \begin{cases} 7 & \text{for } M_s,\, M_t < 2 \\ 14 & \text{for } M_s,\, M_t < 1.4 \end{cases}$$

- The product of a RHP pole and a time delay cannot be too large

$$pT \leq \begin{cases} 0.16 & \text{for } M_s,\, M_t < 2 \\ 0.05 & \text{for } M_s,\, M_t < 1.4. \end{cases}$$

Deriving the Rules

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

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

But $\arg P_{mp}C \approx n\pi/2$, where n is the slope at the crossover frequency. (Exact for Bodes ideal loop transfer function $P_{mp}(s)C(s) = (s/\omega_{gc})^n$). Hence

$$\arg P_{nmp}(i\omega_{gc}) \geq -\pi + \varphi_m - n\frac{\pi}{2}$$

The phase crossover inequality implies that robustness constraints for NMP systems can be expressed in terms of ω_{gc}.

The Crossover Frequency Inequality

The inequality

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

implies that robustness requires that the phase lag of the non-minimum phase component P_{nmp} at the crossover frequency is not too large!

Simple rule of thumb:

- $\varphi_m = 45°, n_{gc} = -1/2 \Rightarrow -\arg P_{nmp}(i\omega_{gc}) \leq \frac{\pi}{2}\ (90°)$
- $\varphi_m = 60°, n_{gc} = -2/3 \Rightarrow -\arg P_{nmp}(i\omega_{gc}) \leq \frac{\pi}{3}\ (60°)$
- $\varphi_m = 45°, n_{gc} = -1 \Rightarrow -\arg P_{nmp}(i\omega_{gc}) \leq \frac{\pi}{4}\ (45°)$

Outline

- Introduction
- Introductory Course
- Feedback Fundamentals
- Advanced Courses
- The Systems View - Laboratories
- Conclusions

Student Categories

- Students who aim for an academic career in control - the future leaders of our field.
- Students who aim for an industrial career in control.
- Students who specialize in other fields such as communication, computing, mathematics, physics, biology, economics, ...

A Challenge

- How to organize a curriculum that focuses on the fundamentals but gives room for broadening in core systems areas like computer science and communication?
- How to make room for a solid mathematics base necessary for research.
- How to bring in enough of process knowledge, sensing, actuation and control practice to make the students useful and aware of engineering issues.

Outline

- Introduction
- Introductory Course
- Feedback Fundamentals
- Advanced Courses
- The Systems View - Laboratories
- Conclusions

The Systems View

Control is a good arena for practicing system engineering

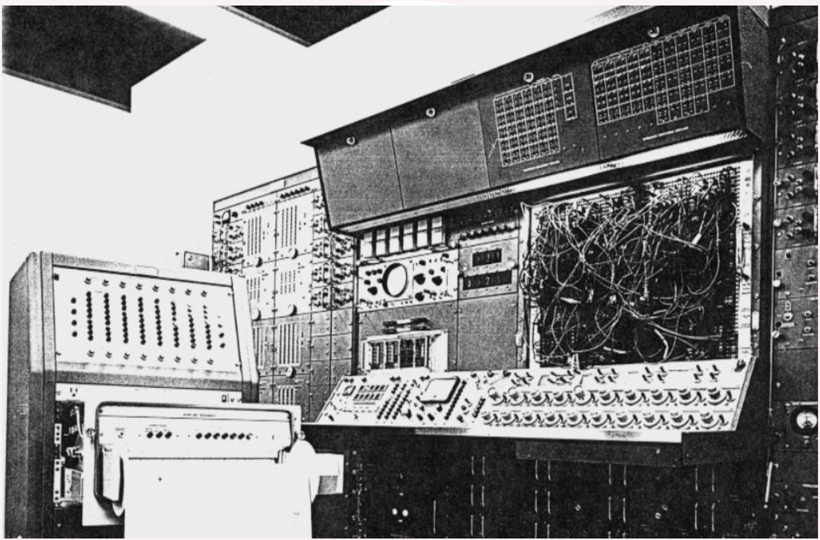
- Requirements
- Specifications
- Modeling
- Design
- Implementation
- Commissioning
- Operation

Essential to provide students with the full picture

Teaching Tools and Laboratories

- Computers change the meaning of solving a problem
  - Analytical solutions sine and Bessel functions
  - Linear systems transfer functions, state space representations and time responses
- Important to balance computing with back of an envelope calculations
- Interactive Learning Tools
  - ICTools, Interactive Learning Modules
- Why laboratories are important?
  - The total systems view
  - Lap top processes
  - Web based
- But always remember to look at results and to think

Return of Interactive Learning



Providing Practical Experience

Labs

- Technology has made labs simple and relatively inexpensive
- The IEEE CSM Special issues.
- Commodity hardware and software
- Lap top processes
- Take-home labs

Industrial projects

T. Hägglund: Case Studies in Control

The Bicycle - A Simple Lab System

Rear steered bike

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

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

RHP zero at V_0/a

Outline

- Introduction
- Introductory Course
- Feedback Fundamentals
- Advanced Courses
- The Systems View - Laboratories
- Conclusions

Summary

- We need to take a serious look at education for our own sake and for others
  - Good introductory course
  - Compactification of available knowledge
- The panel believes that control principles are now a required part of any educated scientist's or engineers background ... If we don't do this it will be done by others and not as well and control may disappear as a discipline.
- The Bologna process gives an excellent opportunity in Europe, let us organize a good collaboration
- Labs are essential